Technology in Strength and Conditioning: Assessing Bodyweight Squat Technique With Wearable Sensors

Martin A. O'Reilly,[1,2] Darragh F. Whelan,[1,2] Tomas E. Ward,[3] Eamonn Delahunt,[2] and Brian M. Caulfield[1,2]

[1]*Insight Center for Data Analytics, University College Dublin, Dublin, Republic of Ireland;* [2]*School of Public Health, Physiotherapy and Sports Science, University College Dublin, Dublin, Republic of Ireland; and* [3]*Insight Center for Data Analytics, Maynooth University, Maynooth, Republic of Ireland*

Abstract

O'Reilly, MA, Whelan, DF, Ward, TE, Delahunt, E, and Caulfield, BM. Technology in strength and conditioning: assessing bodyweight squat technique with wearable sensors. *J Strength Cond Res* 31(8): 2303–2312, 2017—Strength and conditioning (S&C) coaches offer expert guidance to help those they work with achieve their personal fitness goals. However, it is not always practical to operate under the direct supervision of an S&C coach and consequently individuals are often left training without expert oversight. Recent developments in inertial measurement units (IMUs) and mobile computing platforms have allowed for the possibility of unobtrusive motion tracking systems and the provision of real-time individualized feedback regarding exercise performance. These systems could enable S&C coaches to remotely monitor sessions and help individuals record their workout performance. One aspect of such technologies is the ability to assess exercise technique and detect common deviations from acceptable exercise form. In this study, we investigate this ability in the context of a bodyweight (BW) squat exercise. Inertial measurement units were positioned on the lumbar spine, thighs, and shanks of 77 healthy participants. Participants completed repetitions of BW squats with acceptable form and 5 common deviations from acceptable BW squatting technique. Descriptive features were extracted from the IMU signals for each BW squat repetition, and these were used to train a technique classifier. Acceptable or aberrant BW squat technique can be detected with 98% accuracy, 96% sensitivity, and 99% specificity when using features derived from all 5 IMUs. A single IMU system can also distinguish between acceptable and aberrant BW squat biomechanics with excellent accuracy, sensitivity, and specificity. Detecting exact deviations from acceptable BW squatting technique can be achieved with 80% accuracy using a 5 IMU system and 72% accuracy when using a single IMU positioned on the right shank. These results suggest that IMU-based systems can distinguish between acceptable and aberrant BW squat technique with excellent accuracy with a single IMU system. Identification of exact deviations is also possible but multi-IMU systems outperform single IMU systems.

Key Words inertial measurement units, compound exercise, exercise biofeedback, exercise classification

Introduction

In resistance training, the bodyweight (BW) squat is a compound full-body exercise, whose constituent movements are integral to activities of daily living. It commonly features as a fundamental exercise in resistance training and rehabilitation programs. Furthermore, it is incorporated into musculoskeletal injury risk identification protocols (8). The National Strength and Conditioning Association (NSCA) have outlined a number of common deviations from acceptable squat technique (2). These aberrant biomechanics have been shown to increase stress on the joints of the lower extremity (11), potentiating the risk of injury. Thus, the reliable assessment of BW squat biomechanics is necessary to mitigate injury risk.

The assessment of BW squat technique is typically undertaken using 1 of the 2 distinct methods: (a) 3-D motion capture; (b) subjective visual analysis. Both these have a number of limitations. 3-D motion capture systems are expensive, and the application of skin-mounted markers may hinder normal movement (1,3). Furthermore, data processing can be time intensive and specific expertise is often required to interpret the processed data and to make recommendations on the observed results. Therefore, these systems are not frequently used to assess BW squat technique beyond the research laboratory (5). In clinical- and gym-based settings, subjective visual assessment is typically used to assess BW squat technique. This subjective visual

Address correspondence to Martin A. O'Reilly, martin.oreilly@insight-centre.org.

31(8)/2303–2312

Journal of Strength and Conditioning Research

assessment of human biomechanics is not always reliable even among experts as the need to visually assess numerous constituent components simultaneously is challenging (29).

Wearable inertial measurement units (IMUs) may offer the potential to bridge the gap between laboratory and day-to-day "real-world" acquisition and assessment of human movement. These IMUs are small, inexpensive sensors that consist of accelerometers, gyroscopes, and magnetometers. They are able to acquire data pertaining to the inertial motion and 3D orientation of individual limb segments (15,21). Self-contained, wireless IMU devices are easy to set-up and allow for the acquisition of human movement data in unconstrained environments (16). In this article, the term IMU system will be used to describe the IMU sensors, the sensor signals, the associated signal processing applied to them, and the output of the exercise classification algorithms.

Inertial measurement unit systems can robustly track the variety of postures and environmental complexities associated with training, unlike camera-based systems, which are hampered by location, occlusion, and lighting issues in such settings (19). Inertial measurement units have also been shown to be as effective as marker-based systems at measuring joint angles (5,14,25). There are many commercially available examples of IMU systems that monitor physical activity (e.g., Jawbone (San Francisco, CA, USA) and Fitbit (San Francisco, CA, USA)). However, using IMU systems to assess gym-based exercises is less common. Researchers have demonstrated the ability of IMU-based systems to distinguish different gym-based exercises and count repetitions of these exercises with moderate to good levels of accuracy (7,20,22–24). Although these aforementioned systems detail information on the number of repetitions performed, they do not provide instruction on exercise technique and quality of performance. A holistic exercise tracking system should not only recognize the exercise completed, but should also provide technique feedback. Furthermore, in order for IMU systems to be used as an objective method to acquire and assess human movement data as part of a musculoskeletal injury risk-screening protocol, they need to be able to identify aberrant movement patterns and provide easily interpretable data to clinicians and coaches who use them.

