The performance of an algorithm for classifying gym-based tasks across individuals with different body mass index

Abstract

Previous activity classification studies have typically been performed on normal weight individuals. Therefore, it is unclear whether a generic classification algorithm could be developed that would perform consistently across individuals who fall within different BMI categories. Acceleration data were collected from the hip and ankle joints of 50 individuals: 17 normal weight, 14 overweight and 19 obese. Each participant performed a set of 10 dynamic tasks, which included activities of daily living and gym-based exercises. The performance of a generic classification algorithm, developed using linear discriminant analysis, was compared across the three separate BMI groups for each sensor. Higher classification accuracies (92-95%) were observed for the ankle sensor; however, both sensors demonstrated consistent performance across the three groups. This is the first study to demonstrate the effectiveness of a generic classification algorithm across individuals with different BMI and may be a first step towards automated activity profiling in weight-loss programmes.

Key Words:

Classification; activity monitoring; obese

The performance of an algorithm for classifying gym-based tasks across individuals with different body mass index

Introduction

Objective measurement of physical activity in people who are overweight or obese is important for understanding the associations between patterns of physical activity and health-related outcomes (Healy, Winkler, Brakenridge, Reeves, & Eakin, 2015), and also for evaluating and improving exercise interventions aimed at reducing weight (Goode et al., 2016). Using automated measurement of physical activity to provide individual feedback may also enhance adherence to prescribed physical activity during weight loss programmes (Cheatham, Stull, Fantigrassi, & Motel, 2018). In support of this idea, research has shown that participants who receive objective feedback on steps taken and on periods of moderate and vigorous activity, lose significantly more weight at 6-months, compared to participants without access to self-monitoring technology (Hartman et al., 2016; Ross & Wing, 2016).

Objective measurement of physical activity is readily achieved using a body-worn accelerometer, with analysis typically based on the use of predetermined thresholds, referred to as cut-points, which are applied to a measure of signal magnitude (Howe, Moir, & Easton, 2017). With this approach, periods of activity are grouped into signal-intensity categories in order to give an overall measure of physical activity level (Tudor-Locke, Brashear, Johnson, & Katzmarzyk, 2010). Over recent years, support has grown for the use of “smart” approaches for estimating energy expenditure from accelerometer data (Plasqui, 2017). These techniques typically involve an initial recognition of the activity type, after which an activity-specific regression equation is used to obtain a value of energy expenditure (Bonomi, Plasqui, Goris, & Westerterp, 2009). Because activities that require different levels of energy expenditure can sometimes lead to similar magnitudes of the acceleration signal, this approach can be used to improve predictions of energy expenditure (Crouter, Kuffel, Haas, Frongillo, & Bassett, 2010).

In addition to improving energy expenditure estimation, automated recognition of different activities could be used to provide meaningful information to people on a weight loss programme (Stewart, Ferguson, Peng, & Rafferty, 2012). For example, providing people with personalised information on the activities that they undertake may be easier to interpret with respect to behaviour change, than simply providing feedback on the amount of time spent at different intensities of physical activity (Rabbi, Pfammatter, Zhang, Spring, & Choudhury, 2015). Furthermore, the ability to accurately profile individual activities may improve understanding of the type, intensity and duration of activities that are most effective for weight loss and could, in turn, lead to enhanced public health guidelines (Department of Health, 2019). It is therefore important to develop classification algorithms that can differentiate between activities and which will work effectively with accelerometer data collected from people who are overweight or obese.

To date, many studies that have created algorithms that can identify different activities from accelerometer data in normal weight participants (Ahmadi, Pfeiffer, & Trost, 2020; Farrahi, Niemelä, Kangas, Korpelainen, & Jämsä, 2019; Gao, Bourke, & Nelson, 2014). However, there has been minimal research that has aimed to test classification techniques across different body mass index (BMI) groups. This is important, since obese individuals display different movement characteristics to non-obese individuals (da Silva-Hamu et al., 2013; Lai, Leung, Li, & Zhang, 2008; Sibella, Galli, Romei, Montesano, & Crivellini, 2003), which may result in differences in acceleration data for the same task. Furthermore, increased adipose tissue at the point where activity monitors are placed could be responsible for additional movement of the monitor and may lead to differences in signal characteristics between BMI groups. Interestingly, studies have shown that increased BMI can lead to inaccuracy in pedometer-derived step count (Melanson et al., 2004) and that differences in tilt angle may underlie such error (Crouter, Schneider, & Bassett, 2005). Other research suggests that the accuracy of accelerometer-derived step count may be reduced in pregnant women (Connolly, Coe, Kendrick, Bassett, & Thompson, 2011). Given these findings, it is important to understand whether increased BMI affects the performance of activity classification algorithms.

