A Machine Learning Approach for Physical Activity Recognition in Cystic Fibrosis

Silveira Bianchim, Mayara; McNarry, Melitta A.; Barker, Alan; Williams, Craig; Denford, Sarah; Thia, Lena; Evans, Rachel ; Mackintosh, Kelly

Measurement in Physical Education and Exercise Science

DOI:
https://doi.org/10.1080/1091367X.2023.2271444

E-pub ahead of print: 24/10/2023

Publisher's PDF, also known as Version of record

Measurement in Physical Education and Exercise Science

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/hmpe20

A Machine Learning Approach for Physical Activity Recognition in Cystic Fibrosis

Mayara S. Bianchim, Melitta A. McNarry, Alan R. Barker, Craig A. Williams, Sarah Denford, Lena Thia, Rachel Evans & Kelly A Mackintoshon behalf of ActiveYouth SRC group

To cite this article: Mayara S. Bianchim, Melitta A. McNarry, Alan R. Barker, Craig A. Williams, Sarah Denford, Lena Thia, Rachel Evans & Kelly A Mackintoshon behalf of ActiveYouth SRC group (24 Oct 2023): A Machine Learning Approach for Physical Activity Recognition in Cystic Fibrosis, Measurement in Physical Education and Exercise Science, DOI: 10.1080/1091367X.2023.2271444

To link to this article: https://doi.org/10.1080/1091367X.2023.2271444

MEASUREMENT IN PHYSICAL EDUCATION AND EXERCISE SCIENCE
https://doi.org/10.1080/1091367X.2023.2271444

A Machine Learning Approach for Physical Activity Recognition in Cystic Fibrosis

Mayara S. Bianchim[a,b], Melitta A. McNarry[a], Alan R. Barker[c], Craig A. Williams[c], Sarah Denford[c], Lena Thia[d], Rachel Evans[e], Kelly A Mackintosh[a], and on behalf of ActiveYouth SRC group

[a]Applied Sports, Technology, Exercise and Medicine Research Centre, Swansea University, Swansea, UK; [b]School of Medical and Health Sciences, Bangor University, Bangor, UK; [c]Children's Health and Exercise Research Centre, University of Exeter, Exeter, UK; [d]Department of Paediatric Respiratory Medicine and Cystic Fibrosis Unit, Noah's Ark Children's Hospital for Wales, Cardiff, UK; [e]Paediatric Department, Morriston Hospital, Swansea, UK

ABSTRACT

This study aimed to develop and validate machine learning models to predict intensities in children and adolescents with cystic fibrosis (CF) across different accelerometry brands and placements. Thirty-five children and adolescents with CF (11.6 ± 2.8 yrs; 15 girls) and 28 healthy youth (12.2 ± 2.7 yrs; 16 girls) performed six activities whilst wearing GENEActivs (both wrists) and ActiGraphs GT9X (both wrists and waist). Three supervised learning classifiers (K-Nearest Neighbour, Random Forest and eXtreme Gradient Boosted Decision Tree) were used to identify the input signal pattern for each PA type and intensity, with a 10-fold cross-validation utilized to assess the performance of the classifiers. ActiGraph GT9X on the dominant wrist and waist and GENEActiv on the dominant wrist failed to predict vigorous intensity PA activities. All other models, for activity type and intensities, exceeded 97% accuracy, with a sensitivity and specificity of greater than 95%, irrespective of accelerometer brand, placement or health condition.

KEYWORDS

Threshold; Physical Activity; ENMO; MAD; youth

Introduction

Cystic Fibrosis (CF) is the most common life-limiting autosomal recessive disorder in the Caucasian population affecting over 10,500 individuals in the United Kingdom (Trust, 2018).

Limited exercise tolerance in CF is multifactorial and can promote physical inactivity which has significant negative health implications (González et al., 2017). Habitual physical activity (PA), and particularly moderate-to-vigorous physical activity (MVPA), is associated with significant health benefits in youth with CF (Hebestreit et al., 2014), including an increased lifespan and better quality of life (van de Weert van Leeuwen et al., 2013). However, the traditionally derived accelerometer cut-points on which these conclusions, and thus national and international treatment guidelines, are based are associated with limited predictive accuracy and are prone to the misclassification of PA intensities, particularly in clinical populations (Bianchim et al., 2023). Previous research has relied on accelerometer cut-points or prediction equations due, at least in part, to their practicality and simplicity. However, it is pertinent to note that they are highly specific to the population, activities, and accelerometer device and settings on which they were developed (Bassett et al., 2012). Children with CF expend more energy during rest and activities than their peers, due to impaired metabolic and ventilatory responses (Bianchim et al., 2022; Johnson et al., 2006), which means that cut-points and equations are likely to underestimate PA levels in those with CF (Stephens et al., 2016). Indeed cut-points developed in healthy children and adolescents are likely to underestimate PA levels in those with CF (Stephens et al., 2016).

Traditionally, cut-points have been developed using linear methods, resulting in poor prediction when applied to estimate non-linear data, such as PA (Trost et al., 2012). Recent technological advances have facilitated the application of machine learning to non-linear accelerometry data, enabling the analysis of complex accelerometry patterns to identify activity types or PA intensities with excellent accuracy. Whether machine learning approaches can also enhance the prediction of PA and sedentary time (SED) in children and adolescents with clinical conditions, such as CF which may have a higher energy expenditure (EE) for a given activity relative to their healthy peers (Stephens et al., 2016), remains to be elucidated.

CONTACT Melitta A. McNarry M.McNarry@swansea.ac.uk Applied Sports, Technology, Exercise and Medicine Research Centre, Swansea University, Swansea, UK

Supplemental data for this article can be accessed online at https://doi.org/10.1080/1091367X.2023.2271444

The optimal accelerometer placement for PA assessment remains equivocal, with some evidence that machine learning enhances the prediction of EE irrespective of placement (Trost et al., 2014), whereas others found that waist placement provides higher accuracy in healthy children (Mackintosh et al., 2016). Furthermore, PA predictions are reported to vary between placements according to accelerometer brand (Fairclough et al., 2016) and the use of raw or count-based data (Trost et al., 2018). Indeed, two machine learning classifiers derived from raw acceleration data were recently reported to achieve an overall accuracy of 87.5–99.6%, considerably higher than that reported for the models trained on the respective count-based data (57–86%; de Vries et al., 2011; Ruch et al., 2011; Trost et al., 2012). Nonetheless, the majority of studies in healthy children using different classifiers to predict PA types still rely on count-based data despite the low accuracy associated with this approach (de Vries et al., 2011; Ruch et al., 2011; Trost et al., 2012).