A growing body of scientific literature has investigated the utility of IMU systems to assess exercise technique. Taylor et al. (26) used a 5 IMU system to categorize 5 technique deviations during performance of the standing hamstring curl exercise and 4 technique deviations during performance of the straight-leg raise exercise. They were able to identify these deviations with 80% accuracy, 75% sensitivity, and 90% specificity. Pernek et al. (23) also used a 5 IMU system to monitor exercise intensity in 6 dumbbell upper limb exercises (biceps curl, single arm triceps extension, front vertical lift, lateral vertical lift, bent-over row, and military press), demonstrating an average error of just 6% for intensity prediction. Melzi et al. (17) used a wireless body area network of accelerometers to assess performance of the biceps curl. However, their proposed approach was very exercise specific and difficult to transfer to other regularly performed gym-based exercises (23). Research undertaken by Velloso et al. (28) examined the ability of a 4 IMU system to identify 4 deviations from acceptable technique during a unilateral dumbbell bicep curl. Their reported overall accuracy ranged from 74 to 86%. Although these results are encouraging, the multiple IMU systems detailed above are expensive and their use may prove impractical due to the increased risk of placement error and comfort issues (5). In addition, the use of multiple IMUs leads to increased overall power requirements and connectivity complexity with respect to the hosting device (e.g., smartphone, tablet).

A reduced IMU set-up is more desirable for daily environment applications (4). Bonnet et al. (5) reported that a single IMU mounted on the lower back could measure ankle, knee, and hip joint angles in the sagittal plane during performance of the BW squat exercise, with a maximal error of 3.5° compared with a motion capture system. This indicates that relatively accurate quantification of sagittal plane lower limb joint movement can be achieved using a single sensor unit system. Giggins et al. (10) demonstrated an overall accuracy of 79–83% using a single IMU on either the foot, shin, or thigh to identify deviations in the performance of 7 exercises (heel slide, hip abduction, hip flexion, hip extension, knee extension, inner range quads, and straight-leg raise). Pernek et al. (22) analyzed the ability of an accelerometer contained within a smartphone to assess exercise performance during 9 free-weight and machine resistance exercises. They assessed movement quality based on the speed of exercise performance and reported a temporal error of 11% for individual repetition duration.

In summary, the BW squat is a compound full-body exercise that is typically a constituent component of resistance training, rehabilitation programs, and musculoskeletal injury risk-screening protocols. Incorrect BW squat technique can heighten the risk of injury. Traditionally, exercise technique has been evaluated using expensive motion capture systems or via subjective visual inspection from trained professionals. Inertial measurement unit systems offer an opportunity to provide low-cost exercise technique assessment. However, to date, no research has evaluated the capability of IMU systems to assess BW squat technique. The overall aim of this research was to evaluate the ability of a number of IMU system configurations to distinguish between acceptable and aberrant BW squat biomechanics. Aberrant BW squatting biomechanics were defined as BW squats that contained one of the common lower limb deviations identified by the NSCA (2). These deviations are outlined in Table 1.

Methods

Experimental Approach to the Problem

This study used an opportunistic approach to the development of a wearable sensor system for automatically assessing

Table 1. List and description of squat exercise deviations used in this study and the number of repetitions (*n*) extracted of each class.

Deviation	Description	Total reps (*n*)
ACC	Acceptable squat technique	722
KVL	Knees coming together during downward phase	221
KVR	Knees coming apart during downward phase	222
KTF	Knees ahead of toes during downward phase	225
HE	Heels raising off the ground during squat exercise	228
BO	Excessive flexion of hip and torso during squat exercise	231

BW squat technique. Two distinct types of technique assessment were evaluated. "Binary classification" describes systems which classify one's performance as "acceptable" or "aberrant" without detecting the specific deviation occurring. "Multi-label classification" describes systems that can detect the specific deviation occurring (Table 1) and discern between these different aberrant movement patterns. Participants were required to perform 10 BW squats (no external resistance) with acceptable technique, followed by 3 BW squats with each of the predefined deviations from acceptable technique shown in Table 1. During performance of the BW squats, data were acquired from 5 IMUs (SHIMMER; Shimmer Research, Dublin, Ireland) placed on the lumbar spine, right and left thigh, and right and left shank. The IMUs were positioned on each participant by the same researcher using a standardized and repeatable protocol. Participants were allowed a rest interval (around 1 minute) between performances of each set of BW squat repetitions. After data collection, a total of 306 variables were extracted from the sensor signals for every BW squat repetition from each IMU. These variables were used to develop and evaluate the quality of an automated classification system for the analysis of BW squat technique. This was undertaken using data derived from each individual IMU and combinations of multiple IMUs.

Subjects

Seventy-seven healthy volunteers aged 16–40 (55 males, 22 females, age = 22.63 ± 4.87 years, height = 1.75 ± 0.09 m, body mass = 75.03 ± 13.16 kg) participated in the study. No participant reported having a current or recent musculoskeletal injury that would impair his or her performance of the BW squat exercise. All participants reported a level of familiarity with the BW squat exercise. The University College Dublin Human Research Ethics Committee approved the study protocol, and written informed consent was obtained from all participants before testing. In cases where participants were younger than 18 years, written informed consent was also obtained from a parent or guardian.

Procedures

The testing protocol was explained to participants on their arrival at the laboratory. Before formal testing, all participants performed a 10-minute warm-up on a Lode B.V. exercise bike (Groningen, the Netherlands) maintaining a power output of 100 W and constant cadence of 75–85 revolutions per minute. After completion of the warm-up, a Chartered Physiotherapist secured the IMUs to predetermined specific anatomic locations on the participant as follows: over one's clothing at the spinous process of the fifth lumbar vertebra, and directly strapped to the midpoint of both the right and left femurs (determined as half way between the greater trochanter and lateral femoral condyle), and on both shanks 2 cm above the lateral malleolus (Figure 1). The orientation and location of the IMUs was consistent across participants. The neoprene straps used were specifically designed for application in exercise environments. The thick elastic design of the straps minimized the unwanted deviation of IMU position because of loose clothing and movement artifact.

