From Movement to METs: A Validation of ActTrust® for Energy Expenditure Estimation and Physical Activity Classification in Young Adults

Elias dos Santos Batista[1,2,3], Stephania Ruth Basilio Silva Gomes[4], Ayrton Bruno de Morais Ferreira[1], Lucas G. S. França[5,6], John Fontenele Araújo[3,4], Arnaldo Luis Mortatti[1], Mario André Leocadio-Miguel[7*]

1 Programa de Pós-Graduação em Educação Física, Universidade Federal do Rio Grande do Norte, Natal, Rio Grande do Norte, Brazil
2 Instituto Federal de Educação Ciência e Tecnologia do Rio Grande do Norte, Caicó, Rio Grande do Norte, Brazil
3 Programa de Pós Graduação Multicêntrico em Ciências Fisiológicas, Universidade Federal do Rio Grande do Norte, Natal, Rio Grande do Norte, Brazil
4 Programa de Pós-Graduação em Psicobiologia, Universidade Federal do Rio Grande do Norte, Natal, Rio Grande do Norte, Brazil
5 School of Computer Science, Faculty of Science and Environment, Northumbria University, Newcastle upon Tyne, United Kingdom
6 Department of Forensic and Neurodevelopmental Science, Institute of Psychiatry, Psychology & Neuroscience, London, United Kingdom
7 School of Psychology, Faculty of Health and Wellbeing, Northumbria University, Newcastle upon Tyne, United Kingdom

* mario.miguel@northumbria.ac.uk

Abstract

Estimating physical activity (PA) levels is a challenging and expensive task. An alternative could be the use of actigraphy devices to estimate PA. This has been previously done to a number of devices, including ActiGraph® GT3X+. In this study, we validated ActTrust® against the widely used GT3X+ and compared activity counts to metabolic equivalents (METs) derived from indirect calorimetry during treadmill walking and running. Fifty-six young adults (34 men, 22 women) participated in controlled effort exercises including light, moderate, vigorous, and very vigorous activity intensities. We developed a linear model to estimate energy expenditure (EE) from movement count of combinations of devices placed at hip or wrist. We then estimated cut-off points for each intensity range. Our results showed correlations between treadmill speed and both METs ($r = 0.95$, $p < 0.05$) and movement counts from both GT3X+ and ActTrust devices placed either on the hip ($r = 0.94$, $p < 0.05$; $r = 0.93, p < 0.05$) or on the wrist ($r = 0.88$, $p < 0.05$; $r = 0.88$, $p < 0.05$), respectively. Our proposed model performed well with balanced accuracies above 0.77 for all intensity ranges and over 0.9 for light and moderate activity. This is the first study to model estimate and validate PA intensity thresholds on ActTrust® devices. Our findings support the use of ActTrust® devices as simple, cost-effective tool for 24-hour assessments of EE.

Introduction

Physical activity (PA) is recognised for providing several health benefits in humans, mainly preventing and controlling chronic non-communicable diseases [1]. Furthermore, the practice of PA is associated with improved subjective and mental well-being and physical self-perception [2]. The World Health Organization (WHO) recommends 150-300 minutes of moderate-intensity aerobic PA, or at least 75-150 minutes of vigorous-intensity aerobic PA, or an equivalent combination of PA of both intensities during the week for adults. Also important is the WHO's global action plan on physical activity, which aims to decrease the prevalence of sedentarism in adults and adolescents by 15 per cent by 2030 [3].

One way to define the intensity of PA is through the metabolic equivalent (MET) of task. This mechanism expresses energy expenditure (EE) and indicates the oxygen needs for different activities, which at rest is equivalent to 3.5 ml O_2/ kg/min [4, 5]. The gold standard methodological approach to assessing energy expenditure is the doubly labelled water (DLW) technique [6]. However, there are limitations mainly related to this method's high cost, complexity and scalability. Alternatively, accelerometry emerges as a practical, scalable and economically viable method to estimate energy expenditure (EE) and quantify PA based on movement and body acceleration, generating numerical variables of frequency, duration and intensity of the actions performed [7, 8]. The challenges inherent in using accelerometry include the diversity of designs in accelerometer models, the disparities in signal processing between the different sensors and the multiplicity of options regarding possible placements [9–11]. Furthermore, it is possible to verify a variation in the algorithms that calculate the counts, both of the activity and other extracted parameters on different devices [12].