The primary aim of this study was to develop and validate machine learning models to predict PA intensities in children and adolescents with CF using raw acceleration data. The secondary aim was to investigate how these predictions vary according to accelerometer brand and placement.

Materials and methods

Participants

Participants were classified as having CF if they presented with CF-typical symptoms and had either two pathological sweat tests or the identification of two CF-relevant mutations. Sixty-four children and adolescents (35 with CF) participated. One participant was excluded from the analysis for not attending all visits, resulting in 63 (35 CF) participants being included in the analyses. Participants with CF were mainly homozygous (55%) for the *ΔF508* mutation. Children and adolescents with CF (7–18 years) were recruited from pediatric CF outpatient clinics in South Wales. The primary respiratory consultant confirmed the suitability of each patient prior to recruitment. Participants in the healthy group (7–18 years) were recruited through university networks and from the friends and families of the CF participants (Table 1). Their health status was confirmed by a short clinical evaluation to identify the presence of any clinical conditions or medications. Written parental consent and child assent were obtained from all parents/guardians and participants, respectively. This study received ethics approval through the National Health Service Research Ethics Committee (18/WS/0032).

Protocol

Participants completed six activities across three separate visits, with the first two visits separated by seven days. The first visit consisted of the assessment of demographic outcomes and health indicators (anthropometrics, RMR and lung function). The second and third visits were scheduled within a 48-hour gap of each other and comprised of a calibration protocol performed in a laboratory setting. All protocols were completed in the same order. Participants were advised to avoid caffeine and vigorous exercise 24 hours prior to all visits and to arrive at least two hours post-postprandial.

Table 1. Participants characteristics.

	CF			Healthy		
	Total ($n = 35$)	Girls ($n = 15$)	Boys ($n = 20$)	Total ($n = 28$)	Girls ($n = 16$)	Boys ($n = 12$)
Age (years)	11.6 ± 2.8	11.3 ± 2.7	11.8 ± 2.9	12.2 ± 2.7	12.6 ± 2.6	11.5 ± 2.8
Pubertal Stage	23	9	14	15	7	8
Pre-PHV						
Circa-PHV	8	5	3	2	1	1
Post-PHV	4	2	2	11	9	2
Stature (m)	1.46 ± 0.15	1.44 ± 0.12	1.47 ± 0.17	1.53 ± 0.16	1.54 ± 0.10	1.50 ± 0.21$^{\times}$
Body Mass (kg)	39.13 ± 12.	37.3 ± 10.2	40.4 ± 14.2	47.1 ± 15.0	50.1 ± 12.7	43.0 ± 12.2
BMI ($kg{\cdot}m^{-2}$)	18.0 ± 4.2	17.5 ± 2.0	18.4 ± 5.3	19.6 ± 3.5	20.6 ± 3.3	18.2 ± 3.5
zBMI	−0.31 ± 1.10^{+}	−0.12 ± 0.78	−0.47 ± 1.28	0.41 ± 0.80	0.57 ± 0.62	0.19 ± 1.00
RMR ($kcal{\cdot}day^{-1}$)	1,687 ± 480	1,510 ± 257	1,820 ± 566*	1,700 ± 413	1,586 ± 373	1,852 ± 431
RMR ($ml{\cdot}kg^{-1}{\cdot}min^{-1}$)	6.21 ± 1.31	5.86 ± 1.26	6.45 ± 1.24	5.35 ± 1.54	4.51 ± 0.89	6.47 ± 1.51
FVC (L)	2.5 ± 1.0	2.5 ± 1.0	2.5 ± 1.0*	2.8 ± 1.0	2.8 ± 1.0	2.9 ± 1.2$^{\times}$
FVC predicted (%)	99 ± 21	97 ± 20	99 ± 21	105 ± 26	105 ± 26	106 ± 18
FEV_1 (L)	2.0 ± 0.7	2.0 ± 0.8	2.1 ± 0.8*	2.4 ± 0.8	2.4 ± 0.8	2.4 ± 0.9
FEV_1 predicted (%)	94 ± 19	92 ± 20	94 ± 19	99 ± 21	99 ± 22	100 ± 14

Data are presented as mean ± SD. CF: Cystic Fibrosis, RMR: resting metabolic rate, FEV_1: forced expiratory volume in one second, $FEV_1\%_{predicted:}$ forced expiratory volume in one second, BMI: body mass index, zBMI: z-scores body mass index, PHV: peak height velocity. $^{+}$indicates significant difference between healthy and CF; $^{\times}$ indicates significant difference between boys and girls in the healthy group; *indicates significant difference between boys and girls in the CF group ($p \leq .05$).

Measurements

Anthropometry

Body mass (Seca 876, Hamberg, Germany), stature (Holtain Stadiomerter 603VR, Holtain Ltd, UK) and sitting height (Holtain Sitting Height Stadiometer 607VR, Holtain Ltd, UK) were determined to the nearest 0.1 kg, 0.1 cm and 0.1 cm, respectively. Body mass index (BMI) and age- and sex-specific z-scores were calculated according to the World Health Organization reference data (de Onis et al., 2004). Pubertal stage was estimated according to time pre or post the age of peak height velocity (PHV; Mirwald et al., 2002) with pre-pubertal considered > −1 years from PHV, pubertal as −1 to +1 years and post-pubertal as > +1 years post PHV.