A pilot study was undertaken to determine the most appropriate sampling rate and the ranges for the accelerometer and gyroscope on board the IMUs. For the pilot study, data were acquired (512 samples per second) during performance of the BW squat, lunge, deadlift, single-leg BW squat, and tuck jump exercises. A Fourier transform was then used to estimate the spectral extent of the signals which was found to be less than 20 Hz. Therefore, a sampling rate of 51.2 samples per second was chosen based on the Shannon sampling theorem and the Nyquist criterion (12). Each IMU was configured to stream triaxial accelerometer (±16 g), gyroscope ($\pm 500^\circ \cdot s^{-1}$), and magnetometer (±1 Ga) data with the sensor ranges chosen based on data from the pilot study. Each IMU was calibrated for these specific sensor ranges using the Shimmer 9DoF Calibration application (http://www.shimmersensing.com/shop/shimmer-9dof-calibration).

Initially, participants completed 10 repetitions of the BW squat exercise with acceptable technique. The criteria for acceptable technique were based on the recommendations

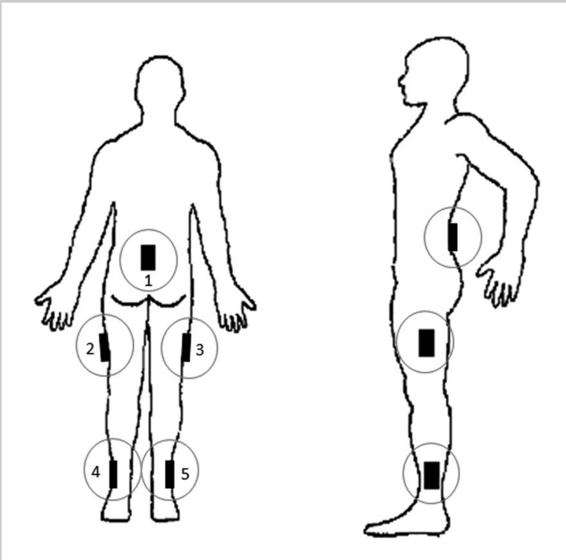

Figure 1. Image showing the 5 IMU positions: (1) The spinous process of the fifth lumbar vertebra, (2 and 3) the midpoint of both femurs on the lateral surface (determined as half way between the greater trochanter and lateral femoral condyle), (4 and 5) and on both shanks 2 cm above the lateral malleolus. IMU = inertial measurement unit.

detailed in the NSCA guidelines (2). This involved participants holding their chest up and out with the head tilted slightly up. As participants moved down into the BW squat position, they were instructed to allow their hips and knees to flex while keeping their torso to floor angle constant. Furthermore, they were required to keep their heels on the floor and knees aligned over their feet. Participants were required to continue flexing at the hips and knees until their thighs were parallel to the floor. As they moved upward, a flat back was to be maintained and they were instructed to keep their chest up and out. Hips and knees were to be extended at the same rate with heels on floor and knees aligned over feet. Participants then extended their hips and knees to reach the starting position. Once the BW squat had been completed with acceptable technique, the participant was instructed to complete the exercise with the technique deviations specified in Table 1. They completed 3 repetitions of each predefined technique deviation. Verbal instructions and a demonstration were provided to all participants, and they were allowed a practice trial to ensure that they were comfortable completing the BW squat with the predefined technique deviations. No external resistance was added during performance of any of the BW squat repetitions. A Chartered Physiotherapist was present during all data collection sessions and ensured that the BW squat and all predefined technique deviations had been completed as instructed. These technique deviations were chosen based on common lower limb deviations outlined by the NSCA (2).

Statistical Analyses

Nine signals were collected from each IMU; accelerometer *x*, *y*, and *z*, gyroscope *x*, *y*, and *z*, and magnetometer *x*, *y*, and *z*. Data were analyzed using MATLAB (2012, The MathWorks, Natick, MA, USA). To ensure that the data analyzed applied to each participant's movement and to eliminate unwanted high-frequency noise, the 9 signals were low pass filtered at $f_c = 20$ Hz using a Butterworth filter of order $n = 8$. Nine additional signals were then calculated. The 3-D orientation of the IMU was computed using the gradient descent algorithm developed by Madgwick et al. (15). The resulting W, X, Y, and Z quaternion values were also converted to pitch, roll, and yaw signals. The pitch, roll, and yaw signals describe the inclination, measured in radians, of each IMU in the sagittal, frontal, and transverse plane respectively. The magnitude of acceleration was also computed using the vector magnitude of accelerometer *x*, *y*, and *z*. The magnitude of acceleration describes the total acceleration of the IMU in any direction. This is the sum of the magnitude of inertial acceleration of the lumbar spine and acceleration due to gravity. In addition, the magnitude of rotational velocity was computed using the vector magnitude of gyroscope *x*, *y*, and *z*.

Each exercise repetition was extracted from the IMU data and resampled to a length of 250 samples; this was undertaken to minimize the influence of the speed of repetition performance on signal feature calculations. It also ensured the computed features related to differences in movement patterns and not the participant's exercise tempo. Repetitions completed by the participant where the IMU's Bluetooth signal dropped were excluded from analysis. The total number of repetitions extracted for each class and used for classification is shown in Table 1. Time-domain and frequency-domain descriptive features were computed to describe the pattern of each of the 18 signals when the 5 different exercises were completed. These features were as follows: "Mean," "RMS," "Standard Deviation," "Kurtosis," "Median," "Skewness," "Range," "Variance," "Max," "Min," "Energy," "25th Percentile," "75th Percentile," "Level Crossing Rate," "Fractal Dimension" (13), and the "variance of both the approximate and detailed wavelet coefficients using the Daubechies 5 mother wavelet to level 7" (http://uk.mathworks.com/help/wavelet/ref/dwt.html). This resulted in 17 features for each of the 18 available signals producing a total of 306 features per IMU.

Figure 2 summarizes the above, whereby 5 IMUs recorded 9 signals each; 9 more signals were derived from these, resulting in a total of 18 signals per IMU. Seventeen features were computed per BW squat repetition for each signal from each IMU, resulting in a total of 1,530 features (306 per IMU, 17 per signal). These features were then used to develop and evaluate a variety of classifiers.

