Journal of Biomechanics 84 (2019) 227–233

Contents lists available at ScienceDirect

Journal of Biomechanics

journal homepage: www.elsevier.com/locate/jbiomech
www.JBiomech.com

Subject-specific and group-based running pattern classification using a single wearable sensor

Nizam Uddin Ahamed [a,*], Dylan Kobsar [a], Lauren C. Benson [a], Christian A. Clermont [a], Sean T. Osis [a,c], Reed Ferber [a,b,c]

[a] Faculty of Kinesiology, University of Calgary, 2500 University Drive N.W., Calgary, Alberta T2N IN4, Canada
[b] Faculty of Nursing, University of Calgary, 2500 University Drive N.W., Calgary, Alberta T2N IN4, Canada
[c] Running Injury Clinic, University of Calgary, 2500 University Drive N.W, Calgary, Alberta T2N IN4, Canada

ARTICLE INFO

Article history:
Accepted 2 January 2019

Keywords:
Accelerometer
Biomechanics
Inclination
Gait
Random forest

ABSTRACT

The objective of this study was to determine whether subject-specific or group-based models provided better classification accuracy to identify changes in biomechanical running gait patterns across different inclination conditions. The classification process was based on measurements from a single wearable sensor using a total of 41,780 strides from eleven recreational runners while running in real-world and uncontrolled environment. Biomechanical variables included pelvic drop, ground contact time, braking, vertical oscillation of pelvis, pelvic rotation, and cadence were recorded during running on three inclination grades: downhill, −2° to −7°; level, −0.2° to +0.2°; and uphill, +2° to +7°. An ensemble and non-linear machine learning algorithm, random forest (RF), was used to classify inclination condition and determine the importance of each of the biomechanical variables. Classification accuracy was determined for subject-specific and group-based RF models. The mean classification accuracy of all subject-specific RF models was 86.29%, while group-based classification accuracy was 76.17%. Braking was identified as the most important variable for all the runners using the group-based model and for most of the runners based on a subject-specific models. In addition, individual runners used different strategies across different inclination conditions and the ranked order of variable importance was unique for each runner. These results demonstrate that subject-specific models can better characterize changes in gait biomechanical patterns compared to a more traditional group-based approach.

1. Introduction

Running is one of the most common sports and recreational activities around the world but despite its popularity, each year approximately 50% of runners experience a running-related musculoskeletal injury (Saragiotto et al., 2014a; 2014b; van Gent et al., 2007). The etiology of overuse running injuries is multifactorial and can result from the interaction of many extrinsic factors, such as running surface, running environment, footwear, and training schedule, as well as intrinsic risk factors, such as age, foot strike pattern, and gait biomechanics (Meeuwisse, 1994; Saragiotto et al., 2014a; 2014b; van Gent et al., 2007). Prolonged exposure to these risk factors may lead to musculoskeletal (MSK) overuse running injuries (Hreljac, 2004). However, most research studies to date have been limited to laboratory-based settings, which limits the ability to study the multifactorial nature of MSK injury.

As an alternative to laboratory-based gait research, inertial measurement units (IMUs) are portable devices that can be used to quantify running biomechanical patterns in any setting, including a runner's real-world environment (Benson et al., 2018a; Reenalda et al., 2016; Wouda et al., 2018). Research studies have shown that IMUs can accurately detect changes in gait biomechanics that may be related to MSK injury (Kiernan et al., 2018; Reenalda et al., 2016; Schütte et al., 2018). For example, Reenalda et al. (2016) placed multiple sensors on three runners whilst data were continuously transmitted wirelessly to a receiver, mounted on the handlebar of an accompanying cyclist, over the course of a single marathon race. These authors (Reenalda et al., 2016) used one-way ANOVAs to determine that subject-specific fatigue-related changes in running velocity, stride length, and step frequency occurred over the course of the marathon. However, traditional biomechanics research has generally investigated changes in gait biomechanics using group-based analyses.

* Corresponding author.
E-mail address: nizam.ahamed1@ucalgary.ca (N.U. Ahamed).

https://doi.org/10.1016/j.jbiomech.2019.01.001
0021-9290/

Several studies have identified differences in running patterns based on different age groups, gender and/or injury status (Bus, 2003; Ferber et al., 2003; Phinyomark et al., 2014). In contrast, a recent systematic review suggested that subject-specific models should be used in future IMU-based gait studies outside of the laboratory setting (Benson et al., 2018a). In support of this, other studies have shown that subject-specific models are advantageous as compared to group-based statistical designs (Bates et al., 2004; Bates and Stergiou, 1996; Stacoff et al., 2000a; Stacoff et al., 2000b). However, to our knowledge no study has directly investigated whether group-based or subject-specific models provide deeper insight into the emerging field of outdoor IMU-based biomechanical investigations.

Therefore, the purpose of this study was to determine whether subject-specific or group-based models provide better classification accuracy to detect changes in biomechanical running gait patterns. Considering that it is well established there are subtle differences in joint kinematics and ground reaction forces between level and graded running (Firminger et al., 2018; Vernillo et al., 2017), we compared subject-specific and group-based classification accuracies based on changes in biomechanical running gait patterns during different inclination conditions: downhill, uphill, and level. The following hypotheses were tested: (i) the best classification accuracy would involve subject-specific models compared to group-based models, and (ii) the importance of the variables used to detect changes in running inclination will differ across the runners based on the subject-specific models. For this classification purpose, a nonlinear machine learning algorithm was used, called random forest (RF), which is an ensemble learning method based on classification and regression trees. The RF classifier was also used to measure variable importance in order to rank the biomechanical variables based on their predictive importance in the classification margin (Breiman, 2001; Degenhardt et al., 2017).