Resting metabolic rate

Resting metabolic rate (RMR) was assessed in the supine position for 20 minutes using an online gas analyzer (MetaMax Cortex 3B, CORTEX Biophysik GmbH, Germany). The analyzer was calibrated according to the manufacture's guidelines prior to each measurement. The measurement started following at least 10 minutes at rest, and participants were required to stay awake throughout the test, with noise kept to a minimum. To ensure a steady state was achieved prior to the averages being derived, the first two and a half minutes of the recording were discarded, with remaining breath-by-breath values of oxygen uptake ($\dot{V}O_2$) and carbon dioxide production ($\dot{V}CO_2$) included in the analyses (Cooper et al., 2009). The Weir equation was then used to calculate the RMR (Weir, 1949).

Lung function

Lung function was assessed by standard spirometry (MetaMax 3B, Cortex Biophysik GmbH, Germany) using a forced vital capacity maneuver in accordance with the American Thoracic Society and European Respiratory Society standards (Graham et al., 2019; Moore, 2012). Percentage predicted lung function was estimated using a reference equation for age, sex and height (Quanjer et al., 2012). Disease severity was classed according to percentage of estimated forced expiratory volume in one second (FEV_1%) as mild (>70%), moderate (40–69%), or severe (<40%; Davies & Alton, 2009).

Accelerometry

Five monitors were used during the activities; three GT9X monitors (ActiGraph, Pensacola, FL) worn on each wrist and the right hip, and two GENEActiv monitors (ActivInsights Ltd., Cambridge, UK), one on each wrist (Filanowski et al., 2022).

Activity protocol

The activity protocol was designed to include activities replicating the participant's daily lives using public and patient involvement (PPI). Specifically, five participants with CF selected by the physiotherapist and 56 healthy children and adolescents completed a survey of common activities from the youth compendium of physical activities (Butte et al., 2018) and were asked to select any that they would typically do during their normal routine. Participants were encouraged to suggest any additional activities that were not listed. The six activities with the highest votes, stratified by intensity and behavior type, were selected to be included in the protocol. Each activity was conducted in a random order for three to ten minutes (Table 2), interspersed with at least three minutes rest. Activities were performed whilst wearing the accelerometers, a metabolic analyzer (MetaMax Cortex 3B, CORTEX Biophysik GmbH) and a pulse oximeter (Nonin® WristOx® Model 3150, Nonin® Medical Inc.).

Data processing and feature extraction

Data from all accelerometers were processed in the same manner. The raw acceleration data were extracted at 100 Hz as.gt3× and.bin files for GT9X (ActiLife V6.10.2) and GENEActiv (GENEActiv PC software V2.2), respectively. All.gt3× files were converted to

Table 2. Activities included in the activity protocol.

Activity	Duration (min)	Description	METy by age-group (years)			
			6–9	10–12	13–15	16–18
Video	10	Watching a video whist sitting	1.4	1.3	1.3	1.2
Colouring/writing	6	Colouring or writing whist sitting	1.8	1.7	1.7	1.7
Handheld device	6	Playing games on the handheld device whist sitting	1.4	1.5	1.5	1.5
Games	6	Playing a variety of self-selected games including football, tennis, badminton, rugby, skipping and mini bowling	6.0	6.2	6.3	6.5
Walking	5	Walking continuously at a self-selected brisk pace	4.6	4.9	5.1	5.4
Climbing stairs	3	Climbing stairs continuously at a self-selected pace	5.5	6.3	7.0	7.7

METy: MET youth developed by McMurray et al. (2015).
MET values extracted from the Youth Compendium of Physical Activities Butte et al. (2018).

time-stamp-free.csv files, and imported with the.bin files into R V3.1.2 (R Foundation for Statistical Computing, Vienna, Austria), which was used for all subsequent analyses. Raw accelerometry data were then auto-calibrated and the x, y and z axes extracted in 5s epochs using the "GGIR" package V1.2–0 (Matthews et al., 2012; Migueles et al., 2019; Vähä-Ypyä et al., 2015). Visual screening tools, such as plots and histograms, were utilized to identify any traits or missing data, and features were extracted from the vector magnitude. Specifically, sliding windows of 1.5 s were created and the components were split into low- and high-frequency using a cutoff of 6 Hz, according to previous recommendations (Zalewski et al., 2020). This is particularly important given the dynamic nature of the signal extracted from the accelerometer. Subsequently, nine time-domain components were calculated for each window using data from the three axes. Specifically, mean, standard deviation, peak-to-peak value, root mean squared value, kurtosis, skewness, crest factor, root mean square velocity and signal entropy were extracted. Metabolic equivalent of task (MET) values were calculated for each activity using the RMR, which was previously described as Youth MET (METy; McMurray et al., 2015). The first and last minute of each activity were excluded to avoid transitional movements (Hurter et al., 2018; Migueles et al., 2021). MET values were subsequently aligned with the raw acceleration data and used to predict PA intensities as sedentary (≤1.5 METs), light (>1.5 to < 3 METs), moderate (≥3 to < 6 METs) or vigorous (≥6; 39).

Machine learning modelling

Three supervised learning classifiers, K-Nearest Neighbour (k-NN), Random Forest and eXtreme Gradient Boosted Decision Tree (XGBoost Decision Trees), were used to identify the input signal pattern for each PA intensity (Friedman, 2001; Kuhn et al., 2013; Patrick & Fischer, 1970; Zertuche, 2014). All models were trained and cross-validated using packages within R ("caret," "randomForest," "xboost," "entropy," "signal" and "knn"). Models were used to identify different PA intensities according to METs. The random forest classifier of 500 trees was trained using the data from all nine features. Specifically, the features were randomly sampled into training and test sets and the whole process repeated 1,000 times. An internal out of bag approach (Winham et al., 2013) was used to test the model accuracy. A decision tree learns from a subset of the data, enabling the remaining data to be used to evaluate the performance of the model (Winham et al., 2013).

For the XGBoost model, the data set was randomly split, with 80% of data used for its own training and 20% for testing. This model evaluates the performance of each round of classification instead of assessing the overall performance of the training set. Specifically, XGBoost is a type of boosting algorithm designed to learn from previous poor predictions in order to use this information to enhance future predictions (Chen & Guestrin, 2016). For this model, 15 consecutive rounds of classification decline were determined prior to halting the learning, with the last best score used as the final outcome. Finally, the weighted k-NN was performed using a kernel function to weight the neighbors of a data point using the distance as a parameter (Zhang, 2016).