Clinical studies that have examined the effects of physical exercise interventions on health-related outcomes for overweight and obese individuals have included many different aerobic activities. These include walking (Kiernan, King, Stefanick, & Killen, 2001), jogging (Wood et al., 1988), cycling (Cox, Burke, Morton, Beilin, & Puddey, 2004), rowing (Raz, Hauser, & Bursztyn, 1994) and stair stepping (Janssen, Fortier, Hudson, & Ross, 2002). In a recent study focused on people who were overweight or obese, Ellis *et al*. (2016) showed it was possible to accurately identify sitting, standing, gait (walking/running), or riding in a vehicle. However, further research is required to understand to potential of creating algorithms which can be used to identify a wide range of activities in people who are obese. Given the limitations of previous research, this study aimed to develop and evaluate the accuracy of a classification algorithm, focused on gym-based exercises and activities of daily living, across different BMI groups. A secondary aim of this study was to compare classification accuracy between two accelerometer placement sites: the hip and the ankle.

Materials and methods

A sample of fifty participants completed the study. Participants were recruited via university advertisement, through online forums, social media platforms and through local weight loss groups. Individuals across a range of BMI categories were approached and recruitment was continued until it was possible to create three groups, with similar numbers of participants, defined by BMI: normal (18.5-24.9 kg/m^2), overweight (25.0-29.9 kg/m^2) and obese (≥30 kg/m^2). To ensure that physical exercise did not pose any risk, participants were required to complete the physical activity readiness questionnaire (Thomas, Reading, & Shephard, 1992). Participants with a history of high or low blood pressure were also excluded to minimise any possible risks during the testing protocol. In addition, people with diabetes were excluded as this condition may affect mobility, altering normal movement patterns.

The study, which was conducted over a period of two years, was approved by the College of Health and Social Care Ethical Approval Panel at the University of Salford, and each participant provided written informed consent. Before activity data were collected, descriptive characteristics were obtained. These included age, weight (using electronic scales), height (using a stadiometer), waist circumference (using a tape measure at the level of the navel) and hip circumference (using a tape measure at widest part of the hips). It also included body fat percentage, measured using the Bodystat 1500 body composition analyser (BodyStat Ltd, Douglas, Isle of Man, UK).

An ActiGraph GT3X+ activity monitor, sampling at 50Hz, was fitted at the right hip, either directly over, or just above, the iliac crest. Where possible, the strap supplied with the monitor, was threaded through the participant's belt loops. If this was not possible, the strap was adjusted appropriately to ensure minimal movement of the monitor during dynamic activities. Another GT3X+ activity monitor (50Hz) was fitted at the right ankle, directly above the lateral malleolus using the elastic strap supplied.

Each participant performed a total of ten different activities during a single testing session. The tasks included three actives of daily living: walking, stair ascent and stair descent. In addition, we included a set of gym-based activities, including treadmill walking, treadmill jogging, cross-training (using an elliptical trainer in a standing position), rowing (on a rowing machine) and static cycling. The gym-based activities also included step aerobics (stepping up and down on a step), side-stretching (alternative leaning to opposite sides with arms raised) and sidestepping (initially standing with legs apart, then moving one foot to meet the other before returning to the initial position and repeating for the other foot). These activities were selected from consideration of previous obesity management studies and to ensure that there was a range of movement patterns large enough to present a sufficient challenge to the activity classification algorithm. Simple postural activities (standing, sitting and lying) were not included in this study as these tasks can be differentiated from dynamic activities using a simple threshold-based approach (Preece et al., 2009). It was felt that a threshold-based classifier could be implemented before a dynamic classification scheme to create a classification approach, which would be able to deal with both simple postural activities as well as dynamic activities.

Walking data were collected both on a treadmill and over ground. For the treadmill walking, each participant was required to walk on the Ergo ELG55 treadmill (WOODWAY GmbH of Weil am Rheine, Germany) at four different speeds, ranging from slow to fast (approximately between 1.0 ms^{-1} and 1.7 ms^{-1}), for five minutes each. These speeds were determined by the participant's capacity for exercise and based upon a timed walk to ascertain normal walking speed. Participants were then asked to walk a designated route outdoors at a self-selected speed. The duration of the walk was approximately three minutes, depending on their walking speed. The walking surface was paved, and in some places uneven or sloped. With this protocol, walking data were collected under different conditions across a range of different speeds; however, for the classification analysis, all walking data were considered the same type of activity.