2. Methods

2.1. Participants

Eleven recreational runners (10 females: age = 45.6 ± 9.1 years, height = 164.6 ± 7.7 cm, mass = 63.1 ± 10.1 kg; and one male: age = 29 years, height = 170 cm, weight = 75 kg) volunteered to participate in this study. The runners were free of any neuromuscular diseases or MSK injuries and all were registered for a half-marathon training program managed by a local running group. This protocol was approved by the University of Calgary Conjoint Health Research Ethics Board (REB16-2035) and all the runners provided their written informed consent.

2.2. Instrumentation

Six gait variables (pelvic drop (PD, deg), vertical oscillation of the pelvis (VOP, cm), ground contact time (GCT, ms), braking (BR, m/s), pelvic rotation (PR, deg), and cadence (CAD, steps/min)) were recorded using a wearable IMU consisting of a 3-dimensional (3D) accelerometer, magnetometer, and gyroscope (Lumo Run®; Lumo Bodytech Inc., Mountain View, CA, USA; Dimensions: 1.96 in × 1.12 in × 0.39 in; sample frequency 100 Hz), which was attached to the posterior aspect of either the runner's waistband or running belt. The IMU performed onboard calculations of the output variables at five-stride intervals. A GPS watch (Garmin vívoactive® HR; Garmin International Inc., KS, USA) was attached to each runner's preferred wrist and recorded speed, distance, and global positioning data, including latitude, longitude and altitude, every second (Table 1). A Matlab-based (Matlab® 2017a) program was developed to synchronize the IMU and GPS data for further analysis.

2.3. Data collection

Data from each runner was recorded on two different days, separated by one week, and involved two different routes, but the same geographical region (Calgary, AB, Canada). The data obtained with the GPS watch were used to determine the inclination condition data for each of the two days (Table 1) and to define the changes in the inclination grades: downhill, $-2°$ to $-7°$; level, $-0.2°$ to $+0.2°$; and uphill, $+2°$ to $+7°$. It should be noted that as the data were recorded from uncontrolled outside of the laboratory settings, we did not fix any single grade similar to an indoor treadmill grade setting (Vernillo et al., 2017).

Weather data from both days were recorded to ensure the similarities between the days (Ahamed et al., 2018). Specifically, data corresponding to the temperature (degree Celsius), precipitation (mm), and humidity (%) for each run were derived from three different weather stations: temperature +2.1 °C and +0.48 °C; humidity 77% and 79%; and precipitation 0.03 mm and 0.05 mm for day 1 and 2, respectively.

On each day, the total running distance recorded for each participant was approximately 14.5 km. The first 0.5 km of each run were discarded, as this distance was considered the warmup period. To minimize the potential effects of fatigue, we chose to only include data from the first 10 km of each run. For this 10 km section, only data within the pre-defined graded running thresholds (downhill, $-2°$ to $-7°$; level, $-0.2°$ to $+0.2°$; and uphill, $+2°$ to $+7°$) of inclination were used for analysis and as a result, average 4.1 km of data from each runner (approximately 1998.6 strides) were used for Day 1 and an average 3.7 km of data from each runner (approximately 1832.3 strides) were used for Day 2. A summary of these data are provided in Table 1.

2.4. Data analysis

A robust, and non-linear machine learning classifier, called Random Forest (RF), was used to develop the classification models, which measured the accuracy and importance of gait biomechanical variables in classifying runs of differing inclination conditions. The RF classifier has been shown to provide a higher classification accuracy than other existing ML classifiers with a faster computation speed, while facilitating complex interactions among predictor variables and providing information about the importance of each

Table 1
Particulars (mean ± SD) for the two running days.

Running platforms	Altitude (m)		Speed (m/s)		Distance (km)		Number of strides (steps)		Inclination (degree)	
	Day 1	Day 2	Day 1	Day 2	Day 1	Day 2	Day 1	Day 2	Day 1	Day 2
Downhill	1094.1 ± 30.1	1075.0 ± 27.1	2.37 ± 0.1	2.44 ± 0.0	1.3 ± 0.1	1.2 ± 0.1	651.3 ± 64.1	577.2 ± 77.4	(−3.8°) ± 1.7	(−4.5°) ± 1.6
Level	1081.6 ± 23.2	1082.9 ± 25.3	2.27 ± 0.0	2.39 ± 0.1	1.5 ± 0.2	1.3 ± 0.1	735.0 ± 85.3	653.9 ± 112.1	(0.07°) ± 0.05	(0.09°) ± 0.04
Uphill	1078.5 ± 24.1	1067.8 ± 22.3	2.19 ± 0.1	2.26 ± 0.0	1.3 ± 0.1	1.2 ± 0.2	612.3 ± 114.1	601.2 ± 89.3	(+5.9°) ± 2.8	(+6.4°) ± 3.8
Overall (mean ± SD)	1084.7 ± 21.3	1075.3 ± 25.3	2.3 ± 0.10	2.4 ± 0.08	4.1 ± 0.4	3.7 ± 0.2	1998.6 ± 156.1	1832.3 ± 198.2		

predictor variable (Cutler et al., 2007; Strobl et al., 2008; Ziegler and König, 2014) and predictive power (Degenhardt et al., 2017) with the most important variable having a relative importance value of 100%.