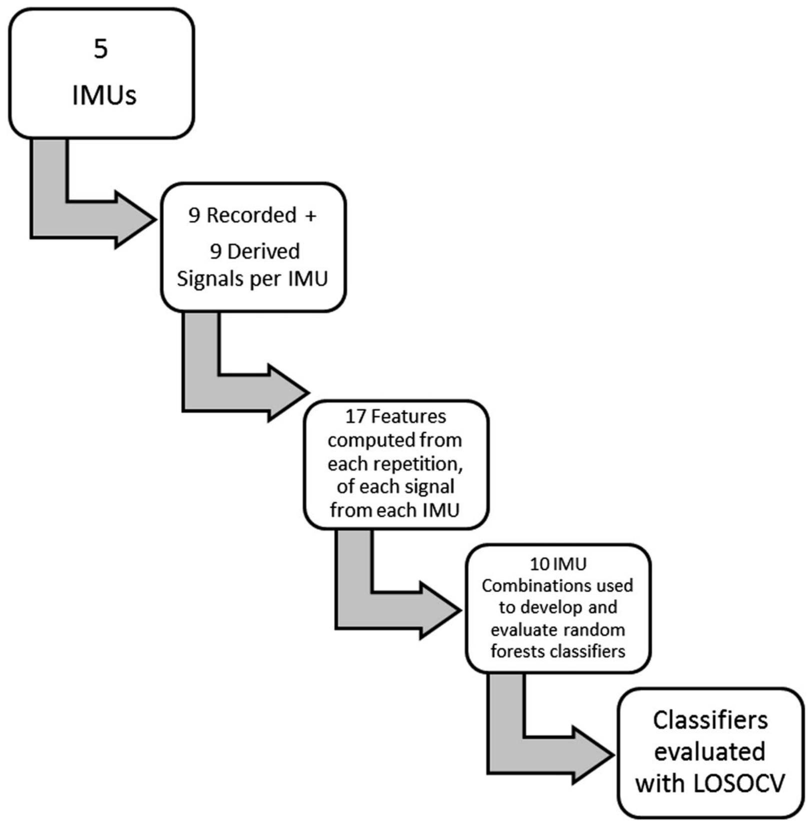

Figure 2. Diagram linking number of IMUs, number of recorded and derived signals, number of features extracted, and the variety of feature combinations used to test classifiers. IMUs = inertial measurement units; LOSOCV = leave-one-subject-out cross-validation.

Initially, binary classification was evaluated to establish how effectively each individual IMU and each combination of IMUs could distinguish between acceptable and aberrant BW squat technique. All repetitions of acceptable performance of the BW squat were labeled "0," and all repetitions of the BW squat performed with one of the predefined deviations as outlined in Table 1 were labeled "1." Multilabel classification was then evaluated on the IMU data to investigate how effectively each individual IMU and each IMU combination could be used to discriminate between acceptable performance of the BW squat exercise and each of the 5 predefined deviations from acceptable technique as described in Table 1. All repetitions of acceptable performance of the BW squat remained labeled as "0," and each of the different deviations was labeled as "1–5."

The random-forests method was used to perform classification (6). This technique was chosen as it has been shown to be effective in analyzing exercise technique with IMUs when compared with the Naive-Bayes and Radial-basis function network techniques (18). Four hundred decision trees were used in each random-forest classifier. Classifiers were developed and evaluated for the 10 combinations of IMUs as shown in Table 2.

The quality of the exercise classification system was established using leave-one-subject-out cross-validation (LOSOCV) and the random-forests classifier with 400 trees (9). Each participant's data correspond to one fold of the cross validation. At each fold, one participant's data are held out as test data, whereas the random-forests classifier is trained with all other participants' data. Where each class in the training data did not have an equal number of instances (i.e., equal number of acceptable and aberrant repetitions in binary classification),

TABLE 2. Inertial measurement unit combinations compared and number of features used for classification.*

Multiple IMUs	*N* features	Individual IMUs	*N* features
All 5 IMUs	1,530 (5 × 306)	Left shank	306
Lumbar and shanks	918 (3 × 306)	Left thigh	306
Lumbar and thighs	918 (3 × 306)	Lumbar	306
Both shanks	612 (2 × 306)	Right thigh	306
Both thighs	612 (2 × 306)	Right shank	306

*IMUs = inertial measurement units.

TABLE 3. Overall accuracy, sensitivity, and specificity in binary classification (acceptable or aberrant technique) for each combination of IMUs following LOSOCV.*

Sensor(s)	Accuracy (%)	Sensitivity (%)	Specificity (%)
All 5 sensors	97.68	96.24	98.55
Lumbar and shanks	96.30	94.26	97.55
Lumbar and thighs	97.26	95.18	98.56
Both shanks	96.24	93.80	97.80
Both thighs	97.41	96.30	98.10
Left shank	95.71	93.19	97.38
Left thigh	97.70	96.02	98.81
Lumbar	95.17	92.73	96.73
Right thigh	95.76	93.63	97.09
Right shank	95.09	93.69	95.95

*IMUs = inertial measurement units; LOSOCV = leave-one-subject-out cross-validation.

random instances of the overrepresented class(es) were removed to balance the training data. The held out data are used to assess the classifier's ability to correctly categorize new data it is presented with. The use of LOSOCV ensures that there is no biasing of the classifiers because the test subjects' data are completely unseen by the classifier before testing. Previous research by Taylor et al. (27) has shown that the use of the test data in training produces results which are not reflective of performance with data not previously seen.