Participants were given some basic instruction on how to use exercise equipment or perform the aerobic activities, though no guidance was given with respect to technique. With the exception of treadmill walking, described above, participants were asked to perform each exercise at their own pace for at least one minute and were instructed to maintain a consistent level of intensity. This included the treadmill jogging for which participants were instructed to select a treadmill speed, which was appropriate to their fitness level. For each activity, start and end times were manually recorded using a clock that was synchronised with the ActiGraph GT3X+ activity monitor.

Three-dimensional acceleration data were obtained from the GT3X+ monitor and imported to MATLAB. The accelerometer data from each participant were segmented into sequential, non-overlapping windows of two seconds in duration. Each window was then associated with a specific activity type using the times recorded during data collection. A maximum of thirty windows (60 seconds data) of activity were selected at random for each of the ten activities and used for subsequent analysis; however, in some shorter duration activities, such as stair walking, fewer windows of data were generated. This procedure was repeated for the data from all 50 participants.

Features were derived from the accelerometer data on a window-by-window basis for use as input to a classification algorithm. Five features were based on those suggested by Baek *et al.* (2004), mean, standard deviation, eccentricity, kurtosis, and skewness. In addition, five discrete cosine transform (DCT) components were select to capture frequency characteristics (He & Jin, 2009) and zero crossing rate (Maurer, Smailagic, Siewiorek, & Deisher, 2006), signal magnitude area (SMA), and percentile values (10^{th}, 25^{th}, 75^{th}, 90^{th}) selected to extend the statistical information. The features were obtained separately from each of the three accelerometer axes. The only exception was where the calculation required a combination of the three axes, as in the case of SMA.

Linear Discriminant Analysis (LDA) was chosen as the classification algorithm (Balakrishnama & Ganapathiraju, 1998). LDA defines a probability distribution for each possible class of activity based on training data. A window of activity data is classified by applying a discriminant function that returns the likelihoods of the window belonging to each activity distribution. The activity that is chosen by the classifier is the one returning the highest likelihood value from the discriminant function. For further details of LDA, the reader is directed to Balakrishnama *et al.* (1998). Example code for the implementation of this algorithm can be downloaded at: https://salford.figshare.com/articles/software/Example_LDA_classifier_zip/12613826

Our primary objective was to understand whether a generic classification algorithm, created using the full cohort, would perform equally well when applied to different BMI groups. To address this objective, a leave-one-out cross-validation was used to obtain a classification accuracy for each of the activities performed by each individual participant. This was achieved by creating a classifier with training data from all participants apart from the individual under test. This algorithm was then applied to the data from the individual under test to obtain a predicted activity class for each activity window. This procedure was repeated until every participant had been tested once, after which average classification accuracies were calculated for four different groups: all participants, normal, overweight, and obese. For each group, the classification accuracy was calculated by dividing the number of correctly classified windows by the total number of activity windows for the group. In addition, sensitivity was obtained for each activity class, by calculating the percentage of correctly identified windows of that activity compared to the total number of windows of that activity. All analyses were performed independently for hip- and ankle-mounted accelerometer data.

In addition to calculating point estimates of classifier accuracy, we used a two-way ANOVA analysis to test for statistical differences in classifier performance across the three groups (normal weight, overweight and obese) and between the two placement sites (hip and ankle). The study was powered to detect an effect size of 0.5 in a three-group ANOVA comparison, which related to our primary objective of investigating differences in classifier accuracy between the three groups. We assumed a power = 0.8 and a critical alpha = 0.05. Using the g-power software, we estimated we would need a total of 42 subjects, at least 14 in each group. Our sample of 50 was therefore sufficient to detect differences between the three groups. Differences in descriptive characteristics between the three groups were investigated using ANOVA or chi-squared as appropriate, again with a critical alpha = 0.05.

Results

Descriptive characteristics for all three groups are presented in Table 1. There were no statistical differences in age, gender or height between the three groups. However, the other BMI-related parameters were shown to differ (P<0.05).