The ActiGraph® GT3X+ accelerometer has the most significant number of validation studies and published thresholds and has been widely used to validate devices with the same purpose [13]. The ActTrust® (Condor Instruments, São Paulo, Brazil) is a device that has a triaxial accelerometer and is specific for inferring sleep from locomotor activity [14]. It offers several advantages over other actigraphy devices: (1) significantly lower cost, making it accessible for large-scale studies in resource-limited settings; (2) longer battery life enabling extended monitoring periods; (3) integrated light and temperature sensors for comprehensive physiological assessment. However, despite these practical advantages, the ActTrust® has not been validated against gold-standard measures of EE, which limits its use in research and clinical contexts.

The validation of accelerometers through indirect calorimetry is crucial for developing mathematical equations capable of predicting energy expenditure using these devices and establishing cutoff points for classifying the intensity of PA derived from accelerometer data [15]. Based on this context, the two main objectives of this study are: (i) to define equations for estimating EE using ActTrust in young adults, and (ii) to establish cutoff points for the ActTrust movement variable (light, moderate, vigorous and very vigorous PA).

Materials and methods

Participants

Fifty-six healthy young adults were recruited for this study by online advertising. We established musculoskeletal or cardiovascular diseases that could hinder PA as excluding criteria. Participants were also excluded if they had any other contraindications to exercise or were taking medication that affected their metabolic rate. All participants completed the Physical Activity Readiness Questionnaire (PAR-Q) [16–18], and no one

was excluded from the study for answering yes to one or more questions. The statistics of the characteristics of the participants are presented in Table 1.

The study was approved by the Human Ethics Committee of the Federal University of Rio Grande do Norte, under protocol number 3.672.214 (CAAE: 15857419.5.0000.5537). It was performed according to the declaration of Helsinki and complied with the Ethical Standards in Sport and Exercise Science Research. Participants provided written consent to participate in the study. Recruitment took place between March 14, 2022 and October 24, 2022.

The sample size of 56 participants used in this study is consistent with previous validation research comparing accelerometer-based activity monitors against indirect calorimetry. Similar studies validating energy expenditure estimates from devices such as the ActiGraph GT3X+ have used samples ranging from 31 to 50 participants [8, 19]. Given the similarity in the study design, controlled laboratory setting and repeated measures across multiple activity intensities and device placements, our sample size provides sufficient statistical power to detect meaningful associations and develop robust predictive models.

Measuring activity

The *ActiGraph GT3X+* (ActiGraph Corporation, headquartered in Pensacola, FL, Firmware version 4.1.0) is a lightweight device that employs triaxial accelerometers and offers a dynamic range ranging from -6 g to +6 g. For the current study, a sampling rate of 30 Hz was chosen, as this range is expected to capture the majority of accelerations arising from human movement sufficiently [20]. After digitisation, the signal passed through a digital filter that restricts the accelerometer's frequency range to 0.25–2.5 Hz. We used the vector magnitude (VM) data (vector representing the summed value of the three axes). Data from the GT3X+ device were retrieved and processed using *ActiLife version 6.13.4*, software provided by ActiGraph Corporation based in Pensacola, FL.

The ActTrust® device (Condor Instruments, São Paulo, SP, Brazil) is an actigraphy tool that features a rechargeable battery, a digital tri-axial accelerometer, a skin temperature sensor and a light sensor. It samples acceleration data at 25 Hz, ranging from 0.03 to 4 g. First, this data undergoes processing through a band-pass filter (0.5 - 2.7 Hz), after which the norm is calculated from the three axes. The resultant value is then integrated over an epoch (1 to 60 seconds) and stored in memory. This equipment is designed for activity and sleep monitoring and processes activity using parameters such as PIM (proportional integration mode), ZCM (zero crossing mode) and TAT (time above threshold) [21]. The *Act Studio software* (Condor Instruments, São Paulo, Brazil) downloads, visualises, and exports the collected data.

Importantly, this study used the highest temporal resolution available for both actigraphy devices, according to their specific firmware. Each 1-second epoch was calculated by aggregating filtered and rectified raw acceleration data into a single 1-second bin. This was consistent with validation approaches for wrist and hip-worn accelerometers across the same range of walking speeds [8, 19].

Measuring oxygen uptake

Oxygen uptake was measured 'breath-by-breath' continuously during each condition using indirect calorimetry (*Quark CPET*, Cosmed®, Italy). Before testing each subject, we calibrated the metabolic cart with a known gas mixture (16% O_2 and 5% CO_2) and volume. One test had to be repeated seven days later due to an error in the treadmill's security system (*Super ALT*, Inbramed®, Porto Alegre, Brazil). Occasional errant breath episodes (e.g., due to coughing, swallowing or talking) were excluded from the

data set when exceeding three standard deviations of the mean, the latter being defined as the average of two following and two preceding sampling intervals [22].