The scores used to measure the quality of classification were total accuracy, average sensitivity, and average specificity. Accuracy is the number of correctly classified repetitions of all the exercises divided by the total number of repetitions completed; this is calculated as the sum of the true positives (TPs) and true negatives (TNs) divided by the sum of the TPs, false positives (FPs), TNs, and false negatives (FNs):

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{TN} + \text{FN}}. \tag{1}$$

In binary classification, acceptable BW squat technique was considered the "positive" class, and aberrant BW squat technique was considered the "negative" class. As such, single sensitivity and specificity values were computed to establish binary classification quality for each IMU combination. In multilabel classification, the sensitivity and specificity were calculated for each of the 6 class labels as outlined in Table 1. Each label was sequentially treated as the "positive" class, and then the mean and *SD* across the 6 values was taken. Sensitivity and specificity were computed using the formulas given below:

TABLE 4. Overall accuracy, average sensitivity, and average specificity in multilabel classification (exact deviation) for each combination of IMUs following LOSOCV.*

Sensor(s)	Accuracy (%)	Sensitivity (%) ± *SD*	Specificity (%) ± *SD*
All 5 sensors	79.91	75.19 ± 12.73	96.04 ± 1.39
Lumbar and shanks	77.18	71.79 ± 14.94	95.52 ± 1.09
Lumbar and thighs	71.25	63.27 ± 20.38	94.38 ± 1.80
Both shanks	74.48	68.52 ± 16.88	94.94 ± 0.90
Both thighs	66.97	56.39 ± 27.19	93.57 ± 1.87
Left shank	72.12	65.03 ± 17.83	94.51 ± 1.47
Left thigh	66.65	55.60 ± 26.22	93.55 ± 2.18
Lumbar	59.75	47.86 ± 29.62	92.15 ± 2.03
Right thigh	63.35	52.98 ± 27.41	92.85 ± 2.34
Right shank	73.10	66.94 ± 16.66	94.71 ± 1.33

*IMUs = inertial measurement units; LOSOCV = leave-one-subject-out cross-validation.

$$\text{Sensitivity} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{2}$$

$$\text{Specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}} \tag{3}$$

Sensitivity measures the effectiveness of a classifier at identifying a desired label, whereas specificity measures the classifier's ability to detect other labels.

In reviewing the accuracy, sensitivity, and specificity scores produced by each classifier, 90% or higher was considered an "excellent" quality result; 80–89% was considered a "good" quality result; 60–79% was considered a "moderate" result; and anything less than 59% was deemed a poor result. The authors chose these values after reviewing the aforementioned literature on identifying deviations from acceptable exercise performance using data derived from IMUs. In reviewing such literature, an existing accepted standard for an excellent, good, moderate, or poor classifier could not be found. Therefore, the above system was agreed on by the authors to facilitate interpretation of results.

Results

Table 3 highlights the overall binary classification accuracy, sensitivity, and specificity achieved by each IMU combination and each individual IMU. Acceptable or aberrant BW squat technique can be detected with 98% accuracy, 96% sensitivity, and 99% specificity when using features derived from all 5 IMUs for classification. It is evident that systems developed using data derived from fewer IMUs can produce equivalent binary classification quality. For instance, a single IMU positioned on the left thigh also achieves 98% accuracy, 96% sensitivity, and 99% specificity.

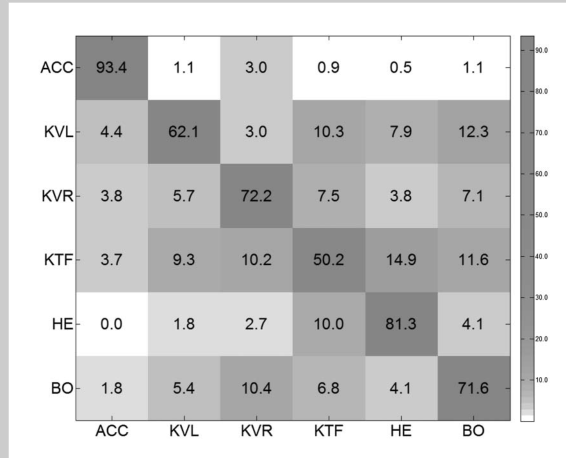

Figure 4. Heat map confusion matrix for multilabel (exact deviation) classification results using data derived from 3 IMUs placed on both shanks and the lumbar spine. IMUs = inertial measurement units.

Table 4 demonstrates the multilabel classification results determined with LOSOCV. The overall accuracy, average sensitivity, and average specificity are displayed for multiple and single IMU systems positioned on various body segments.

A 5 IMU system achieves 80% overall accuracy and an average of 75% sensitivity and 96% specificity. Figure 3 shows that for this IMU set-up, acceptable BW squat technique repetitions are detected with a 95% TP rate. The predefined deviation KTF is the least detected deviation with only a 57% TP rate.

A system which uses 3 IMUs positioned on both shanks and the lumbar spine achieved 77% overall accuracy. This set-up averaged 72% and 96% sensitivity and specificity,

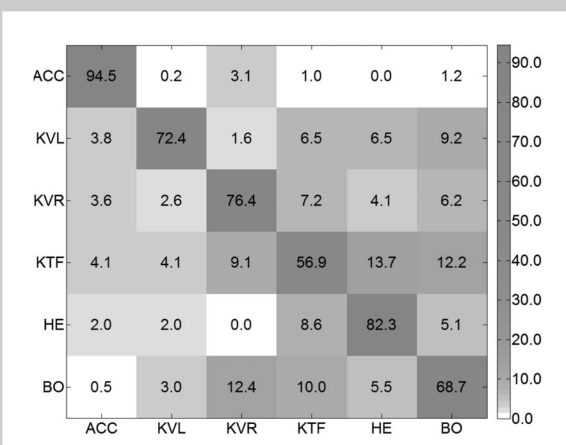

Figure 3. Heat map confusion matrix for multilabel (exact deviation) classification results using data derived from all 5 IMUs. IMUs = inertial measurement units.

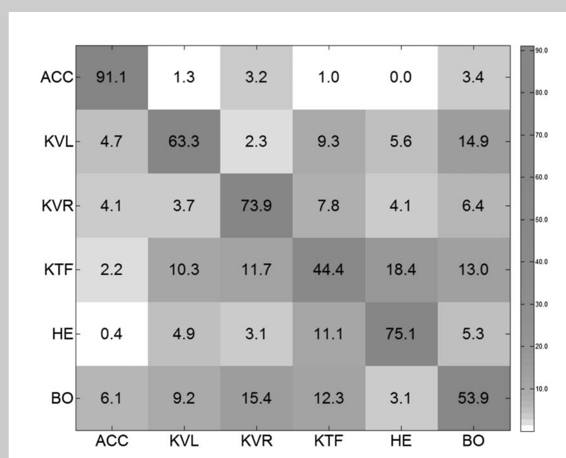

Figure 5. Heat map confusion matrix for multilabel (exact deviation) classification results using data derived from a single IMU placed on the right shank. IMU = inertial measurement unit.

respectively. Figure 4 shows that acceptable BW squat technique is again detected with the highest TP rate; 94%. The predefined deviations KVL and KTF were the most poorly classified. This 3 IMU set-up misclassifies 15% of KTF repetitions to be HE and 12% of KVL repetitions as BO.