A 10-fold cross-validation (Little et al., 2017) was utilized to assess the performance of all classifiers. The average of the results was used to indicate the accuracy of the model. In addition, the percentage agreement, 95% confidence intervals and kappa scores were calculated. The detection rate (also called true positive rate) was calculated as the proportion of the sample where events were detected correctly. The balanced accuracy was calculated for each model and compared using the McNemar test in SPSS version 29.0 (IBM Corp., USA). Statistical significance was accepted when $p \leq .05$.

Results

CF was classified as mild for all 63 (35 CF) participants. There were no differences in age, stature, body weight or RMR between the CF and healthy participants.

XGBoost and k-NN algorithms were reported as one outcome; they both achieved the same performance and provided the same classification values. Confusion matrices are provided as supplementary material for Random Forest (Supplementary Table S1) and XGBoost/k-NN models (Supplementary Table S2). Models using features extracted from the GT9X worn on the dominant wrist and waist and GENEActiv worn on the dominant wrist failed to predict VPA in those with CF, whilst the GENEActiv on the non-dominant wrist failed to recognize patterns related to vigorous intensity in healthy children.

As shown in Table 3, those with CF had higher EE than healthy participants while watching a video, whist sitting, and climbing stairs. It is noteworthy that the healthy group did not reach VPA (≥6 METs) for any of the activities.

All models provided excellent accuracy (97–100%) and low error (Table 4). Comparisons demonstrated that XGBoost and k-NN yielded statistically higher accuracy ($p = .02$) than Random Forest, and

Table 3. Metabolic equivalent of task during each activity from the activity protocol.

	METs	
	CF ($n = 35$)	Healthy ($n = 28$)
Video	1.3* ±0.9	1.0 ± 0.3
Colouring/writing	1.3 ± 0.3	1.2 ± 0.5
Handheld Device	1.1 ± 0.3	1.1 ± 0.4
Games	4.1 ± 2.0	4.2 ± 1.8
Walking	2.9 ± 1.3	2.5 ± 1.1
Stairs	5.1* ±2.0	4.7 ± 0.7

Data are presented as mean ± SD. CF: Cystic Fibrosis, METs: metabolic equivalent of task. *indicates significant difference between groups ($p \leq .05$).

Table 4. Performance metrics (%) of different models to classify physical activity intensities.

		CF				Healthy			
	Placement	Sensitivity	Specificity	Accuracy	Detection Rate	Sensitivity	Specificity	Accuracy	Detection Rate
Random Forest									
SED	GE non-dominant wrist	99.7	100	99.8	65.1	100	100	100	49.6
	GE dominant wrist	100	100	100	60.6	100	100	100	41.2
	AG non-dominant wrist	100	100	100	63.5	100	100	100	77.4
	AG dominant wrist	100	99.5	99.7	53.2	98.5	98.3	98.4	43.7
	AG waist	100	100	99.8	42.9	100	99.4	99.7	67.6
LPA	GE non-dominant wrist	100	99.7	99.8	23.8	100	100	100	27.3
	GE dominant wrist	100	100	100	12.5	100	100	100	41.5
	AG non-dominant wrist	100	100	100	15.6	100	100	100	17.9
	AG dominant wrist	99.0	100	99.5	23.1	97.7	98.4	98.1	39.1
	AG waist	100	100	100	44.4	99.1	100	99.5	19.6
MPA	GE non-dominant wrist	100	100	100	9.6	100	100	100	20.7
	GE dominant wrist	100	100	100	24.2	100	100	100	9.9
	AG non-dominant wrist	100	100	100	19.0	100	100	100	4.6
	AG dominant wrist	100	100	100	23.4	95.8	99.5	97.6	14.9
	AG waist	100	100	100	12.6	100	100	100	4.9
VPA	GE non-dominant wrist	100	100	100	1.2	100	100	100	2.3
	GE dominant wrist	100	100	100	2.5	100	100	100	7.2
	AG non-dominant wrist	100	100	100	1.7	0	0	0	0
	AG dominant wrist	0	0	0	0	0	0	0	0
	AG waist	0	0	0	0	100	100	100	7.6
XGBoost/k-NN									
SED	GE non-dominant wrist	100	100	100	65.5	100	100	100	50.0
	GE dominant wrist	100	100	100	60.9	100	98.1	99.0	40.4
	AG non-dominant wrist	100	100	100	64.0	100	100	100	78.7
	AG dominant wrist	100	100	100	53.7	100	100	100	44.7
	AG waist	100	100	100	42.9	100	100	100	68.9
LPA	GE non-dominant wrist	100	100	100	24.0	100	100	100	27.0
	GE dominant wrist	100	100	100	12.6	97.3	100	98.6	41.5
	AG non-dominant wrist	100	100	100	15.6	100	100	100	17.7
	AG dominant wrist	100	100	100	22.7	100	100	100	40.3
	AG waist	100	100	100	44.4	100	100	100	19.5
MPA	GE non-dominant wrist	100	100	100	24.0	100	100	100	20.8
	GE dominant wrist	100	100	100	24.1	100	100	100	10.1
	AG non-dominant wrist	100	100	100	18.7	100	100	100	4.1
	AG dominant wrist	100	100	100	23.4	100	100	100	14.9
	AG waist	100	100	100	12.6	100	100	100	4.5
VPA	GE non-dominant wrist	100	100	100	1.0	100	100	100	2.0
	GE dominant wrist	100	100	100	2.2	100	100	100	6.7
	AG non-dominant wrist	100	100	100	1.5	0	0	0	0
	AG dominant wrist	0	0	0	0	0	0	0	0
	AG waist	0	0	0	0	100	100	100	6.8

CF: Cystic Fibrosis, GE: GENEActiv, AG: ActiGraph, SED: sedentary, LPA: light physical activity, MPA: moderate physical activity, VPA: vigorous physical activity, XGBoost: eXtreme Gradient Boosting Trees; k-NN: k-Nearest Neighbour; NA: non-applicable (not enough data to validate the models).

GENEActiv demonstrated higher accuracy than ActiGraph ($p = .002$). We found no differences in prediction accuracy between CF and healthy groups, or between placements (Table 4). Finally, all classifications presented sensitivity and specificity higher than 95%, independent of accelerometer brand, placement and health status for all models performed.