Experimental procedure

The entire data collection took place between March 21, 2022, and November 1, 2022. All *GT3X+* and *ActTrust* devices were initialiased via computer interfaces to collect data in 1-second epochs [19]. The GT3X+ was initialised to collect the VM data, and the ActTrust to collect activity in the PIM mode. Each participant wore two pairs of GT3X+ and ActTrust devices simultaneously, securely placed on the hip on the side of the participant's dominant lower limb through an elastic belt and the other on the non-dominant wrist.

Participants refrained from exercise and fasted for at least 4 hours before the testing session. The protocol consisted of 5 conditions (of 10-min duration each), using a progressive incremental design, interspersed with 5-minute rest periods: (i) resting; (ii) treadmill (Quasar Med 4.0, h p cosmos, Nussdorf-Traunstein, Germany) walking at 3 $km \cdot h^{-1}$; (iii) treadmill walking at 5 $km \cdot h^{-1}$; (iv) treadmill walking or running at 7 $km \cdot h^{-1}$; and (v) treadmill running at 9 $km \cdot h^{-1}$. Prior to initiating the protocol, devices were synchronised using a manual time-matching approach [23], allowing precise temporal alignment between the accelerometers and breath-by-breath metabolic data.

Data analysis

To ensure that data reflected true physiological steady states, we adopted the protocol described in previous studies [8, 19]. Thus, we calculated the mean activity counts from the four central minutes of the Vector Magnitude (VM) for GT3X+ and the counts (proportional integral mode - PIM) for *ActTrust* to estimate activity levels. METs (Metabolic Equivalents) were individually calculated (VO^2) similarly.

First, we assessed the effect of the different devices and placement with an ANOVA according to the model Treadmill speed$\sim$ $\text{Activity}^{1/2}$: $\text{Device}_{\text{placement}}$. Secondly, we then modelled the association between METs and activity for each device using a general linear model defined in equation 1. A square root transformation was applied to both METs and activity counts due to the non-linear relationship. Model assumptions were verified through examination of residual plots and Q-Q plots for normality – included in supplementary material S3 Figure.

$$\text{MET}^{1/2} \sim \text{Activity}^{1/2} * \text{Device}_{\text{placement}} \tag{1}$$

We evaluated two devices according to their placement, hence our model can be written as in equation 2.

$$\text{MET}^{1/2} = \beta_0^{\dagger} + \beta_1^{\dagger}\text{Activity}^{1/2} \tag{2}$$

with $\beta_0^{\dagger} = \beta_0 + \beta_2\text{ACTT}_{hip} + \beta_3\text{ACTT}_{wrist} + \beta_4\text{GT3X+}_{wrist}$ and $\beta_1^{\dagger} = \beta_1 + \beta_5\text{ACTT}_{hip} + \beta_6\text{ACTT}_{wrist} + \beta_7\text{GT3X+}_{wrist}$. ACTT_{hip}, ACTT_{wrist}, and GT3X+_{wrist} are binary variables set to adjust the model values to each configuration. GT3X+_{hip} was defined as the reference level.

Analyses were conducted using R v. 4.4.1 (R Foundation for Statistical Computing, Vienna, Austria) [24] and auxiliary packages *tidyverse* v. 2.0.0 [25], *caret* v. 6.0.94 [26], *MuMIn* v. 1.48.4 [27], *MetBrewer* v. 0.2.0 [28], *cowplot* v. 1.1.3 [29], *pROC* v. 1.18.5 [30], *broom* v. 1.0.6 [31], *lmtest* v. 0.9.40 [32], *car* v. 3.1.2 [33], *effectSize* v. 1.0.2 [34], and *msm* v. 1.8.2 [35]. The significance level was set at $\alpha = 0.05$. The data and scripts used in the

Fig 1. Scores for each treadmill speed band. a) Metabolic equivalents (VO^2) obtained from all participants for each speed band. b) Activity counts obtained for each speed band for every combination of placement and device, i.e., *GT3X+ (hip)*, *ACTT (hip)*, *ACTT (wrist)*, *GT3X+ (wrist)*. The scores across the different devices show a similar profiles increasing towards higher speeds, which is compatible to the variation exhibited by the ergo-spirometry metabolic equivalents measures. Point shape denotes sex (square: male, $n = 34$; circle: female, $n = 22$).

analyses featured in this article are available at `https://github.com/circadia-bio/ACTT_validation_study` [36].