Two validation methods were used to ensure that the proposed random forest (RF)-based classification approach was robust and that the data were not overfit (Ferber et al., 2016; Reynard and Terrier, 2017). Method 1 (one-against-another) was a subject-specific approach. Data from one of the runs (day 1 or day 2) was randomly chosen as the training set, and the test set consisted of the data from the other run (day 1 or day 2). As a result, individual training and test sets were generated for each subject. Method 2 was a group-based approach and employed a leave-one-out approach. Using this method, all data from both runs were combined for each runner, the classifiers were trained on all the data except those from one runner, and the data from the remaining runners were then used for testing. These steps were performed 11 times (i.e., once for each runner). Both classification methods were applied using the standalone Python programming language (version 3.6, www.python.org) (Pedregosa et al., 2011).

The developed RF models were trained and cross-validated using the built-in Anaconda distribution (open-source) of Python programming language with notable packages including matplotlib, numpy, scipy, and scikit-learn ("sklearn.ensemble.RandomForestClassifier") (Gupta et al., 2017). The most significant of the selected parameters was *n_estimators*, which refers to the number of trees in the forest. The other parameters were *max_depth*, which refers to the maximum depth of the tree, and *max_features*, which is the number of features selected to split each node (Krauss et al., 2017; Lawson et al., 2017). All these parameters were tuned to improve the overall classification accuracy. Additionally, the RF used a Gini index to calculate the impurity of a node from the CART (classification and regression tree) learning system in order to construct the decision trees (Strobl et al., 2008). These RF trees compute a heuristic for determining the significance of the six biomechanical variables in predicting different inclination conditions and have been commonly used for classification with high-dimensional data in order to rank variable predictors through its built-in variable importance measures (Janitza et al., 2016). In addition, statistical analyses were performed using repeated measures ANOVA ($P < 0.05$) and Cohen's d effects size estimates were calculated for each difference on the outcome measures between each inclination.

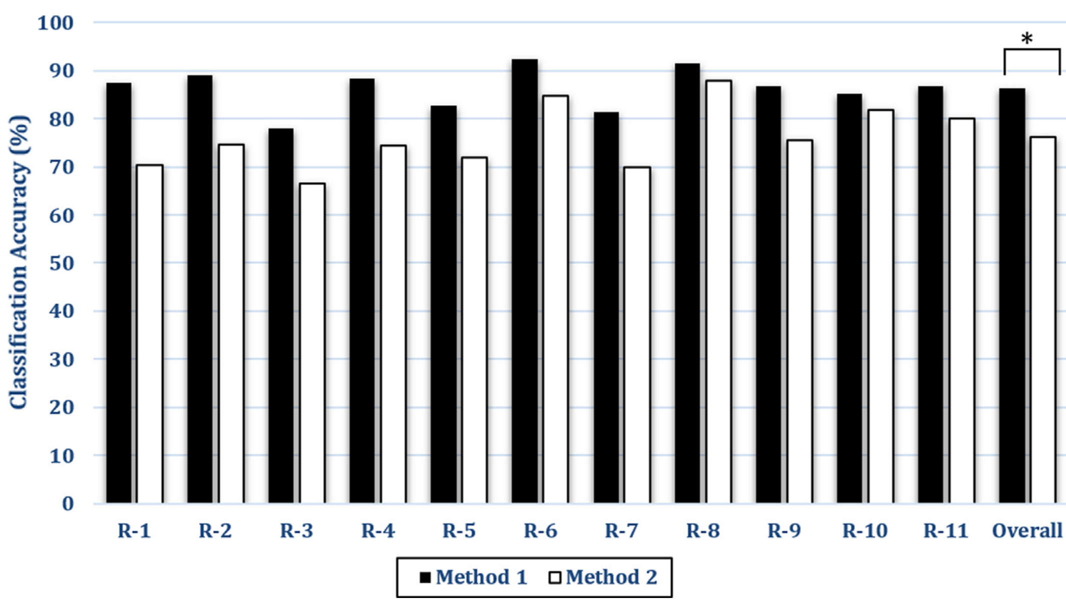

Fig. 1. Classification accuracy generated with Method 1 and Method 2; *: $\boldsymbol{P} < 0.05$.

Table 2
Classification accuracy and variable importance obtained for both methodologies and for each runner.