Discussion

This study provides novel insights into the suitability of using three machine learning classifiers with raw accelerometry data to predict PA intensities in children and adolescents with CF and their healthy peers. Overall, the use of machine learning yielded high accuracy (>97%) to predict different PA intensities when trained on time-domain features, irrespective of accelerometer brand or placement. Although it is important to note that none of the healthy children reached VPA, which affected the prediction of this intensity. Future studies evaluating PA in CF should consider using raw acceleration data with machine learning algorithms. A cross-validation of the models in a larger sample within a free-living setting is warranted.

Accelerometry data

Akin to previous studies using machine learning on raw accelerometry data (Ahmadi et al., 2020; Kühnhausen et al., 2017), all three classifiers trained and tested in this study outperformed previous models trained on accelerometry counts. Studies comparing count based data to raw data have yielded conflicting results regarding accuracy. For instance, Kühnhausen et al. (2017) found a significantly higher accuracy (92.7%) for predicting PA in healthy children using machine learning models developed from raw accelerometry data in comparison to those using counts (70.9–71.2%). These discrepancies may be atributtable to the elimination of vital information during the data reduction process to transform raw accelerometry data into counts (Kühnhausen et al., 2017). However, others have achieved higher accuracy with count-based in comparison to raw data (Clevenger et al., 2020; Montoye et al., 2019). Moreover, recently developed harmonization methods can aid comparison of different accelerometry metrics facilitating transferability between raw and count-based data findings (Karas et al., 2022).

Machine learning classifiers

Whilst the three activity classifiers tested in the present study demonstrated similar overall performance, suggesting that any of these models could be used to predict PA in children with CF or, indeed, healthy children, the XGBoost and k-NN achieved higher sensitivity and classification accuracy. This is contrary to previous studies in healthy children, which found that Random Forest performed marginally better to predict ambulatory activities in comparison with SED and LPA (Ahmadi et al., 2020). It is also noteworthy that children with mild CF might present metabolic adaptations and an altered muscle function that could impact the predictions of EE during walking and climbing stairs (Erickson et al., 2015). Indeed, differences in accelerometry raw outputs were observed between CF and healthy participants during walking (Bianchim et al., 2023). XGBoost is a type of boosting algorithm designed to learn from previous weak predictions, adjusting the future predictions accordingly (Friedman, 2001). Unsurprisingly, this study demonstrated that XGBoost enhanced all of the Random Forest predictions, independent of intensity. This finding is corroborated by previous research demonstrating that Gradient Boosting classifiers perform better compared to other types of boosting classifiers for PA classification (Rahman et al., 2020).

Accelerometer brand and placement

All three placements yielded comparable overall accuracy across classifiers, in line with previous research in healthy children (Mackintosh et al., 2016). This finding differs from previous research (Trost et al., 2018) reporting that Random Forest models achieved marginally higher activity classification accuracy for GT3X+ placed at the waist in comparison with wrist in healthy children and adults. Interestingly, the waist-worn GT9X also performed marginally better than the wrist to predict PA in children with cerebral palsy using different machine learning models (Ahmadi et al., 2018). Nevertheless, wrist-worn accelerometry protocols are associated with enhanced compliance (Fairclough et al., 2016), which is extremely important in studies assessing PA in free-living conditions.

It is important to highlight that GT9X placed on the dominant wrist and waist failed to predict VPA in those with CF. Similar results were found for models using features extracted from the GT9X placed on dominant and non-dominant wrist in healthy children. Visual inspection suggested that whilst the models failed to predict VPA in healthy participants due to the scarcity of data points associated with this intensity, the same was not observed in those with CF. This discrepancy therefore indicates that whilst intended vigorous components of the protocol were classified as moderate intensity for healthy participants, they were classified as vigorous for the majority of the CF group (see Table 3). Models using data from GENEActiv models predicted VPA in both groups, albeit with low detection rates. Indeed, GENEActiv models performed with higher accuracy overall when compared to GT9X. This not only reiterates the need for CF-specific approaches to evaluate PA, but highlights the importance of

considering placement and accelerometer brand when measuring PA. Therefore, future studies should further investigate how accelerometer brand and placement can affect the measurement of PA when using machine learning.

Disease specific factors

Youth with CF are known to require higher EE during rest due to the enhanced cost of breathing and higher RMR in comparison with their healthy peers (Tomezsko et al., 1994), although the latter was not observed in this study. While this might raise the question of whether a specific model is warranted in mild CF, it is important to acknowledge that other factors associated with exercise intolerance, such as chronic inflammation and impaired muscle metabolism, were not controlled for (Erickson et al., 2015).

Recent studies have demonstrated the feasibility of using machine learning in adults to characterize sport-related PA (Pfeiffer et al., 2023) and to describe characteristics of overall sedentary behavior and sitting (Tjurin et al., 2019) with varying degrees of success. Building on these studies (Pfeiffer et al., 2023; Tjurin et al., 2019), which have focused on characterizing specific aspects of the PA spectrum in healthy adults. The present study demonstrated the feasibility of using machine learning models to accurately predict SED and PA from EE in youth with CF. These findings have significant importance for clinical practice, with PA recognized as a valuable component of CF treatment (Rand & Prasad, 2012). Machine learning algorithms could be used to identify daily patterns of PA in youth with CF, advancing our understanding of the association between daily behaviors (PA and sleep patterns) and health outcomes. This study has the potential to contribute to the design of PA interventions and specific recommendations in this population.

Strengths and limitations

This is the first study to utilize machine learning models to identify PA intensities using raw acceleration data from both wrist- and waist-worn accelerometers in youth with CF, with a 10-fold cross validation which has a lower variance than a leave-one-out cross validation. Moreover, all activities incorporated in the study protocol were selected by the participants through an initial survey, ensuring ecological validity. Additionally, comparisons between models and features across multiple accelerometer placements and brands were made, with RMR individually assessed. There were, however, several limitations. Specifically, all activities were performed in a structured laboratory setting and might not be representative of free-living conditions and did not include sleep. It is also noteworthy that none of the healthy participants reached vigorous intensity during the protocol. Furthermore, whilst this study did not include frequency domain features in accord with previous recommendations (Montoye et al., 2018), this omission might hinder inter-study comparisons. However, previous research has demonstrated that the inclusion of frequency domain features in the activity models does not improve the overall accuracy of the predictions and can lead to overfitting (Montoye et al., 2018). It is important to highlight that previous studies might have evaluated the performance of their algorithms differently and used different parameters to predict PA levels. Finally, this study included youth with mild CF and, therefore, might not be representative of those with more severe forms of the condition.