TABLE 1 HERE

There were relatively small differences in overall classification accuracy between the three BMI groups (Table 2). Specifically, data from the ankle placement demonstrated a difference of 3.6 percentage points between the normal and overweight group. Although there was slightly more variability for the hip placement, classification accuracies were still relatively consistent, with a difference of 8.2 percentage points between the overweight group and the obese group (Table 2). The ANOVA analysis showed no main effect of group (p=0.15), nor group-placement interaction (p=0.15), confirming that classification accuracy did not differ statistically between the three groups.

Figure 1a shows the sensitivity of the algorithm, broken down by activity, for the ankle placement. This plot illustrates that the algorithm was associated with similar levels of accuracy, across the groups, for most activities. With the exception of cycling and stepping, which were associated with lower sensitivities for the normal weight participants, the other eight activities were associated with minimal (<6 percentage point) differences in sensitivity between the BMI groups. In contrast, there was more variability between the groups at the hip site (Figure 1b), with six out of the ten activities being associated with moderate (>10 percentage point) differences in algorithm sensitivity.

TABLE 2 HERE

FIGURE 1A AND 1B HERE

Classification accuracy for all 50 participants (entire group) was 8.9 percentage points higher for the ankle compared to the hip placement. Importantly, higher classification for the ankle placement was observed for each of the separate BMI groups (Table 2). The ANOVA analysis showed a main effect of placement (p<0.001), confirming that classification accuracy was statistically higher for the ankle. The observation of higher classification accuracies is clearly visible in Figure 1, which shows algorithm sensitivity, broken down by activity. These data illustrate that ankle sensitivities were up to 37 percentage points higher and typically 5-10 percentage points higher than the corresponding hip sensitivity. Confusion matrices for all participants for the hip (Table 3) and the ankle (Table 4) illustrate the source of activity misclassification. For the hip sensor, there was confusion between upstairs and stepping, between cycling and side-stretching and between side-stepping and side-stretching. For the hip, misclassification was primarily between upstairs and stepping and between the cross-trainer and cycling (Table 4).

TABLES 3 AND 4 HERE

Discussion

The primary objective of this study was to establish if a generic activity classification algorithm would perform consistently across individuals with different BMI. To address this objective, we tested our algorithm on three separate groups, defined by BMI, showing similar levels of classification accuracy across the groups. When the data were broken down into different activities, we observed consistency across the BMI groups; however, algorithm sensitivities were more consistent for the ankle when compared to the hip sensor. Taken together, these findings suggest that differences in BMI are unlikely to affect classifier performance if the aim is to differentiate between gym-based exercises.

The motivation for this study came from previous research that has demonstrated that BMI may impact on movement patterns (da Silva-Hamu et al., 2013; Lai et al., 2008; Sibella et al., 2003), and that body shape may influence the accuracy of accelerometer-derived step count (Connolly et al., 2011). It is possible that increased adipose tissue may create additional variability in the acceleration signals during the performance of the same activity, potentially lowering classification accuracy. In line with this idea, we observed a degree of variability in classifier performance at the hip, when the results were broken down into different activities (Figure 1); however, there was minimal variation for the ankle placement. This observation may reflect the idea that hip accelerations are more likely to be affected by soft tissue motion than accelerations measured at the ankle. As our overall classification accuracies were consistent across the different BMI groups, the findings suggest that it should be possible to a create a generic classification algorithm using data from either a hip or ankle mounted accelerometer.

A secondary objective of this study was to understand potential differences in classification accuracy between the ankle and hip site. In addition to demonstrating lower variability across the three BMI groups, we also observed higher overall performance from the ankle placement (Figure 1). It is likely that this improved performance is a result of the larger range of motion, and therefore distinct acceleration patterns, which are likely to be associated with the ankle in comparison to the hip site. We suggest that this leads to improved classifier performance. We would therefore advocate the practice of placing an accelerometer at the ankle if the objective is to differentiate between multiple different activities. However, we acknowledge that this current study was only tested to a gym setting and therefore further research is required to investigate a more comprehensive set of daily activities.