We estimated metabolic rate in METs to derive the cut-points for moderate, vigorous and very vigorous exercise intensity levels according to the intervals (3.0-5.99, 6.0-8.99 and above 9.0 METs).

We employed performance metrics for model diagnostics and computed R^2 and confusion matrix. Finally, we evaluated the ability of the new cut points from the linear regression model to accurately classify the physical activity intensity levels derived from EE objective measurements, according to the different MET bands with sensitivity, specificity, balanced accuracy, and area-under-the-curve (AUC) bench marks in a similar fashion to previous works in literature [8].

Results

The sample comprised young adults (34 men, 22 women) aged 18–35. The participants' characteristics are displayed in Table 1. Energy expenditure values obtained with indirect calorimetry reflected each condition, increasing proportionally according to the treadmill speed, see Figure 1.

Table 1. Demographic and physical characteristics of the participants. Mean (st. dev.).

Sex	Sample size	Age (y)	Height (cm)	Weight (kg)	Body mass index (kg/m^2)
Male	34	28.3 (4.6)	172 (6.2)	78.2 (11.9)	23.1 (1.6)
Female	22	26.8 (5.0)	160 (7.3)	59.9 (9.4)	26.7 (2.5)

Activity counts and treadmill velocity

We performed a two-way ANOVA to understand the outputs of each device (GT3X+ and ActTrust) according to the placement and condition (treadmill velocity). Movement counts from the GT3X+ devices were mostly distinct ($F(19, 1087) = 1347$, $p < 0.001$, $\eta^2 = 0.96$, 95% CI $[0.96, 1.00]$) with significantly lower activity magnitudes than those from the ActTrust devices at 3, 5, 7 and 9 $km \cdot h^{-1}$ ($p < 0.05$).

In contrast, placement did not affect the measures for GT3X+ regardless of hip or wrist acquisitions for 3 $km \cdot h^{-1}$ and 5 $km \cdot h^{-1}$. Movement counts of both devices (GT3X+ and ActTrust) placed on the hip or wrist increased with the increment of the treadmill speed (see Figure 1). Statistical details of the ANOVA analysis and post-hoc Tukey's Honestly Significant Difference (HSD) [37, 38] tests were included in the supplementary table S1.

Definition of equations to estimate energy expenditure

Our main objective was to analyse the relationship between motion measurements obtained with the ActTrust® in comparison to the GT3X+® and metabolic equivalent

based on gas exchange in young adults. First, we explored how treadmill speed is associated with METs and movement counts. Specifically, speed was highly correlated to both METs ($r = 0.95, p < 0.05$) and movement counts from both GT3X+ and ActTrust devices placed either on the hip ($r = 0.94, p < 0.05; r = 0.93, p < 0.05$) or on the wrist ($r = 0.88, p < 0.05; r = 0.88, p < 0.05$), respectively. Our second step was to develop a linear regression model to estimate METs based on the movement counts recorded by each device at each placement (see Figure 2c). The resulting equations are crucial for estimating physical activity in METs from the movement counts of the GT3X+ and ActTrust devices placed on the hip and wrist (Table **??**).

GT3X+® and ActTrust® movement count cut points to classify physical activity intensity

Finally, we derived cut-points based on MET ranges [0,3) or light, [3,6) or moderate, [6,9) or vigorous, [9,∞) or very vigorous physical activity intensity ranges based on the movement counts from both devices and placement. We then assessed the ability of the new cut points from our linear regression model to accurately classify the physical activity intensity levels objectively derived from EE measurement (METs). We estimated the sensitivity (recall), specificity, balanced accuracy, area-under-the-curve (AUC), and confusion matrix for each intensity zone (for metrics see Table 3). The obtained confusion matrices show a consistent high accuracy for the [0,3); [3,6); and [9, ∞) classes, featuring true positive rates above 75%, both when aggregating all the devices and placements (see Figure 2b) and for specific combinations of device and placement (see Figure 3). Sensitivity and balanced accuracy are reduced for vigorous and very vigorous activity but remained well above 0.6, see Table 3.

Fig 2. Cross device score analyses. a) Pearson correlation between the activity counts magnitudes measured at each device and placement combination. b) Model of Activity count and METs following the relationship $\textit{Treadmill speed}^{1/2} \sim \textit{Activity}^{1/2} \times \textit{Device}_{\textit{placement}}$. c) Confusion matrix relating predicted intensity ranges from model in b) and metabolic equivalents (VO^2) obtained from simultaneous ergo-spirometry measurements, with a high true positive rate for the classes [0,3), [3,6) and [9,∞)

Table 2. Cut-points and regression equations with 95% confidence intervals per device and placement.