Runner	Method 1 (Subject-specific) (One-against-another)							Method 2 (Group-based) (Leave-one-out)						
	CA (%)	Variable importance (%)						CA (%)	Variable importance (%)					
		BR	PD	VOP	PR	GCT	CAD		BR	PD	VOP	PR	GCT	CAD
R-1	87.50	19.88	16.32	21.38	*25.62*	13.56	3.25	70.30	*44.54*	15.44	16.99	10.79	7.40	4.85
R-2	89.00	27.40	14.45	*33.02*	3.64	18.28	3.21	74.70	*30.87*	21.54	25.67	9.23	11.37	1.33
R-3	78.00	21.10	*32.51*	6.87	16.70	11.68	11.14	66.50	*43.51*	13.54	19.76	5.32	11.52	6.35
R-4	88.26	13.92	*32.59*	27.62	9.69	3.83	12.34	74.40	*37.30*	21.91	14.61	6.79	10.31	9.08
R-5	82.66	9.28	*36.66*	8.25	17.38	20.99	7.44	72.00	*38.34*	17.41	17.49	5.80	15.14	5.82
R-6	92.45	*40.07*	32.41	8.75	7.37	6.63	4.76	84.70	*32.38*	15.35	17.46	12.33	11.57	10.91
R-7	81.34	26.24	*27.63*	23.57	12.69	7.34	2.53	70.00	*34.22*	19.51	11.23	8.70	14.98	11.36
R-8	91.39	*35.08*	14.90	24.89	10.09	4.53	10.51	88.00	*25.36*	18.14	19.56	11.00	14.55	11.40
R-9	86.68	*50.12*	10.88	10.28	13.86	5.35	9.51	75.50	*41.85*	17.23	15.37	8.74	10.63	6.17
R-10	85.23	26.39	13.35	*28.81*	19.65	4.04	7.77	81.80	*30.07*	15.73	17.50	9.57	13.32	13.81
R-11	86.73	*30.63*	6.90	29.92	3.61	3.59	25.34	80.00	*50.31*	15.07	13.50	8.06	8.61	4.46
Overall	86.29	*27.28*	21.69	20.31	12.75	9.07	8.89	76.17	*37.16*	17.35	17.19	8.76	11.76	7.78

CA: classification accuracy; The most important variable identified by each method and for each runner is highlighted in bold and italic font.
BR: braking, VOP: vertical oscillation of the pelvis, PD: pelvic drop, PR: pelvic rotation, GCT: ground contact time, CAD: cadence.

Table 3
Descriptive statistics for both runs ran by each runner and summary of gait biomechanical data obtained from the three running platforms.

Runner	Braking (m/s)			Vertical oscillation of pelvis (cm)			Pelvic drop (deg)			Pelvic rotation (deg)			Ground contact time (ms)			Cadence (steps/min)		
	DH	Level	UH	DH	Level	UH	DH	Level	UH	DH	Level	UH	DH	Level	UH	DH	Level	UH
R-1	0.34	0.27	0.25	6.35	5.58	5.41	9.29	7.89	8.35	13.84	14.48	13.19	255.04	259.51	253.82	174.39	175.40	176.30
R-2	0.49	0.38	0.38	9.67	9.15	8.46	10.75	10.02	9.87	14.85	15.78	14.82	270.39	257.73	269.66	162.92	167.56	165.47
R-3	0.22	0.22	0.25	5.19	5.09	5.03	8.11	8.95	8.39	14.90	16.79	17.15	262.38	267.98	263.64	181.43	180.46	181.44
R-4	0.32	0.35	0.32	9.91	8.63	8.42	10.97	10.82	13.09	11.52	10.59	9.36	310.26	299.49	320.97	136.67	144.98	152.07
R-5	0.37	0.32	0.31	6.65	6.51	6.64	6.57	8.10	9.52	10.80	13.68	13.76	254.47	258.18	247.84	173.10	173.26	172.65
R-6	0.41	0.31	0.26	7.05	7.24	6.22	13.83	14.72	16.24	9.81	7.60	7.56	224.60	234.46	239.45	186.71	182.84	185.57
R-7	0.30	0.32	0.29	12.58	12.39	11.48	13.75	11.86	11.00	16.69	15.26	17.75	315.26	310.01	309.50	148.39	149.09	150.60
R-8	0.40	0.35	0.25	5.84	7.01	6.54	8.67	9.58	10.71	7.63	10.10	10.31	260.07	258.79	269.96	176.27	176.08	171.82
R-9	0.27	0.25	0.22	9.10	9.32	8.68	7.34	7.80	7.13	20.31	20.83	18.20	274.86	276.75	277.71	164.00	161.84	163.69
R-10	0.30	0.30	0.22	7.61	8.24	7.03	6.35	7.33	7.15	8.96	10.07	9.40	273.67	269.92	271.30	167.05	168.13	165.56
R-11	0.39	0.34	0.27	7.17	6.39	6.12	8.97	8.01	8.38	17.60	17.70	17.63	258.21	245.34	251.47	179.14	185.63	181.31
Mean ± SD	0.35 ± 0.1	0.31 ± 0.1	0.27 ± 0.1	8.11 ± 2.2	7.78 ± 2.1	7.27 ± 1.9	9.51 ± 2.6	9.55 ± 2.2	9.98 ± 2.7	13.35 ± 3.9	13.90 ± 3.9	13.56 ± 3.9	269.02 ± 25.5	267.11 ± 22.0	270.48 ± 25.01	168.19 ± 14.8	169.57 ± 13.2	169.68 ± 11.5
P	DL = 0.03*; LU = 0.01*; UD = 0.00*			DL = 0.72; LU = 0.00*; UD = 0.01*			DL = 0.90; LU = 0.20; UD = 0.39			DL = 0.28; LU = 0.40 UD = 0.73			DL = 0.44; LU = 0.25; UD = 0.53			DL = 0.24; LU = 0.92; UD = 0.35		
ES	DL = 0.77; LU = 0.96; UD = −1.2			DL = 0.20; LU = 1.16 UD = −1.01			DL = −0.04; LU = −0.42 UD = 0.27			DL = −0.34; LU = 0.26; UD = 0.11			DL = 0.24; LU = −0.37; UD = 0.20			DL = −0.38; LU = −0.03; UD = 0.29		

DH: Downhill; **UH**: Uphill; **DL**: Downhill vs. Level; **LU**: Level vs. Uphill; **UD**: Uphill vs. Downhill; *: ***P*** < 0.05; **ES**: Effect size.