Conclusion

This study demonstrated the feasibility of using three different machine learning classifiers to estimate different PA intensities from waist- and wrist-mounted ActiGraph GT9X and GENEActiv accelerometers in youth with CF. The accuracy achieved in this study was comparable or higher than studies in youth using various machine learning classifiers. Thus, this study provides support for the use of machine learning to predict complex pattern variables such as PA in youth with CF. Future studies assessing PA levels in those with CF should consider using raw acceleration data with machine learning algorithms to enhance prediction accuracy.

Acknowledgments

We acknowledge Mari Powell for her support during the recruitment and testing stages of this study, and all participants that have volunteered to this study. The Cystic Fibrosis Trust UK funded the study (grant reference No RP-PG-0108-10011), and MSB PhD. The funder did not participate in the conduct of the study, the writing of the manuscript, or the decision to submit it for publication.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by the Cystic Fibrosis Trust UK under its programme grant for Strategic Research Centres (grant reference number RP-PG-0108-10011).

ORCID

Melitta A. McNarry http://orcid.org/0000-0003-0813-7477
Kelly A Mackintosh http://orcid.org/0000-0003-0355-6357

Data availability statement

The datasets generated and/or analyzed during the current study are not publicly available due to GDPR regulations and to protect individual privacy but are available from the corresponding author on reasonable request.

References

Ahmadi, M., O'Neil, M., Fragala-Pinkham, M., Lennon, N., & Trost, S. G. (2018). Machine learning algorithms for activity recognition in ambulant children and adolescents with cerebral palsy. *Journal of Neuroengineering and Rehabilitation*, *15*(1), 1–9. https://doi.org/10.1186/s12984-018-0456-x

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*, *24*(2), 129–136. https://doi.org/10.1080/1091367X.2020.1716768

Bassett, D. R., Jr., Rowlands, A., & Trost, S. G. (2012). Calibration and validation of wearable monitors. *Medicine & Science in Sports & Exercise*, *44*(1 Suppl 1), S32–38. https://doi.org/10.1249/MSS.0b013e3182399cf7

Bianchim, M. S., McNarry, M. A., Barker, A. R., Williams, C. A., Denford, S., Holland, A. E., Cox, N. S., Dreger, J., Evans, R., Thia, L., & Mackintosh, K. A. (2022). Sleep, sedentary time and physical activity levels in children with Cystic Fibrosis. *International Journal of Environmental Research and Public Health*, *19*(12), 7133. https://doi.org/10.3390/ijerph19127133

Bianchim, M. S., McNarry, M. A., Evans, R., Thia, L., Barker, A. R., Williams, C. A., Denford, S., & Mackintosh, K. A. (2023). Calibration and cross-validation of accelerometry in children and adolescents with Cystic Fibrosis. *Measurement in Physical Education and Exercise Science*, *27*(1), 51–59. https://doi.org/10.1080/1091367X.2022.2065919

Butte, N. F., Watson, K. B., Ridley, K., Zakeri, I. F., McMurray, R. G., Pfeiffer, K. A., Crouter, S. E., Herrmann, S. D., Bassett, D. R., Long, A., Berhane, Z., Trost, S. G., Ainsworth, B. E., Berrigan, D., & Fulton, J. E. (2018). A youth compendium of physical activities: Activity codes and metabolic intensities. *Medicine & Science in Sports & Exercise*, *50*(2), 246–256. https://doi.org/10.1249/mss.0000000000001430

Chen, T., & Guestrin, C. (2016). Xgboost: A scalable tree boosting system. *KDD '16: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785–794. https://doi.org/10.1145/2939672.2939785

Clevenger, K. A., Pfeiffer, K. A., & Montoye, A. H. K. (2020). Cross-generational comparability of raw and count-based metrics from ActiGraph GT9X and wGt3X-BT accelerometers during free-living in youth. *Measurement in Physical Education and Exercise Science*, *24*(3), 194–204. https://doi.org/10.1080/1091367X.2020.1773827

Cooper, J. A., Watras, A. C., O'Brien, M. J., Luke, A., Dobratz, J. R., Earthman, C. P., & Schoeller, D. A. (2009). Assessing validity and reliability of resting metabolic rate in six gas analysis systems. *Journal of the American Dietetic Association*, *109*(1), 128–132. https://doi.org/10.1016/j.jada.2008.10.004

Davies, J. C., & Alton, E. W. (2009). Monitoring respiratory disease severity in Cystic Fibrosis. *Respiratory Care*, *54*(5), 606–617. https://doi.org/10.4187/aarc0493

de Onis, M., Onyango, A. W., Van den Broeck, J., Chumlea, W. C., & Martorell, R. (2004). Measurement and standardization protocols for anthropometry used in the construction of a new international growth reference. *Food and Nutrition Bulletin*, *25*(1 Suppl), S27–S36. https://doi.org/10.1177/15648265040251S104

de Vries, S. I., Engels, M., & Garre, F. G. (2011). Identification of children's activity type with accelerometer-based neural networks. *Medicine & Science in Sports & Exercise*, *43*(10), 1994–1999. https://doi.org/10.1249/MSS.0b013e318219d939

Erickson, M. L., Seigler, N., McKie, K. T., McCully, K. K., & Harris, R. A. (2015). Skeletal muscle oxidative capacity in patients with cystic fibrosis. *Experimental Physiology*, *100*(5), 545–552. https://doi.org/10.1113/ep085037

Fairclough, S. J., Noonan, R., Rowlands, A. V., Van Hees, V., Knowles, Z., & Boddy, L. M. (2016). Wear compliance and activity in children wearing wrist- and hip-mounted accelerometers. *Medicine and Science in Sports and Exercise*, *48*(2), 245–253. https://doi.org/10.1249/MSS.0000000000000771