A classifier developed with data from a single IMU positioned on the right shank produced 73, 67, and 95% overall accuracy, average sensitivity, and average specificity, respectively. Perhaps counterintuitively this IMU positioned on the shank correctly identifies more BO repetitions than KTF repetitions. In this set, almost 18% of KTF repetitions are misclassified as HE. KVR and HE repetitions are correctly identified with a 74 and 75% TP rate, respectively.

Discussion

This article describes a study to assess the ability of an IMU-based system to distinguish between acceptable and aberrant BW squat technique and compares results using different IMU set-ups. The results presented in this article indicate that an IMU-based system is capable of distinguishing between acceptable and aberrant BW squat technique (binary classification) with an excellent level of overall accuracy with 5, 3, 2, or single IMU set-ups (Table 3). Overall accuracy, sensitivity, and specificity scores are reduced when attempting to identify specific deviations (multilabel classification) as seen in Table 4. All 5 IMUs are capable of identifying specific deviations with a good level of accuracy (80%). Overall accuracy with a reduced IMU set-up (less than 5 sensors) is moderate, with right shank proving the best location for a single IMU. An IMU system at this location is able to distinguish between 5 deviations with just 7% less overall accuracy compared with a 5 IMU set-up.

Multilabel scores are likely to be reduced compared with binary level scores because of the number of classes in the multilabel set. The multilabel classifier needs to be able to distinguish between 6 classes (acceptable, KVL, KVR, KTF, HE, and BO) compared with just 2 in binary classification (acceptable and aberrant). Table 4 demonstrates that multilabel sensitivity scores are far less than specificity scores, impacting on the overall accuracy. Accuracy is related to both sensitivity and specificity (equation 1). Sensitivity measures the effectiveness of a classifier at identifying a desired label (equation 2), whereas specificity measures the classifiers ability to identify if a desired label has not occurred (equation 3). In a real-world application, the ability to identify a desired label is likely to be more important than the ability to detect other labels (i.e., the IMU system would function better if it were able to identify if knee valgus was present rather than rule out knee valgus). Therefore, higher sensitivity scores would likely improve user experience.

The confusion matrices for 5 (Figure 3), 3 (Figure 4), and single (Figure 5) IMU systems show that all are able to detect acceptable BW squatting technique with an excellent TP rate. In each IMU set-up, KTF is the label with the lowest TP rate and it is most commonly confused with HE. The HE deviation is quite similar in nature to KTF and feedback given to prevent this may be comparable. Therefore, an IMU system that looks to provide BW squat technique feedback may be able to group these deviation classes and provide similar feedback once this deviation grouping is detected. This would likely improve the multilabel scores shown in Table 4, as it would reduce the total number of classes to 5 (acceptable, KVL, KVR, KTF+HE, and BO).

The presented results compare favorably with other research in the area. We were able to demonstrate the same overall accuracy (80%), sensitivity (75%), and a 16% improvement in specificity (96%) using a 5 IMU set-up compared with Taylor et al. (26) who analyzed 2 exercises (standing hamstring curl and straight-leg raise). Pernek et al. (23) also used a 5 IMU set-up to monitor exercise intensity in 6 dumbbell upper limb exercises. They did not identify specific deviations but rather exercise intensity as scored by Borg's rating of perceived exertion. As such, their system would not be able to feedback the exact deviations to the user. Melzi et al. (17) presented a demonstration article on detecting errors in the biceps curl, however did not report on overall accuracy, sensitivity, and specificity scores making it impossible to compare the results. Velloso et al. (28) demonstrated an overall accuracy of 78% detecting 4 deviations from acceptable form in the unilateral dumbbell curl. The overall accuracy presented here is higher with a 5 IMU system and just 1% lower with the best 3 IMU system. It is worth noting that one of the IMUs used by Velloso et al. must be placed on the dumbbell being used. This may make it harder to implement in a real-world setting, as the IMU would need to be swapped between dumbbells during a training session.

As discussed in the introduction, a single IMU system is more desirable than a multiple IMU set-up because of ease of use and power considerations. The results in Table 3 indicate that a single IMU system can classify BW squats as acceptable or aberrant with excellent overall accuracy scores (>95% regardless of sensor position). The ability of a single IMU system to identify which deviation has occurred (multilabel classification as shown in Table 4) is moderate, with the right shank showing the highest overall accuracy (73%). A possible reason for this may be that the deviations investigated may involve a high degree of movement in the shanks to complete the aberrant movement. The left and right shank show similar overall classification results. Overall classification scores are around 1–2% higher using the right shank sensor compared with the left. This discrepancy could be attributed to the fact that the majority of the participants were right foot dominant, leading to overcompensation on this side. The confusion matrix in Figure 5 shows that the right shank is able to classify normal reps with an excellent TP rate, but deviations such as KTF and BO are less clearly classified, possibly due to the similar movement profile needed to complete these deviations.