It is difficult to compare the classification accuracies found in this study with those reported in other studies because of differences in activity sets and metrics used to calculate accuracy. Nevertheless, overall classification accuracies greater than 90% compare favourably with other studies on normal weight individuals (Gao et al., 2014; Gupta & Dallas, 2014; Moncada-Torres, Leuenberger, Gonzenbach, Luft, & Gassert, 2014; Parkka et al., 2006). However, it is important to point out that several previous studies have generated and tested classification algorithms on small numbers of participants, typically n=10 or fewer (Chang, Chen, & Canny, 2007; Gupta & Dallas, 2014; Moncada-Torres et al., 2014; Qi, Yang, Hanneghan, Tang, & Zhou, 2019). By developing and testing an algorithm on 50 participants, we have shown that our classification approach can handle a wide range of individual variation, which is likely to be representative of the general population. This study purposefully excluded static activities (sitting, standing and lying), as it was felt that differentiation between static and dynamic activities is a well-studied problem (Preece et al., 2009), and that inclusion of static activities would have given an inflated picture of the effectiveness of the classification scheme. Importantly, despite excluding sitting, standing and lying, it was still possible to obtain over 90% accuracy. When combined with classification schemes that can be used to differentiate between static and dynamic activities, it is likely that much higher levels of accuracy would be achieved.

There are two primary limitations to this study, which should be highlighted. First, data were not collected under free-living conditions, but were obtained using a predefined protocol under laboratory conditions. However, a wide range of different activities were studied and participants were given minimal instruction on how to perform each task; thereby, presenting the classification algorithm with considerable variability in accelerometer data. Although the algorithm performed well, we acknowledge that further development and testing is required to create a system that would be able to deal with real-world data. Another limitation was that our testing was limited to sensors on the ankle and hip. Sensor data were not collected from the wrist, as our objective was to be confident that we obtained dynamic signals for each activity, which could easily be differentiated from signals associated with sedentary activities (e.g. sitting). With the inclusion of cycling in our protocol, the ankle and hip were deemed to be placement sites that would provide dynamic signals, whilst at the same time being acceptable to people in a real-word setting.

This study demonstrates that it is possible to identify activities of daily living and gym-based exercises with a single accelerometer in normal weight, overweight, and obese individuals. The highest accuracy was obtained with the ankle sensor (92–95%); however, classification accuracies of 82–90% were also obtained from the hip accelerometer. Further research is required to integrate the proposed classification scheme into an algorithm that can deal with real-world data. This will enable the generation of continuous activity profiles for overweight and obese individuals undertaking programmes of prescribed physical activity. Such profiles should help improve adherence to prescribed physical activity programmes and provide greater insight into the relationship between physical activity type and weight loss.

Acknowledgements

The research leading to these results received funding from the European Community's Seventh Framework Programme under SSHOES project, Grant Agreement no. NMP2-SE-2009-229261

References

Ahmadi, M. N., Pfeiffer, K. A., & Trost, S. G. (2020). Physical Activity Classification in Youth Using Raw Accelerometer Data from the Hip. *Measurement in Physical Education and Exercise Science*, 1-7. http://dx.doi.org/10.1080/1091367X.2020.1716768.

Baek, J., Lee, G., Park, W., & Yun, B. J. (2004). *Accelerometer signal processing for user activity detection.* Knowledge-Based Intelligent Information and Engineering Systems. KES 2004. Lecture Notes in Computer Science, vol 3215. Springer, Berlin, Heidelberg. http://dx.doi.org/https://doi.org/10.1007/978-3-540-30134-9_82.

Balakrishnama, S., & Ganapathiraju, A. (1998). Linear discriminant analysis-a brief tutorial. *Institute for Signal and information Processing, 18*, 1-8.

Bonomi, A. G., Plasqui, G., Goris, A. H., & Westerterp, K. R. (2009). Improving assessment of daily energy expenditure by identifying types of physical activity with a single accelerometer. *J Appl Physiol (1985), 107*(3), 655-661. http://dx.doi.org/10.1152/japplphysiol.00150.2009.

Chang, K.-h., Chen, M. Y., & Canny, J. (2007). *Tracking Free-Weight Exercises*, Berlin, Heidelberg.

Cheatham, S. W., Stull, K. R., Fantigrassi, M., & Motel, I. (2018). The efficacy of wearable activity tracking technology as part of a weight loss program: a systematic review. *The Journal of sports medicine and physical fitness, 58*(4), 534-548. http://dx.doi.org/10.23736/s0022-4707.17.07437-0

Connolly, C. P., Coe, D. P., Kendrick, J. M., Bassett, D. R., Jr., & Thompson, D. L. (2011). Accuracy of physical activity monitors in pregnant women. *Med Sci Sports Exerc, 43*(6), 1100-1105. http://dx.doi.org/10.1249/MSS.0b013e3182058883.

Cox, K. L., Burke, V., Morton, A. R., Beilin, L. J., & Puddey, I. B. (2004). Independent and additive effects of energy restriction and exercise on glucose and insulin concentrations in sedentary overweight men. *The American journal of clinical nutrition, 80*(2), 308-316. http://dx.doi.org/10.1093/ajcn/80.2.308.