3. Results

The mean classification accuracy of the subject-specific RF models (Method 1) was 86.29%. All the runners had individual classification accuracy higher than 80% with the exception of R-3, which was 78%. In contrast, the mean classification accuracy obtained with group-based approach (Method 2) was significantly lower at 76.17% ($P = 0.01$). In the group-based models, only four runners had individual classification accuracy higher than 80% and the remaining runners exhibited classification accuracy between 66.5% and 80% (Fig. 1).

Similarities in the gait variables were observed for two runners (R-6 and R-8) that yielded the highest classification accuracy with both methods (92.45% and 91.39% in Method 1 and 84.7% and 88% in Method 2, respectively). For example, the RF variable importance measure showed that BR was the important variable for these two runners for both methods (Table 2). In addition, both runners generated greater BR and CAD values for downhill running, as compared to level and uphill, followed by VOP for level running as compared to uphill and downhill, and also greater PD and GCT for uphill running conditions, as compared to downhill and level (Table 3).

BR was ranked highest for both methods based on the average values across the eleven runners (Table 2). Although, BR in Method 2 was consistently ranked highest across all runners, different trends were observed for Method 1. For example, BR was ranked highest in four runners, second for four runners and third (R-4 and 1) and fourth (R-5) for the remaining three runners. PD was the second most important variable in both methods according to the overall average values, but its importance of ranking differed according to the individual results for both methods. For example, with Method 1, PD was the most important for four of the runners (R-3, 4, 5 and 7) but was ranked second, third or fourth most important for the remaining runners. Similarly, with Method 2, this variable was second most important for four of the runners and the third most important for the remaining seven runners. Similar trends were observed in VOP. While it was the third most important variable identified using both methods according to the overall average results, this result was not same for the individual runners. Specifically, with Method 1, VOP was identified as the most important for two runners (R-2 and 10), second most important for four runners and third most important for - two runners (R-6 and 7). In contrast, with Method 2, VOP was identified as the second and third most important variable for seven and three runners, respectively (Table 2).

Fig. 2 shows an example of the subject-specific approach (Method 1) and depicts the difference in BR between two runners using the two methods. Specifically, R-6 (white bar) yielded the highest classification accuracy (92.45%) in Method 1, and BR was the most important variable and showed maximal values in the downhill and level running compared with the uphill running platforms. In contrast, R-3 (black bar) yielded the lowest accuracy (78%) with Method 1, and BR was not important and generated the lowest values in the downhill and level platform. Finally, the statistical results for the six gait biomechanical activities of running according to subject-specific and group-based are shown in Table 3.

4. Discussion

The objective of this study was to determine whether subject-specific or group-based models provide better classification accuracy based on changes in biomechanical running gait patterns across different inclination conditions. We utilised a machine learning algorithm, called Random Forest, to detect changes in tri-

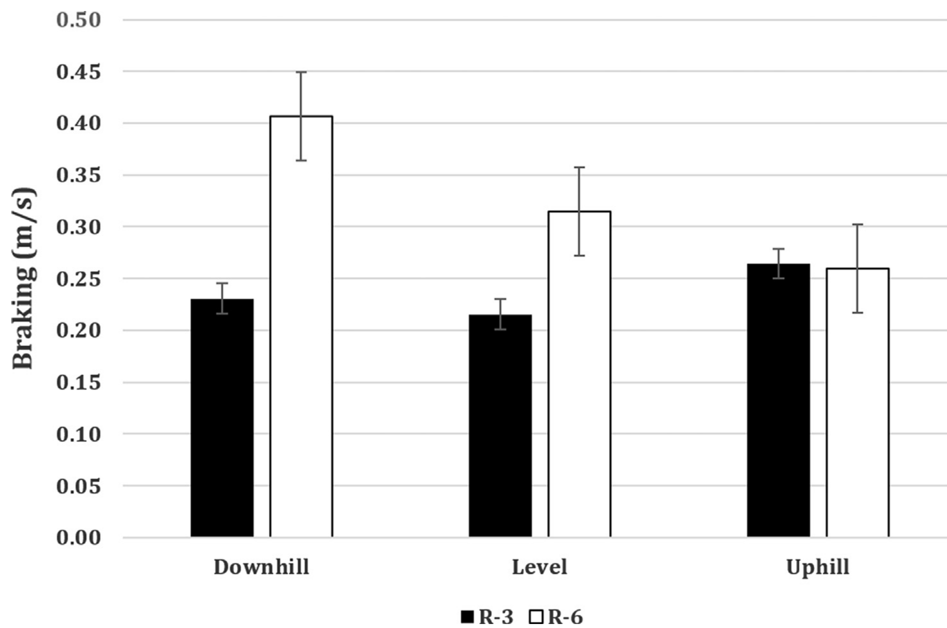

Fig. 2. An example of the differing effects of braking for the three inclination conditions for two runners (R-6; white and R-3; black), who exhibited the highest (92.45%) and lowest (78%) classification accuracies, respectively, using Method 1 (subject-specific).