	Cut-point, counts/min (95% CI)			$\text{MET}^{1/2} = \beta_0 + \beta_1 \times \text{Activity}^{1/2}$	
Device	3 METs	6 METs	9 METs	β_0 (95% CI)	β_1 (95% CI)
GT3X+ (hip)	1132 (1041, 1223)	4853 (4688, 5018)	9468 (9135, 9800)	1.062 (1.019, 1.105)	0.0199 (0.0193, 0.0206)
ACTT (hip)	5057 (4624, 5491)	23339 (22522, 24157)	46410 (44764, 48055)	1.107 (1.066, 1.148)	0.0088 (0.0085, 0.0091)
ACTT (wrist)	3761 (3357, 4165)	22368 (21503, 23232)	47203 (45409, 48997)	1.233 (1.195, 1.271)	0.0081 (0.0079, 0.0084)
GT3X+ (wrist)	1698 (1525, 1871)	9503 (9143, 9862)	19787 (19047, 20527)	1.207 (1.168, 1.245)	0.0127 (0.0123, 0.0132)

Table 3. Sensitivity, specificity, balanced accuracy, and area under the curve for the [0,3), [3,6), [6,9), and [9,∞) intensity ranges

Metric	**Class: [0,3)**	**Class: [3,6)**	**Class: [6,9)**	**Class: [9,∞)**
Sensitivity	0.8810	0.9190	0.6630	0.6441
Specificity	0.9762	0.9182	0.8932	0.9556
Balanced Accuracy	0.9286	0.9186	0.7781	0.7998
AUC	0.9820	0.6307	0.7813	0.9550

Fig 3. Confusion matrices for each combination of device and placement, i.e., a) *GT3X+ (hip)*, b) *GT3X+ (wrist)*, c) *ACTT (hip)*, *ACTT (wrist)*. The individual devices confusion matrices reproduce a similar profile to the one observed for the entire model, with a higher true positive rate for the classes [0,3), [3,6) and [9,∞)

Discussion

This study compared the *ActTrust*® and *ActiGraph GT3X+*® activity counts during resting, treadmill walking and running. The main findings are as follows. Firstly, we established the suitability of our indirect calorimetry setup by measuring the metabolic equivalent as a function of the conditions (treadmill speed, resting). The speed of the treadmill was highly and significantly correlated with the estimated METs. Secondly, we demonstrated that movement counts increased as speed increased, and both *ActTrust*® and *GT3X+*® could differentiate among the different conditions, independently of the placement, either on the wrist or the hip. As a next step, we then modelled the association between METs and activity for each device X placement using a general linear model. In addition, we established cut points for movement counts of GT3X+ and ActTrust devices to classify the intensity of physical activity. These points were derived for the categories of moderate, intense and very intense physical activity (see Table **??**), considering both the movement counts of the devices and the placement. Finally, we evaluated the ability of the new cut points from the linear regression models to accurately classify the intensity levels of physical activity objectively derived from EE measurement (METs). For this analysis, we calculated the balanced accuracy, sensitivity (recall), and specificity for each cut point and each linear model (device X placement).

Our results demonstrated that the movement count of both devices (GT3X+ and ActTrust), placed on both the hip and the wrist, increased with the increase in the velocity of the treadmill, as did the energy expenditure (METs) obtained from indirect calorimetry. Our results are consistent with similar validation studies [39, 40]. If we compare our results with studies which used different devices [8, 19], while our model's sensitivity and accuracy are reduced for vigorous and very vigorous activity, it still achieves accuracies above 0.90 and higher AUC scores than previous studies. Precisely, our model achieved area under the ROC curve of 0.98, 0.63, 0.78 and 0.95 for light, moderate, vigorous and very vigorous physical activity, respectively. As a comparison, the artificial neural network model developed by Santos-Lozano et al., resulted in AUC of 0.8, 0.6, 0.7 and 0.6 using a very similar protocol [8]. Therefore, we found that the model we proposed is accurate with a simpler method for the estimation of energy expenditure (EE) than the equations previously developed. Furthermore, our approach is fully explainable unlike neural network models.

The cut-points derived in the present study for the GT3X+ placed on the hip were lower than most published GT3X reference values across all thresholds. Compared to Sasaki et al. [19], our cut-points were 58%, 21%, and 2% lower at the MET 3, 6, and 9 thresholds, respectively, and 65%, 43%, and 18% lower than Santos-Lozano et al. [8] respective cut-off points for adults. Importantly, however, there is substantial variability in cut-points between these two quite similar studies, as they differ by 16% at MET 3, 39% at MET 6 and 20% at MET 9. The overall variability suggests that sample-level characteristics are the main drivers of the observed discrepancies rather than device or methodological artefacts.