Pernek et al. (22) used a single IMU to capture resistance training information and were able to recognize repetition duration with a temporal detection error of about 11%. However, different exercise goals may require varying movement speeds. As such, assessment based on speed alone is not a holistic way of evaluating exercise technique. Bonnet et al. (5) demonstrated that a single IMU system on the lower back could measure ankle, knee, and hip joint angles in the sagittal plane during the BW squat with a maximal error of 3.5 degrees compared with a Vicon motion capture system in human participants. However, the ability to display angles to the end user may not be actionable information to an individual not trained in biomechanics as they may not be able to distinguish what angle range represents an aberrant movement pattern. Furthermore, the authors were only able to identify sagittal plane angles. Many common BW squat deviations can occur in different or multiple planes simultaneously, such as knee valgus and varus deviations, which occur predominantly in the frontal plane. Giggins et al. (10) assessed the ability of a single IMU to identify deviations in 7 exercises (heel slide, hip abduction, hip flexion, hip extension, knee extension, inner range quads, and straight-leg raise) with an overall accuracy of 79–83% depending on where the IMU was positioned. These results are slightly higher than those presented in this work; however, the authors looked at a maximum of 3 deviations for each exercise, whereas some of the exercises only had one deviation (i.e., binary level classification). Our work sought to identify 5 deviations from normal, and this may go some way to explaining the lower overall accuracy scores compared with Giggins et al. Furthermore, the BW squat exercise is more complex than the exercises classified by Giggins et al. with deviations occurring in multiple joints simultaneously.

Comparing results presented in this article with the above work is challenging because of differences in exercises investigated, sensor positions, and feedback given to end users. However, these results build on the previous work. The majority of research to date has investigated the ability of IMU systems to monitor technique in simple exercises such as heel slides (10), dumbbell curls (28), or straight-leg raises (26). This article describes an evaluation of an IMU system's ability to quantify BW squatting performance, a complex exercise that involves multiple joints. This system has also demonstrated the ability to identify 5 deviations from normal technique (Table 4). The reduced number of deviations in some of the studies (10,26,28) may make it easier for classifiers to identify specific deviations and subsequently produce higher accuracy, sensitivity, and specificity scores. Finally, the main focus of the IMU system analyzed in this article is to identify specific technique deviations and not just angles of movement (5), tempo (22), or exercise intensity (23). This information may be more actionable to gym users, particularly those without biomechanics training.

It is difficult to ascertain whether the moderate levels of multilevel classification presented in this article are sufficient for real-life strength and conditioning (S&C) settings. Further research is being undertaken to determine usability, functionality, and user perceptions of using wearable technology to assess exercise biomechanics. It is hoped this will give a greater indication as to the levels of accuracy S&C coaches, experienced and novice gym users would define as acceptable. It is possible that these levels may vary depending on the clients S&C coaches work with and the training goals of gym users.

A number of contextual factors must be taken into account when interpreting the results. All deviations were deliberately induced and completed by healthy individuals. When deviations occur naturally, the exact way in which they present may differ from the induced deviations investigated in this study. Moreover, there were no controls as to the severity with which each participant performed a deviation, and therefore it is possible that naturally occurring BW squat deviations in a "real life" application may be more acute, or occur in a more idiosyncratic fashion than those used in the presented classification systems. No gold standard 3-dimensional motion capture system was used to confirm that each deviation occurred. However, a Chartered Physiotherapist and an individual trained in S&C were present for all data collection and ensured the deviations occurred through visual observation. The participant was asked to repeat the movement if the investigators felt it had not been completed satisfactorily. A motion capture system was not used as researchers have already shown the reliability of IMU set-ups compared with such systems (14,25). In addition, the 5 deviations identified by the IMU systems is a nonexhaustive list of those that can occur during the BW squat exercise. These deviations were chosen in consultation with sports and medicine practitioners, S&C coaches, and the NSCA guidelines (2).

In conclusion, it is shown that a system based on data derived from body-worn IMUs can classify acceptable and aberrant BW squat biomechanics with excellent overall accuracy, sensitivity, and specificity. These excellent classification levels are maintained even with a single IMU. The ability to identify specific stimulated deviations is more difficult but can be achieved with a good level of overall accuracy with a 5 IMU system. A single IMU system can identify specific deviations with a moderate level of accuracy. These results are comparable with current research in the area, despite the BW squat being a more complex exercise than many of those previously investigated. However, it must be stressed that this is not a fully operational system at present. Such a system should be able to recognize and evaluate multiple exercises. Furthermore, the deviations investigated in this study are induced, meaning how they appear in a natural setting may be different. However, the results presented in this article are promising and further research is warranted to investigate an IMU system's capability of monitoring technique in various movements. This work is currently ongoing with future research focusing on the analysis of natural deviations and exercises such as the deadlift and tuck jump.

Practical Applications

The BW squat is an important movement in S&C, musculoskeletal injury, risk screening, and rehabilitation. The ability to objectively quantify BW squatting technique using low-cost IMU technology would have practical advantages in all these settings. In an S&C setting, the ability to remotely monitor form may help exercise goals be achieved and reduce the risk of injury. In musculoskeletal screening, an IMU system that can identify aberrant movement patterns would allow for quicker and more objective risk identification and stratification. In a rehabilitation setting, monitoring technique would be important to prevent further injury and possibly allow exercises to be completed at home without the need for constant supervision, reducing overall health care costs.

Acknowledgments

This project is partly funded by the Irish Research Council as part of a Postgraduate Enterprise Partnership Scheme with Shimmer (EPSPG/2013/574) and partly funded by Science Foundation Ireland. M. A. O'Reilly and D. F. Whelan are Joint lead authors.