axial accelerometer, gyroscope, and magnetometer data from an IMU in order to classify three running inclination conditions (level, uphill and downhill). In support of our hypothesis, the group-based classification model (Method 2) achieved a mean classification rate (76.17%) that was lower than a subject-specific approach, which demonstrated a mean classification rate of 86.29%. These results support the case for a subject-specific experimental design when analyzing movement patterns based on data collected outdoors using an IMU (Ahamed et al., 2018; Bates and Stergiou, 1996; Lenaerts et al., 2008; Luu et al., 2014). Further, the results of the current study suggest that the strategies used to accomplish the tasks of uphill, downhill, or level running are different for individual runners.

In support of a subject-specific approach to understanding changes in running gait biomechanics (Ahamed et al., 2018), the current study also revealed that individual runners use different strategies across different inclination conditions and the ranked order of variable importance was unique for each runner. Moreover, no runner's ranked order of variable importance matched the overall ranked order of variable importance and some runners even appeared to have more distinct differences between running patterns on different inclined surfaces than others. For example, the two runners with the greatest classification accuracy in the group-based model (R-6 and 8) also had the greatest classification accuracy in the subject-specific model. Likewise, the lowest classification accuracy for each model occurred for the same pair of runners (R-3 and 7). Therefore, future analyses of running patterns under different external conditions should consider subject-specific analysis models.

For the group-based model, we were able to classify the running patterns of all the runners with 76.1% accuracy. These results are consistent with previous research from Schuldhaus et al. (Schuldhaus et al., 2012) who also classified uphill, downhill, and level running conditions with 76.5–81.2% classification accuracy, depending on the classifier used, with time- and frequency-based features extracted from foot-mounted 3D accelerometers and gyroscopes. The current study significantly builds upon this research by collecting data under real-world environments as compared to the repeated, predefined 2.1-km controlled loop used by these authors (Schuldhaus et al., 2012). Moreover, while foot-mounted wearable technology can be relatively unobtrusive and implemented into running footwear, the computational load for calculating and extracting the 66 features used by Schuldhaus et al. (Schuldhaus et al., 2012) may be too complex for the current capabilities of commercial-based wearable technology (Benson et al., 2018b). The time- and frequency-based features selected in the classification models used by these authors were also not specified, and it is therefore unclear whether these features have meaning for clinical and performance-based applications. Therefore, to provide insight into this knowledge gap and open new research directions, the current study developed and evaluated subject-specific and group-based methods, using an RF classifier using data from a single IMU, and achieved excellent classification accuracy results using only six gait variables.

With respect to the RF variable importance measurement and statistical analysis, we demonstrated that braking was the most important variable for all the runners in group-based model and most of the runners based on a subject-specific model. Additionally, inferential statistics further demonstrated that braking was significantly different between the three inclination conditions: greatest in downhill running, followed by level then uphill running. These findings are in agreement with previous literature, in which downhill running resulted higher impact/braking forces compared to level or uphill inclination (Vernillo et al., 2017). Additionally, Telhan et al. (Telhan et al., 2010) reported greater impact peak forces during downhill running, and Gottschall and Kram (Gottschall and Kram, 2005) reported higher impact forces, loading rates, and anteroposterior braking impulses during downhill running compared to level running at controlled speeds. Thus, it is evident that the braking force was greater during downhill running compared to level and uphill inclinations, but the current study is the first, to our knowledge, to find this result in real-world and uncontrolled outdoor setting using a single IMU-based wearable sensor.

It is our hope that these results are useful to clinicians and coaches alike in order to assist the former in understanding and preventing running-related injuries and for the latter to help improve athletic performance. For example, the subject-specific models utilised in the current study can enable a coach to reduce an athlete's braking force during downhill running in order to optimize running velocity. Moreover, it is possible for a clinician to optimize an injured athlete's rehabilitation by altering and/or gradually changing braking force over time.

While the results of the current study are encouraging, it is important to recognize the limitations of this work. For example, there were a limited number of subjects assessed in the current analysis. However, while limited sample sizes can lead to a risk of overfitting, we are confident that the results presented here are not overfit for several reasons. First, each subject-specific model involved a dataset of approximately 4,000 observations and these models were trained on one run and tested on a completely different run, which the model was blinded. This is a critical point, as the success of this method demonstrates that the classification of running patterns based on changes in gait patterns was not overfit to a single dataset but remains successful amid the potential presence of various sources of between-day variability. Second, the use of a leave-one-out cross validation method for the group-specific model ensured that all models generated were based on data blinded to the training sets. Nevertheless, we also acknowledge that these results still only represent the gait patterns of eleven runners wherein data were only collected from a single wearable sensor mounted near the body's center of mass. Therefore, there is a need to further evaluate a larger and more diverse cohort of runners. Moreover, it is recommended that investigations using a larger sample size are necessary to determine if homogenous sub-groups, or clusters, will form as a result of consistent within-group biomechanical changes due to changes in the real-world environment (Phinyomark et al., 2015).