Nevertheless, this is the first study to develop movement-based cut-off points for wrist and hip-worn ActTrust® devices. These cut-offs reflect moderate, intense and very intense physical activity categories. The study successfully classified physical activity intensity across the spectrum, demonstrating excellent accuracy. However, it is not possible to compare these values with other studies. For the discrimination between intensity categories based on METs, we propose thresholds to be used for the hip and wrist-worn ActTrust® devices. This allows researchers to examine how much walking people perform and distinguish it from running using these accelerometers.

Beyond estimating the intensity of physical activity, the value of validating ActTrust® lies in its ability to capture the entire 24-hour rest-activity cycle within a single device. Therefore, establishing PA cut-points for ActTrust® is a necessary first

step toward leveraging its broader monitoring capabilities, which extend beyond exercise to the assessment of sleep and circadian rhythms. Actigraphy devices, such as the *ActTrust*® accelerometers have been used in sleep and circadian research for decades [41–43]. Actually, the sleep and wake states have been classically inferred by algorithms based on activity counts, provided frequency, duration and intensity of movement [44]. Although sleep studies are designed to collect 24-hour data, including rest and physical activity throughout the day, the simultaneous analysis of sleep/circadian data, physical activity, and sedentary behaviour is limited [13]. This gap is notable given strong evidence linking regular exercise, reduced sedentary behaviour, and healthy sleep to improved longevity and disease prevention [45, 46].

Despite this clear dependence between physical activity and sleep, it is essential to move one step further and take advantage of the complete 24-hour rest-activity cycle data obtained through actigraphy. Precise health guidelines can only be determined by combining measurements of sleep, sedentary behaviour, and physical activity [13]. That said, expanding the availability of cut-off points for physical activity based on multiple actigraphy devices increases the potential to improve recommendations, guidelines and, ultimately, health and longevity. Therefore, studies like ours and the introduction of more afordable actigraphy devices, such as the ActTrust®, have the potential to stimulate future research by providing tools to extract physical activity and sleep information, even retrospectively. In fact, one of the main advantages of using accelerometry over self-reported (retrospective) physical activity is that it is not prone to recognition or memory biases, as recall measures are [47]. The validation of devices that were exclusively dedicated to inferring sleep-wake cycles as tools to identify and classify physical activity according to the metabolic equivalent opens up the possibility of using such data not only prospectively, but also retrospectively. Studies such as ours are crucial to developing personalised programmes that maximise long-term adherence to physical activity and the long-term benefits of exercise. We acknowledge that adults should aim for at least 150 to 300 minutes of moderate-intensity aerobic physical activity, or 75 to 150 minutes of vigorous-intensity aerobic physical activity, or an equivalent combination throughout the week, to achieve substantial health benefits [48]. Likewise, adults should usually aim for 7 or more hours of regular sleep each night to promote optimal health [49]. However, only validated accelerometer devices can provide information on these two essential proxies of health simultaneously (sleep and physical activity).

It is important to acknowledge the limitations in the findings of this study. Although a significantly large number of participants (N=56) was considered, the participants were young healthy adults, with no functional limitations, which limits generalisability. As a result, the findings may not be applicable to individuals with health conditions and other age groups without further validation. In addition, the study was carried out in a laboratory setting and focused on specific physical activities, walking and running, which may also limit the universality of our findings. Although these activities are common among populations around the world, it is crucial to assess physical activities in real life and home environments. Moreover, the fixed progressive protocol may have introduced order effects such as those linked to warm-up or fatigue. While this approach was chosen for safety, and mirrors clinical exercise testing protocols, future studies could employ counterbalanced designs or include longer rest periods to eliminate carryover effects.

Further work is needed to define appropriate cut-off points for other groups, such as children and older people, as well as different device placements, as our findings are limited to devices worn on the wrist and hip. Furthermore, the exercise protocol followed a fixed incremental order, which precludes ruling out potential order effects such as progressive fatigue or warm-up adaptation; future studies should consider

counterbalanced or randomised condition sequences to address this limitation.