Crouter, S. E., Kuffel, E., Haas, J. D., Frongillo, E. A., & Bassett, D. R., Jr. (2010). Refined two-regression model for the ActiGraph accelerometer. *Med Sci Sports Exerc, 42*(5), 1029-1037. http://dx.doi.org/10.1249/MSS.0b013e3181c37458.

Crouter, S. E., Schneider, P. L., & Bassett, D. R., Jr. (2005). Spring-levered versus piezo-electric pedometer accuracy in overweight and obese adults. *Med Sci Sports Exerc, 37*(10), 1673-1679. http://dx.doi.org/10.1249/01.mss.0000181677.36658.a8.

da Silva-Hamu, T. C. D., Formiga, C. K. M. R., Gervasio, F. M., Ribeiro, D. M., Christofoletti, G., & de Franca Barros, J. (2013). The impact of obesity in the kinematic parameters of gait in young women. *International journal of general medicine, 6*, 507-513. http://dx.doi.org/10.2147/ijgm.s44768.

Department of Health. (2019). *A report from the Chief Medical Officers in the UK on the amount and type of physical activity people should be doing to improve their health.* https://www.gov.uk/government/publications/physical-activity-guidelines-uk-chief-medical-officers-report.

Ellis, K., Kerr, J., Godbole, S., Staudenmayer, J., & Lanckriet, G. (2016). Hip and Wrist Accelerometer Algorithms for Free-Living Behavior Classification. *Medicine and Science in Sports and Exercise, 48*(5), 933-940. http://dx.doi.org/10.1249/mss.0000000000000840.

Farrahi, V., Niemelä, M., Kangas, M., Korpelainen, R., & Jämsä, T. (2019). Calibration and validation of accelerometer-based activity monitors: A systematic review of machine-learning approaches. *Gait & Posture, 68*, 285-299. http://dx.doi.org/https://doi.org/10.1016/j.gaitpost.2018.12.003.

Gao, L., Bourke, A. K., & Nelson, J. (2014). Evaluation of accelerometer based multi-sensor versus single-sensor activity recognition systems. *Medical Engineering & Physics, 36*(6), 779-785. http://dx.doi.org/https://doi.org/10.1016/j.medengphy.2014.02.012.

Goode, A. P., Hall, K. S., Batch, B. C., Huffman, K. M., Hastings, S. N., Allen, K. D., . . . Kosinski, A. S. (2016). The impact of interventions that integrate accelerometers on physical activity and weight loss: a systematic review. *Annals of Behavioral Medicine, 51*(1), 79-93. http://dx.doi.org/10.1007/s12160-016-9829-1.

Gupta, P., & Dallas, T. (2014). Feature selection and activity recognition system using a single triaxial accelerometer. *IEEE Trans Biomed Eng, 61*(6), 1780-1786. http://dx.doi.org/10.1109/tbme.2014.2307069.

Hartman, S. J., Nelson, S. H., Cadmus-Bertram, L. A., Patterson, R. E., Parker, B. A., & Pierce, J. P. (2016). Technology-and phone-based weight loss intervention: pilot RCT in women at elevated breast cancer risk. *American journal of preventive medicine, 51*(5), 714-721. http://dx.doi.org/10.1016/j.amepre.2016.06.024.

He, Z., & Jin, L. (2009). *Activity recognition from acceleration data based on discrete consine transform and SVM.*IEEE International Conference on Systems, Man and Cybernetics. http://dx.doi.org/10.1109/ICSMC.2009.5346042.

Healy, G. N., Winkler, E. A., Brakenridge, C. L., Reeves, M. M., & Eakin, E. G. (2015). Accelerometer-derived sedentary and physical activity time in overweight/obese adults with type 2 diabetes: cross-sectional associations with cardiometabolic biomarkers. *PloS one, 10*(3), e0119140. http://dx.doi.org/10.1371/journal.pone.0119140.

Howe, C. C. F., Moir, H. J., & Easton, C. (2017). Classification of Physical Activity Cut-Points and the Estimation of Energy Expenditure During Walking Using the GT3X+ Accelerometer in Overweight and Obese Adults. *Measurement in Physical Education and Exercise Science, 21*(3), 127-133. http://dx.doi.org/10.1080/1091367x.2016.1271801.