Another potential limitation is that a limited number of spatiotemporal and biomechanical variables were obtained from a commercially available wearable sensor device and used for the current classification problem. While we acknowledge that additional IMU sensors may yield different results, it is important to consider that the current study involved collecting data from runners over multiple days and in real-world environments. Thus, we had to strike a balance between a runner's comfort, and what they will tolerate in terms of attached devices, and the research data to be collected. Considering that the use of wearable sensors for analysing gait biomechanics in real-world environments is in very early stages of development (Benson et al., 2018a), we encourage future research to collect as much gait-related data as can be tolerated by the research subjects and dependent on comfort, distance and time of each running event, and the research question. Furthermore, while it is likely that additional or more complex variables from one or more wearable sensors could improve the classification accuracy of the current study, we posit that the simplicity and translatability to the current market of wearable sensors is a significant advantage that should not be overlooked. Regardless, future research should include a broader range of variables, and possibly more wearable sensor devices, in order to gain a deeper understanding for subject-specific changes in gait patterns during out-of-laboratory data collections.

Lastly, considering that the present experiment involved an IMU that was clipped to the posterior aspect of either the runner's waistband or running belt and considering that data collection occurred in an outdoor and uncontrolled environment, there was the possibility that the IMU sensor shifted slightly during the run trials. However, even if there was a shift in the position of the IMU, there is an equal chance of this occurring to any participant and at any time, regardless of whether they were running uphill, downhill, or on level ground. Therefore, we do not feel that the results of the study would be significantly influenced by such a minor occurrence.

5. Conclusion

Our results support the use of a subject-specific machine learning approach for determining changes in running gait patterns across three different inclination conditions. As well, we demonstrated that braking was the most important variable in the classification model but that individual runners use different strategies across different inclination conditions and the ranked order of variable importance was unique for each runner. These results could help improve our understanding of running-related MSK injuries and possibly provide valuable real-time information to the individual runner, coaches, and/or sport scientists for future research involving IMU data obtained in real-world settings.

Conflict of interest statement

We wish to confirm that there are no known conflicts of interest associated with this publication and there has been no significant financial support for this work that could have influenced its outcome.

Acknowledgments

This study was partially funded by the Natural Sciences and Engineering Research Council of Canada (NSERC), Canada; Idea-2-Innovation grant, Canada; a University of Calgary Eyes High Postdoctoral Research award, Canada and a Strategic Research Grant from the Vice-President (Research) at the University of Calgary, Canada.

References

Ahamed, N.U., Kobsar, D., Benson, L., Clermont, C., Kohrs, R., Osis, S.T., Ferber, R., 2018. Using wearable sensors to classify subject-specific running biomechanical gait patterns based on changes in environmental weather conditions. PLoS One 13, e0203839.

Bates, B., James, C., Dufek, J., 2004. Single-subject analysis. In: Stergiou, N. (Ed.), Innovative analyses of human movement, pp. 3–28.

Bates, B.T., Stergiou, N., 1996. Performance accomodation to midsole hardness during running. J. Human Movement Studies 31, 189–210.

Benson, L.C., Clermont, C.A., Bošnjak, E., Ferber, R., 2018a. The use of wearable devices for walking and running gait analysis outside of the lab: a systematic review. Gait Posture 63, 124–138.

Benson, L.C., Clermont, C.A., Osis, S.T., Kobsar, D., Ferber, R., 2018b. Classifying running speed conditions using a single wearable sensor: optimal segmentation and feature extraction methods. J. Biomech. 71, 94–99.

Breiman, L., 2001. Random forests. Machine learning 45, 5–32.

Bus, S.A., 2003. Ground reaction forces and kinematics in distance running in older-aged men. Med. Sci. Sports Exerc. 35, 1167–1175.

Cutler, D.R., Edwards, T.C., Beard, K.H., Cutler, A., Hess, K.T., Gibson, J., Lawler, J.J., 2007. Random forests for classification in ecology. Ecology 88, 2783–2792.

Degenhardt, F., Seifert, S., Szymczak, S., 2017. Evaluation of variable selection methods for random forests and omics data sets. Briefings Bioinf.

Ferber, R., Davis, I.M., Williams iii, D.S., 2003. Gender differences in lower extremity mechanics during running. Clin. Biomech. 18, 350–357.

Ferber, R., Osis, S.T., Hicks, J.L., Delp, S.L., 2016. Gait biomechanics in the era of data science. J. Biomech. 49, 3759–3761.

Firminger, C.R., Vernillo, G., Savoldelli, A., Stefanyshyn, D.J., Millet, G.Y., Edwards, W.B., 2018. Joint kinematics and ground reaction forces in overground versus treadmill graded running. Gait posture 63, 109–113.

Gottschall, J.S., Kram, R., 2005. Ground reaction forces during downhill and uphill running. J. Biomech. 38, 445–452.

Gupta, V.P., Garton, A.L., Sisti, J.A., Christophe, B.R., Lord, A.S., Lewis, A.K., Frey, H.-P., Claassen, J., Connolly, E.S., 2017. Prognosticating functional outcome after intracerebral hemorrhage: the ICHOP score. World Neurosurg. 101, 577–583.

Hreljac, A., 2004. Impact and overuse injuries in runners. Med. Sci. Sports Exerc. 36, 845–849.

Janitza, S., Tutz, G., Boulesteix, A.-L., 2016. Random forest for ordinal responses: prediction and variable selection. Comput. Stat. Data Anal. 96, 57–73.