Conclusion

In this study we have for the first time validated measures of the ActTrust®. When paired with a simple linear model, we can estimate energy expenditure and classify the intensity of physical activity. Our model distinguished between light, moderate, vigorous, and very vigorous activity intensity ranges with benchmark metrics comparable to or exceeding those reported in the literature, whilst adopting a much simpler and explainable model than black-box approaches such as neural networks, facilitating interpretation and implementation by practitioners. Nevertheless, we did not directly compare performance against such alternative modelling approaches in this study.

These results demonstrated the feasibility of using ActTrust® as a cost effective and scalable tool to estimate PA. It also has potential for potential for clinical use following validation in patient populations with diverse health conditions and functional capacities. Our results also provide hints on how to harmonise outcomes from different devices whilst estimating PA and analysing activity measures. Future work should extend our analyses to free-living conditions, comparison with commercial wearable devices other than GT3X+® and more naturalistic activity patterns.

Declarations

Author contribution. Conceptualization: ESB, ML-M. Formal analysis: LGSF, ML-M. Investigation: ESB, ABMF. Methodology: ESB, ABMF, LGSF, ML-M. Project administration: ML-M. Supervision: ML-M, JFA, ALM. Writing – original draft: ESB, ABMF, SRBG, LGSG, ML-M. Writing – review & editing: JFA, ALM, LGSF, ML-M.

Funding. The authors received no specific funding for this work.

Ethics approval and consent to participate. This study was approved by the Human Ethics Committee of the Federal University of Rio Grande do Norte, under protocol number 3.672.214 (CAAE: 15857419.5.0000.5537). Participants provided written consent to participate in the study.

Conflict of interest/Competing interests. The authors declare no competing interests.

Data availability. The data used in the analyses featured in this article are available at `https://github.com/circadia-bio/ACTT` validation study.

Materials availability. Not applicable.

Code availability. The code used in the analyses featured in this article are available at `https://github.com/circadia-bio/ACTT` validation study.

Supporting information

S1 Table. ANOVA analysis and post-hoc Tukey's Honestly Significant Difference (HSD) table.

S2 Table. Area under the curve (AUC) for one-vs-rest physical activity intensity classification stratified by sex. AUC values were computed for each intensity class using the model prediction score against the true class derived from indirect calorimetry for male (n=34) and female (n=22) participants. Intensity classes are defined as light [0,3), moderate [3,6), vigorous [6,9), and very vigorous [9,∞) METs.

S1 Figure. Confusion matrices for physical activity intensity classification stratified by sex. Row-normalised confusion matrices showing the proportion of observations assigned to each predicted intensity class relative to the true class derived from indirect calorimetry for (A) male (n=34) and (B) female (n=22) participants. Intensity classes are defined as light [0,3), moderate [3,6), vigorous [6,9), and very vigorous [9,∞) METs.

S2 Figure. Bland-Altman plots comparing model-predicted and measured metabolic equivalents (METs) for each device and placement combination. Each panel plots the difference between predicted and measured METs against their mean for each device and placement with GT3X+ hip (A), GT3X+ wrist (B), ActTrust hip (C), and ActTrust wrist (D). Red dashed lines indicate the mean bias and blue dotted lines indicate the 95% limits of agreement. There is no evidence of systematic proportional bias across the range of measured values.

S3 Figure. Regression diagnostics for the general model. A) Normal Q-Q plot of model residuals with points following the reference blue line, featuring minor deviations at the tails. B) Residuals versus fitted values with no systematic trend supporting the assumption of homoscedasticity.

References

1. Anderson E, Durstine JL. Physical activity, exercise, and chronic diseases: A brief review. Sports Medicine and Health Science. 2019;1(1):3–10. doi:https://doi.org/10.1016/j.smhs.2019.08.006.

2. Buecker S, Simacek T, Ingwersen B, Terwiel S, and BAS. Physical activity and subjective well-being in healthy individuals: a meta-analytic review. Health Psychology Review. 2021;15(4):574–592. doi:10.1080/17437199.2020.1760728.

3. Bull FC, Al-Ansari SS, Biddle S, Borodulin K, Buman MP, Cardon G, et al. World Health Organization 2020 guidelines on physical activity and sedentary behaviour. British Journal of Sports Medicine. 2020;54(24):1451–1462. doi:10.1136/bjsports-2020-102955.

4. deJong A. The metabolic equivalent: reevaluating what we know about the MET. ACSM's Health & Fitness Journal. 2010;14(4):43–46.

5. Franklin BA, Brinks J, Berra K, Lavie CJ, Gordon NF, Sperling LS. Using metabolic equivalents in clinical practice. The American journal of cardiology. 2018;121(3):382–387.