References

1. Ahmadi, A, Mitchell, E, Destelle, F, Gowing, M, O'Connor, NE, Richter, C, and Moran, K. Automatic activity classification and movement assessment during a sports training session using wearable inertial sensors. In: Presented at Proceedings of the 11th International Conference on Wearable and Implantable Body Sensor Networks (BSN), Zurich, Switzerland, June 16–19, 2014.
2. Baechle, TR and Earle, RW. *NSCA's Essentials of Personal Training.* Champaign, IL: Human Kinetics, 2004.
3. Bonnechere, B, Jansen, B, Salvia, P, Bouzahouene, H, Omelina, L, Moiseev, F, Sholukha, V, Cornelis, J, Rooze, M, and Van Sint Jan, S. Validity and reliability of the Kinect within functional assessment activities: Comparison with standard stereophotogrammetry. *Gait Posture* 39: 593–598, 2014.
4. Bonnet, V, Mazza, C, Fraisse, P, and Cappozzo, A. A least-squares identification algorithm for estimating squat exercise mechanics using a single inertial measurement unit. *J Biomech* 45: 1472–1477, 2012.
5. Bonnet, V, Mazza, C, Fraisse, P, and Cappozzo, A. Real-time estimate of body kinematics during a planar squat task using a single inertial measurement unit. *IEEE Trans Biomed Eng* 60: 1920–1926, 2013.
6. Breiman, L. Random forests. *Machine Learning* 45: 5–32, 2001.
7. Chang, KH, Chen, MY, and Canny, J. Tracking free-weight exercises. In: *UbiComp 2007: Ubiquitous Computing.* Berlin, Germany: Springer, 2007. pp. 19–37.
8. Cook, G, Burton, L, and Hoogenboom, B. Pre-participation screening: The use of fundamental movements as an assessment of function–Part 1. *N Am J Sports Phys Ther* 1: 62–72, 2006.
9. Fushiki, T. Estimation of prediction error by using K-fold cross-validation. *Stat Comput* 21: 137–146, 2011.
10. Giggins, OM, Sweeney, KT, and Caulfield, B. Rehabilitation exercise assessment using inertial sensors: A cross-sectional analytical study. *J Neuroeng Rehabil* 11: 158–168, 2014.
11. Hall, M, Nielsen, JH, Holsgaard-Larsen, A, Nielsen, DB, Creaby, MW, and Thorlund, JB. Forward lunge knee biomechanics before and after partial meniscectomy. *Knee* 22: 506–509, 2015.
12. Jerri, AJ. The Shannon sampling theorem–Its various extensions and applications: A tutorial review. *Proc IEEE* 65: 1565–1596, 1977.
13. Katz, MJ and George, EB. Fractals and the analysis of growth paths. *Bull Math Biol* 47: 273–286, 1985.
14. Leardini, A, Lullini, G, Giannini, S, Berti, L, Ortolani, M, and Caravaggi, P. Validation of the angular measurements of a new inertial-measurement-unit based rehabilitation system: Comparison with state-of-the-art gait analysis. *J Neuroeng Rehabil* 11: 1–7, 2014.
15. Madgwick, SOH, Harrison, AJL, and Vaidyanathan, R. Estimation of IMU and MARG orientation using a gradient descent algorithm. In: Presented at Proceedings of the IEEE International Conference on Rehabilitation Robotics (ICORR), Zurich, Switzerland, June 29–July 1, 2011.
16. McGrath, D, Greene, BR, O'Donovan, KJ, and Caulfield, B. Gyroscope-based assessment of temporal gait parameters during treadmill walking and running. *Sports Eng* 15: 207–213, 2012.
17. Melzi, S, Borsani, L, and Cesana, M. The virtual trainer: Supervising movements through a wearable wireless sensor network. In: Presented at Sensor, Mesh and Ad Hoc Communications and Networks Workshops, 2009 SECON Workshops' 09 6th Annual IEEE Communications Society Conference on, Rome, Italy, June 22–26, 2009.
18. Mitchell, E, Ahmadi, A, O'Connor, NE, Richter, C, Farrell, E, Kavanagh, J, and Moran, K. Automatically detecting asymmetric running using time and frequency domain features. Presented at Proceedings of the 12th International Conference on Wearable and Implantable Body Sensor Networks (BSN), Boston, MA, June 9–12, 2015.
19. Morris, D, Saponas, TS, Guillory, A, and Kelner, I. RecoFit: Using a wearable sensor to find, recognize, and count repetitive exercises. In: Presented at Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Toronto, ON, Canada, April 26–May 1, 2014.
20. Muehlbauer, M, Bahle, G, and Lukowicz, P. What can an arm holster worn smart phone do for activity recognition? In: Proceedings of the International Symposium on Wearable Computers (ISWC). IEEE, San Francisco, CA, June 12–15, 2011, pp 79–82.
21. O'Donovan, KJ, Greene, BR, McGrath, D, O'Neill, R, Burns, A, and Caulfield, B. SHIMMER: A new tool for temporal gait analysis. Presented at Engineering in Medicine and Biology Society, 2009 EMBC 2009 Annual International Conference of the IEEE, Minneapolis, USA, September 2nd–6th, 2009.
22. Pernek, I, Hummel, KA, and Kokol, P. Exercise repetition detection for resistance training based on smartphones. *Personal Ubiquitous Computing* 17: 771–782, 2013.
23. Pernek, I, Kurillo, G, Stiglic, G, and Bajcsy, R. Recognizing the intensity of strength training exercises with wearable sensors. *J Biomed Inform* 58: 145–155, 2015.
24. Seeger, C, Buchmann, A, and Van Laerhoven, K. myHealthAssistant: A phone-based body sensor network that captures the wearer's exercises throughout the day. In: Proceedings of the 6th International Conference on Body Area Networks. ICST, Beijng, China, November 7–10, 2011, pp 1–7.
25. Tang, Z, Sekine, M, Tamura, T, Tanaka, N, Yoshida, M, and Chen, W. Measurement and estimation of 3D orientation using magnetic and inertial sensors. *Adv Biomed Eng* 4: 135–143, 2015.
26. Taylor, PE, Almeida, GJ, Hodgins, JK, and Kanade, T. Multi-label classification for the analysis of human motion quality. In: Presented at Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), San Diego, CA, August 28–September 1, 2012.
27. Taylor, PE, Almeida, GJ, Kanade, T, and Hodgins, JK. Classifying human motion quality for knee osteoarthritis using accelerometers. Presented at Engineering in Medicine and Biology Society (EMBC), Buenos Aires, Argentina, August 31–September 4, 2010.
28. Velloso, E, Bulling, A, Gellersen, H, Ugulino, W, and Fuks, H. Qualitative activity recognition of weight lifting exercises. In: Presented at Proceedings of the 4th Augmented Human International Conference, Stuttgart, Germany, March 7–8, 2013.
29. Whiteside, D, Deneweth, JM, Pohorence, MA, Sandoval, B, Russell, JR, McLean, SG, Zernicke, RF, and Goulet, GC. Grading the functional movement screen: A comparison of manual (real-time) and objective methods. *J Strength Cond Res* 30: 924–933, 2016.