Kiernan, D., Hawkins, D.A., Manoukian, M.A.C., McKallip, M., Oelsner, L., Caskey, C.F., Coolbaugh, C.L., 2018. Accelerometer-based prediction of running injury in National Collegiate Athletic Association track athletes. J. Biomech. 73, 201–209.

Krauss, C., Do, X.A., Huck, N., 2017. Deep neural networks, gradient-boosted trees, random forests: Statistical arbitrage on the S&P 500. Eur. J. Oper. Res. 259, 689–702.

Lawson, E., Smith, D., Sofge, D., Elmore, P., Petry, F., 2017. Decision forests for machine learning classification of large, noisy seafloor feature sets. Comput. Geosci. 99, 116–124.
Lenaerts, G., De Groote, F., Demeulenaere, B., Mulier, M., Van der Perre, G., Spaepen, A., Jonkers, I., 2008. Subject-specific hip geometry affects predicted hip joint contact forces during gait. J. Biomech. 41, 1243–1252.
Luu, T.P., Low, K., Qu, X., Lim, H., Hoon, K., 2014. An individual-specific gait pattern prediction model based on generalized regression neural networks. Gait posture 39, 443–448.
Meeuwisse, W.H., 1994. Assessing causation in sport injury: a multifactorial model. Clin. J. Sport Med. 4, 166–170.
Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., 2011. Scikit-learn: machine learning in Python. J. Mach. Learning Res. 12, 2825–2830.
Phinyomark, A., Hettinga, B.A., Osis, S.T., Ferber, R., 2014. Gender and age-related differences in bilateral lower extremity mechanics during treadmill running. PLoS ONE 9, e105246.
Phinyomark, A., Osis, S., Hettinga, B., Leigh, R., Ferber, R., 2015. Gender differences in gait kinematics in runners with iliotibial band syndrome. Scand. J. Med. Sci. Sports 25, 744–753.
Reenalda, J., Maartens, E., Homan, L., Buurke, J.J., 2016. Continuous three dimensional analysis of running mechanics during a marathon by means of inertial magnetic measurement units to objectify changes in running mechanics. J. Biomech. 49, 3362–3367.
Reynard, F., Terrier, P., 2017. Determinants of gait stability while walking on a treadmill: a machine learning approach. J. Biomech. 65, 212–215.
Saragiotto, B.T., Di Pierro, C., Lopes, A.D., 2014a. Risk factors and injury prevention in elite athletes: a descriptive study of the opinions of physical therapists, doctors and trainers. Braz. J. Phys. Therapy 18, 137–143.
Saragiotto, B.T., Yamato, T.P., Junior, L.C.H., Rainbow, M.J., Davis, I.S., Lopes, A.D., 2014b. What are the main risk factors for running-related injuries? Sports Med. 44, 1153–1163.
Schuldhaus, D., Kugler, P., Jensen, U., Eskofier, B., Schlarb, H., Leible, M., 2012. Classification of surfaces and inclinations during outdoor running using shoe-mounted inertial sensors. In: 21st IEEE International Conference on Pattern Recognition (ICPR), Tsukuba City Japan, pp. 2258–2261.
Schütte, K.H., Seerden, S., Venter, R., Vanwanseele, B., 2018. Influence of outdoor running fatigue and medial tibial stress syndrome on accelerometer-based loading and stability. Gait posture 59, 222–228.
Stacoff, A., Nigg, B.M., Reinschmidt, C., Van Den Bogert, A.J., Lundberg, A., 2000a. Tibiocalcaneal kinematics of barefoot versus shod running. J. Biomech. 33, 1387–1395.
Stacoff, A., Reinschmidt, C., Nigg, B., van den Bogert, A.J., Lundberg, A., Denoth, J., Stüssi, E., 2000b. Effects of foot orthoses on skeletal motion during running. Clin. Biomech. 15, 54–64.
Strobl, C., Boulesteix, A.-L., Kneib, T., Augustin, T., Zeileis, A., 2008. Conditional variable importance for random forests. BMC Bioinf. 9, 307.
Telhan, G., Franz, J.R., Dicharry, J., Wilder, R.P., Riley, P.O., Kerrigan, D.C., 2010. Lower limb joint kinetics during moderately sloped running. J. Athlet. Train. 45, 16–21.
van Gent, B.R., Siem, D.D., van Middelkoop, M., van Os, T.A., Bierma-Zeinstra, S.S., Koes, B.B., 2007. Incidence and determinants of lower extremity running injuries in long distance runners: a systematic review. Br. J. Sports Med. 41, 469–480.
Vernillo, G., Giandolini, M., Edwards, W.B., Morin, J.-B., Samozino, P., Horvais, N., Millet, G.Y., 2017. Biomechanics and physiology of uphill and downhill running. Sports Med. 47, 615–629.
Wouda, F.J., Giuberti, M., Bellusci, G., Maartens, E., Reenalda, J., Van Beijnum, B.-J.F., Veltink, P.H., 2018. Estimation of vertical ground reaction forces and sagittal knee kinematics during running using three inertial sensors. Front. Physiol. 9, 218.
Ziegler, A., König, I.R., 2014. Mining data with random forests: current options for real-world applications. Wiley Interdisciplin. Rev.: Data Mining Knowledge Discov. 4, 55–63.