Janssen, I., Fortier, A., Hudson, R., & Ross, R. (2002). Effects of an energy-restrictive diet with or without exercise on abdominal fat, intermuscular fat, and metabolic risk factors in obese women. *Diabetes care, 25*(3), 431-438. http://dx.doi.org/10.2337/diacare.25.3.431.

Kiernan, M., King, A. C., Stefanick, M. L., & Killen, J. D. (2001). Men Gain Additional Psychological Benefits by Adding Exercise to a Weight‐Loss Program. *Obesity Research, 9*(12), 770-777. http://dx.doi.org/10.1038/oby.2001.106.

Lai, P. P. K., Leung, A. K. L., Li, A. N. M., & Zhang, M. (2008). Three-dimensional gait analysis of obese adults. *Clinical Biomechanics, 23*, S2-S6. http://dx.doi.org/10.1016/j.clinbiomech.2008.02.004.

Maurer, U., Smailagic, A., Siewiorek, D. P., & Deisher, M. (2006). *Activity recognition and monitoring using multiple sensors on different body positions.*International Workshop on Wearable and Implantable Body Sensor Networks. http://dx.doi.org/10.1109/BSN.2006.6.

Melanson, E. L., Knoll, J. R., Bell, M. L., Donahoo, W. T., Hill, J. O., Nysse, L. J., . . . Levine, J. A. (2004). Commercially available pedometers: considerations for accurate step counting. *Prev Med, 39*(2), 361-368. http://dx.doi.org/10.1016/j.ypmed.2004.01.032.

Moncada-Torres, A., Leuenberger, K., Gonzenbach, R., Luft, A., & Gassert, R. (2014). Activity classification based on inertial and barometric pressure sensors at different anatomical locations. *Physiol Meas, 35*(7), 1245-1263. http://dx.doi.org/10.1088/0967-3334/35/7/1245.

Parkka, J., Ermes, M., Korpipaa, P., Mantyjarvi, J., Peltola, J., & Korhonen, I. (2006). Activity classification using realistic data from wearable sensors. *IEEE Trans. Inf. Technol. Biomed., 10*(1), 119-128. http://dx.doi.org/10.1109/titb.2005.856863.

Plasqui, G. (2017). Smart approaches for assessing free-living energy expenditure following identification of types of physical activity. *Obesity Reviews, 18*, 50-55. http://dx.doi.org/10.1111/obr.12506.

Preece, S. J., Goulermas, J. Y., Kenney, L. P., Howard, D., Meijer, K., & Crompton, R. (2009). Activity identification using body-mounted sensors--a review of classification techniques. *Physiol Meas, 30*(4), R1-33. http://dx.doi.org/10.1088/0967-3334/30/4/R01.

Qi, J., Yang, P., Hanneghan, M., Tang, S., & Zhou, B. (2019). A Hybrid Hierarchical Framework for Gym Physical Activity Recognition and Measurement Using Wearable Sensors. *Ieee Internet of Things Journal, 6*(2), 1384-1393. http://dx.doi.org/10.1109/jiot.2018.2846359.

Rabbi, M., Pfammatter, A., Zhang, M., Spring, B., & Choudhury, T. (2015). Automated personalized feedback for physical activity and dietary behavior change with mobile phones: a randomized controlled trial on adults. *JMIR mHealth and uHealth, 3*(2), e42. http://dx.doi.org/0.2196/mhealth.4160.

Raz, I., Hauser, E., & Bursztyn, M. (1994). Moderate exercise improves glucose metabolism in uncontrolled elderly patients with non-insulin-dependent diabetes mellitus. *Israel journal of medical sciences, 30*(10), 766-770.

Ross, K. M., & Wing, R. R. (2016). Impact of newer self-monitoring technology and brief phone-based intervention on weight loss: A randomized pilot study. *Obesity (Silver Spring), 24*(8), 1653-1659. http://dx.doi.org/10.1002/oby.21536.

Sibella, F., Galli, M., Romei, M., Montesano, A., & Crivellini, M. (2003). Biomechanical analysis of sit-to-stand movement in normal and obese subjects. *Clinical Biomechanics, 18*(8), 745-750. http://dx.doi.org/10.1016/s0268-0033(03)00144-x.

Stewart, V., Ferguson, S., Peng, J., & Rafferty, K. (2012). *Practical automated activity recognition using standard smartphones.*IEEE International Conference on Pervasive Computing and Communications Workshops. http://dx.doi.org/10.1109/PerComW.2012.6197485.