Filanowski, P. M., Slade, E., Iannotti, R. J., Camhi, S. M., & Milliken, L. A. (2022). The impact of ActiGraph's low-frequency extension filter on measurement of children's physical activity. *Journal of Sports Sciences*, *40*(12), 1406–1411. https://doi.org/10.1080/02640414.2022.2081404

Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *Annals of Statistics*, *29*(5), 1189–1232. https://doi.org/10.1214/aos/1013203451

González, K., Fuentes, J., & Márquez, J. L. (2017). Physical inactivity, sedentary behavior and chronic diseases. *Korean Journal of Family Medicine*, *38*(3), 111–115. https://doi.org/10.4082/kjfm.2017.38.3.111

Graham, B. L., Steenbruggen, I., Miller, M. R., Barjaktarevic, I. Z., Cooper, B. G., Hall, G. L., Hallstrand, T. S., Kaminsky, D. A., McCarthy, K., McCormack, M. C., & Oropez, C. E. (2019). Standardization of spirometry 2019 update. An official American thoracic society and European respiratory society technical statement. *American Journal of Respiratory and Critial Care*, *200*(8), e70–e88. https://doi.org/10.1164/rccm.201908-1590ST

Hebestreit, H., Schmid, K., Kieser, S., Junge, S., Ballmann, M., Roth, K. . . . Kriemler, S. (2014). Quality of life is associated with physical activity and fitness in cystic fibrosis. *BMC Pulmonary Medicine*, *14*(1), 26. https://doi.org/10.1186/1471-2466-14-26

Hurter, L., Fairclough, S. J., Knowles, Z. R., Porcellato, L. A., Cooper-Ryan, A. M., & Boddy, L. M. (2018). Establishing raw acceleration thresholds to classify sedentary and stationary behaviour in children. *Children (Basel, Switzerland)*, *5*(12), 172. https://doi.org/10.3390/children5120172

Johnson, M. R., Ferkol, T. W., & Shepherd, R. W. (2006). Energy cost of activity and exercise in children and adolescents with cystic fibrosis. *Journal of Cystic Fibrosis: Official*

Journal of the European Cystic Fibrosis Society, *5*(1), 53–58. https://doi.org/10.1016/j.jcf.2005.10.001

Karas, M., Muschelli, J., Leroux, A., Urbanek, J. K., Wanigatunga, A. A., Bai, J., Crainiceanu, C. M., & Schrack, J. A. (2022). Comparison of accelerometry-based measures of physical activity: Retrospective observational data analysis study. *JMIR MHealth and UHealth*, *10*(7), e38077. https://doi.org/10.2196/38077

Kühnhausen, J., Dirk, J., & Schmiedek, F. (2017). Individual classification of elementary school children's physical activity: A time-efficient, group-based approach to reference measurements. *Behavior Research Methods*, *49*(2), 685–697. https://doi.org/10.3758/s13428-016-0724-2

Kuhn, L., Page, K., Ward, J., & Worral-Carter, L. (2013). The process and utility of classification and regression tree methodology in nursing research. *Journal of Advanced Nursin*, *70* (6), 1276–1286. https://doi.org/10.1111/jan.12288

Little, M. A., Varoquaux, G., Saeb, S., Lonini, L., Jayaraman, A., Mohr, D. C., & Kording, K. P. (2017). Using and understanding cross-validation strategies. Perspectives on saeb et al. *GigaScience*, *6*(5), 1–6. https://doi.org/10.1093/gigascience/gix020

Mackintosh, K. A., Montoye, A. H. K., Pfeiffer, K. A., & McNarry, M. A. (2016). *Investigating optimal accelerometer placement for energy expenditure prediction in children using a machine learning approach - IOPscience*. https://doi.org/10.1088/0967-3334/37/10/1728

Matthews, C. E., Hagströmer, M., Pober, D. M., & Bowles, H. R. (2012). Best practices for using physical activity monitors in population-based research. *Medicine and Science in Sports and Exercise*, *44*(1 Suppl 1), S68–S76. https://doi.org/10.1249/MSS.0b013e3182399e5b

McMurray, R. G., Butte, N. F., Crouter, S. E., Trost, S. G., Pfeiffer, K. A., Bassett, D. R., Puyau, M. R., Berrigan, D., Watson, K. B., Fulton, J. E., & Lof, M. (2015). Exploring metrics to express energy expenditure of physical activity in youth. *PLoS ONE*, *10*(6), e0130869. https://doi.org/10.1371/journal.pone.0130869

Migueles, J. H., Cadenas-Sanchez, C., Alcantara, J. M. A., Leal-Martín, J., Mañas, A., Ara, I., Glynn, N. W., & Shiroma, E. J. (2021). Calibration and cross-validation of accelerometer cut-points to classify sedentary time and physical activity from hip and non-dominant and dominant wrists in older adults. *Sensors*, *21*(10), 3326. MDPI AGRetrieved from. https://doi.org/10.3390/s21103326

Migueles, J. H., Rowlands, A. V., Huber, F., Sabia, S., & van Hees, V. T. (2019). GGIR: A research community–driven open source R package for generating physical activity and sleep outcomes from multi-day raw accelerometer data in. *Journal for the Measurement of Physical Behaviour*, *2*(3), 188–196. https://doi.org/10.1123/jmpb.2018-0063

Mirwald, R. L., Baxter-Jones, A. D., Bailey, D. A., & Beunen, G. P. (2002). An assessment of maturity from anthropometric measurements. *Medicine and Science in Sports and Exercise*, *34*(4), 689–694. https://doi.org/10.1097/00005768-200204000-00020

Montoye, A. H. K., Clevenger, K. A., Mackintosh, K. A., McNarry, M. A., & Pfeiffer, K. A. (2019). Cross-validation and comparison of energy expenditure prediction models using count-based and raw accelerometer data in youth. *Journal for the Measurement of Physical Behaviour*, *2*(4), 237–246. https://doi.org/10.1123/jmpb.2018-0011