Thomas, S., Reading, J., & Shephard, R. J. (1992). Revision of the Physical Activity Readiness Questionnaire (PAR-Q). *Can J Sport Sci, 17*(4), 338-345.

Tudor-Locke, C., Brashear, M. M., Johnson, W. D., & Katzmarzyk, P. T. (2010). Accelerometer profiles of physical activity and inactivity in normal weight, overweight, and obese US men and women. *International Journal of Behavioral Nutrition and Physical Activity, 7*. http://dx.doi.org/10.1186/1479-5868-7-60.

Wood, P. D., Stefanick, M. L., Dreon, D. M., Frey-Hewitt, B., Garay, S. C., Williams, P. T., . . . Vranizan, K. M. (1988). Changes in plasma lipids and lipoproteins in overweight men during weight loss through dieting as compared with exercise. *New England Journal of Medicine, 319*(18), 1173-1179. http://dx.doi.org/10.1056/NEJM198811033191801.

Tables

Table 1: Mean (SD) descriptive characteristics for the three BMI groups. Significant differences between the groups are indicated by *(p<0.05).

	All participants	Normal weight	Overweight	Obese
Number of participants	50	17	14	19
Gender	21M, 29F	9M, 8F	4M, 10F	8M, 11F
Age (years)	34.6 (11.2)	32.1 (10.1)	38.2 (14.4)	34.2 (9.0)
Height (m)	1.68 (.09)	1.72 (.09)	1.67 (.06)	1.67 (0.1)
Body Mass (kg)*	81.3 (16.7)	66.7 (8.6)	77.4 (4.5)	97.2 (14.0)
BMI (kg/m^2)*	28.7 (6.2)	22.5 (1.6)	27.7 (1.2)	35.0 (4.5)
BMI range	18.9 - 43.9	18.89 - 24.96	25.26 - 29.97	30.12 - 43.93
Body fat (%)*	30.2 (11.1)	21.0 (8.1)	31.3 (9.2)	37.5 (8.8)
Waist circumference (m)*	0.95 (0.15)	0.89 (0.07)	0.93 (0.06)	1.09 (0.12)
Hip circumference (m)*	1.06 (0.11)	0.96 (0.06)	1.04 (0.06)	1.16 (0.08)

Table 2: Overall classification accuracy for the four groups for each accelerometer site

Test dataset	Hip placement	Ankle placement
Entire Group (n=50)	85.0%	93.9%
Normal (n=17)	84.5%	91.6%
Obese (n=19)	81.9%	95.0%
Overweight (n=14)	90.0%	95.2%

Table 3: Confusion matrix for the hip placement for all participants.

Prediction / True activity	Walk	Downstairs	Upstairs	Cycling	Rowing	Cross-trainer	Jog	Stepping	Side-stepping	Side-stretching
Walk	1309	71	72	19	0	0	3	4	7	15
Downstairs	93	694	68	0	0	2	10	6	0	12
Upstairs	56	42	741	1	0	14	10	43	9	15
Cycling	9	0	22	1419	0	4	0	0	8	35
Rowing	0	0	1	39	1419	0	0	4	0	35
Cross-trainer	1	6	70	39	0	1249	10	0	79	41
Jog	59	19	0	0	0	0	1386	0	0	30
Stepping	18	34	105	9	0	0	0	1245	41	39
Sidestepping	1	17	26	46	0	5	0	26	1190	189
Side-stretching	0	0	2	213	0	0	0	36	173	1073

Table 4: Confusion matrix for the ankle placement for all participants.

Prediction / True activity	Walk	Downstairs	Upstairs	Cycling	Rowing	Cross-trainer	Jog	Stepping	Side-stepping	Side-stretching
Walk	1435	23	0	0	0	0	17	20	2	3
Downstairs	0	853	15	0	0	0	1	13	0	0
Upstairs	0	16	821	1	0	1	0	139	0	0
Cycling	0	0	0	1372	0	120	0	0	0	0
Rowing	0	0	0	0	1496	0	0	0	0	0
Cross-trainer	0	0	0	16	30	1452	0	0	0	0
Jog	20	6	0	0	0	0	1469	0	0	0
Stepping	0	23	148	0	0	1	0	1321	5	2
Sidestepping	0	0	0	0	0	0	0	4	1352	144
Side-stretching	0	0	0	0	0	8	0	1	60	1414

Figures

Figure 1: Sensitivity results for each activity across the different BMI groups for the (a) ankle and (b) hip accelerometer site.