Montoye, A. H. K., Westgate, B. S., Fonley, M. R., & Pfeiffer, K. A. (2018). Cross-validation and out-of-sample testing of physical activity intensity predictions with a wrist-worn accelerometer. *Journal of Applied Physiology (Bethesda, Md.: 1985)*, (5), 1284–1293. https://doi.org/10.1152/japplphysiol.00760.2017

Moore, V. C. (2012). Spirometry: Step by step. *Breathe*, *8*(3), 232–240. https://doi.org/10.1183/20734735.0021711

Patrick, E. A., & Fischer, F. P. (1970). A generalized k-nearest neighbor rule *information and control*. *Information & Control*, *16*(2), 128–152. https://doi.org/10.1016/S0019-9958(70)90081-1

Pfeiffer, K. A., Lisee, C., Westgate, B. S., Kalfsbeek, C., Kuenze, C., Bell, D., Cadmus-Bertram, L., & Montoye, A. H. (2023). Using accelerometers to detect activity type in a sport setting: Challenges with using multiple types of conventional machine learning approaches. *Measurement in Physical Education and Exercise Science*, *27*(1), 60–72. https://doi.org/10.1080/1091367X.2022.2069467

Quanjer, P. Q., Sanja, S., Tim, J. C., Xaver, B., Graham, L. H., & Bruce, H. C. (2012). *Multi-ethnic reference values for spirometry for the 3–95-yr age range: The global lung function 2012 equations*. https://doi.org/10.1183/09031936.00080312

Rahman, S., Irfan, M., Raza, M., Moyeezullah Ghori, K., Yaqoob, S., & Awais, M. (2020). Performance analysis of boosting classifiers in recognizing activities of daily living. *International Journal of Environmental Research and Public Health*, *17*(3), 1082. https://doi.org/10.3390/ijerph17031082

Rand, S., & Prasad, S. A. (2012). Exercise as part of a cystic fibrosis therapeutic routine. *Expert Review of Respiratory Medicine*, *6*(3), 341–352. https://doi.org/10.1586/ers.12.19

Ruch, N., Rumo, M., & Mader, U. (2011). Recognition of activities in children by two uniaxial accelerometers in free-living conditions. *European Journal of Applied Physiology*, *111*(8), 1917–1927. https://doi.org/10.1007/s00421-011-1828-0

Stephens, S., Takken, T., Esliger, D. W., Pullenayegum, E., Beyene, J., Tremblay, M., Schneiderman, J., Biggar, D., Longmuir, P., McCrindle, B., & Abad, A. (2016). Validation of accelerometer prediction equations in children with chronic disease. *Pediatric Exercise Science*, *28*(1), 117–132. https://doi.org/10.1123/pes.2014-0185

Tjurin, P., Niemelä, M., Kangas, M., Vähä-Ypyä, H., Sievänen, H., Jämsä, T., & Korpelainen, R. (2019). Accelerometry-based characteristics of overall sedentary behavior and sitting in middle-aged adults. *Measurement in Physical Education and Exercise Science*, *23*(3), 249–257. https://doi.org/10.1080/1091367X.2019.1613997

Tomezsko, J. L., Stallings, V. A., Kawchak, D. A., Goin, J. E., Diamond, G., & Scanlin, T. F. (1994). Energy expenditure and genotype of children with cystic fibrosis. *Pediatric Research*, *35*(4), 451–460. https://doi.org/10.1203/00006450-199404000-00013

Trost, S. G., Cliff, D. P., Ahmadi, M. N., Tuc, N. V., & Hagenbuchner, M. (2018). Sensor-enabled activity class recognition in preschoolers: Hip versus wrist data. *Medicine & Science in Sports & Exercise*, *50*(3), 634–641. https://doi.org/10.1249/mss.0000000000001460

Trost, S. G., Wong, W. K., Pfeiffer, K. A., & Zheng, Y. (2012). Artificial neural networks to predict activity type and energy expenditure in youth. *Medicine & Science in*

Sports & Exercise, *44*(9), 1801–1809. https://doi.org/10.1249/MSS.0b013e318258ac11

Trost, S. G., Zheng, Y., & Wong, W. (2014). Machine learning for activity recognition: Hip versus wrist data. *Institute of Physics and Engineering in Medici*, *35*(11), 2183–2189. https://doi.org/10.1088/0967-3334/35/11/2183

Trust, C. F. (2018). *UK cystic fibrosis registry 2018 annual data report*. https://www.cysticfibrosis.org.uk/sites/default/files/2020-12/2018%20Registry%20Annual%20Data%20Report.pdf

Vähä-Ypyä, H., Vasankari, T., Husu, P., Suni, J., Sievänen, & Sievänen, H. (2015). A universal, accurate intensity-based classification of different physical activities using raw data of accelerometer. *Clinical Physiology and Functional Imaging*, *35*(1), 64–70. https://doi.org/10.1111/cpf.12127

van de Weert van Leeuwen, P., Arets, H. G. M., van der Ent, C. K., & Beekman, J. M. (2013). Infection, inflammation and exercise in cystic fibrosis. *Respiratory Research*, *14*(1), 1–10.

Weir, J. B. (1949). New methods for calculating metabolic rate with special reference to protein metabolism. *Journal of Physiology, Paris*, *109*(1–2), https://doi.org/10.1113/jphysiol.1949.sp004363

Winham, S. J., Freimuth, R. R., & Biernacka, J. M. (2013). A weighted random forests approach to improve predictive performance. *Statistical Analysis and Data Mining*, *6*(6), 496–505. https://doi.org/10.1002/sam.11196

Zalewski, P., Marchegiani, L., Elsts, A., Piechocki, R., Craddock, I., & Fafoutis, X. (2020 6). From bits of data to bits of knowledge—an on-board classification framework for wearable sensing systems. *Sensors (Basel)*, *20*(6), 1655. https://doi.org/10.3390/s20061655

Zertuche, L. (2014). *Human activity recognition from accelerometer data*. Fellowship AI.

Zhang, Z. (2016). Introduction to machine learning: K-nearest neighbors. *Ann Transl Med*, *4*(11), 218–218. https://doi.org/10.21037/atm.2016.03.37