6. Speakman J. Doubly Labelled Water: Theory and Practice. Springer Science & Business Media;.

7. Migueles JH, Cadenas-Sanchez C, Ekelund U, Delisle Nyström C, Mora-Gonzalez J, Löf M, et al. Accelerometer data collection and processing criteria to assess physical activity and other outcomes: a systematic review and practical considerations. Sports medicine. 2017;47:1821–1845.

8. Santos-Lozano A, Santin-Medeiros F, Cardon G, Torres-Luque G, Bailon R, Bergmeir C, et al. Actigraph GT3X: validation and determination of physical activity intensity cut points. International journal of sports medicine. 2013;34(11):975–982.

9. van Hees VT, Thaler-Kall K, Wolf KH, Brønd JC, Bonomi A, Schulze M, et al. Challenges and opportunities for harmonizing research methodology: Raw accelerometry. Methods of information in medicine. 2016;55(06):525–532.

10. Rosenberger ME, Haskell WL, Albinali F, Mota S, Nawyn J, Intille S. Estimating activity and sedentary behavior from an accelerometer on the hip or wrist. Medicine and science in sports and exercise. 2013;45(5):964.

11. Redenius N, Kim Y, Byun W. Concurrent validity of the Fitbit for assessing sedentary behavior and moderate-to-vigorous physical activity. BMC medical research methodology. 2019;19:1–9.

12. Neishabouri A, Nguyen J, Samuelsson J, Guthrie T, Biggs M, Wyatt J, et al. Quantification of acceleration as activity counts in ActiGraph wearable. Scientific Reports. 2022;12(1):11958.

13. Rosenberger ME, Buman MP, Haskell WL, McConnell MV, Carstensen LL. 24 hours of sleep, sedentary behavior, and physical activity with nine wearable devices. Medicine and science in sports and exercise. 2016;48(3):457.

14. Bellone GJ, Plano SA, Cardinali DP, Chada DP, Vigo DE, Golombek DA. Comparative analysis of actigraphy performance in healthy young subjects. Sleep Science. 2016;9(4):272–279.

15. Chen KY, David R Bassett J. The technology of accelerometry-based activity monitors: current and future. Medicine & Science in Sports & Exercise. 2005;37(11):S490–S500.

16. Chisholm D, Collis M, Kulak L, Davenport W, Gruber N. Physical activity readiness. BC Med J. 1975;17(2):375–8.

17. Warburton DE, Gledhill N, Jamnik VK, Bredin SS, McKenzie DC, Stone J, et al. Evidence-based risk assessment and recommendations for physical activity clearance: Consensus Document 2011. Applied Physiology, Nutrition, and Metabolism. 2011;36(S1):S266–S298.

18. Schwartz J, Oh P, Takito MY, Saunders B, Dolan E, Franchini E, et al. Translation, cultural adaptation, and reproducibility of the physical activity readiness questionnaire for everyone (PAR-Q+): the Brazilian Portuguese version. Frontiers in cardiovascular medicine. 2021;8:712696.

19. Sasaki JE, John D, Freedson PS. Validation and comparison of ActiGraph activity monitors. Journal of science and medicine in sport. 2011;14(5):411–416.

20. Matlary RED, Holme PA, Glosli H, Rueegg CS, Grydeland M. Comparison of free-living physical activity measurements between ActiGraph GT3X-BT and Fitbit Charge 3 in young people with haemophilia. Haemophilia. 2022;28(6):e172–e180.

21. Ancoli-Israel S, Cole R, Alessi C, Chambers M, Moorcroft W, Pollak CP. The role of actigraphy in the study of sleep and circadian rhythms. Sleep. 2003;26(3):342–392.

22. Lamarra N, Whipp BJ, Ward SA, Wasserman K. Effect of interbreath fluctuations on characterizing exercise gas exchange kinetics. Journal of Applied Physiology. 1987;62(5):2003–2012.

23. Crouter SE, Kuffel E, Haas JD, Frongillo EA, Bassett DR. Refined Two-Regression Model for the ActiGraph Accelerometer;42(5):1029–1037. doi:10.1249/MSS.0b013e3181c37458.

24. R Core Team. R: A Language and Environment for Statistical Computing; 2024. Available from: `https://www.R-project.org/`.

25. Wickham H, Averick M, Bryan J, Chang W, McGowan LD, François R, et al. Welcome to the tidyverse. Journal of Open Source Software. 2019;4(43):1686. doi:10.21105/joss.01686.

26. Kuhn, Max. Building Predictive Models in R Using the caret Package. Journal of Statistical Software. 2008;28(5):1–26. doi:10.18637/jss.v028.i05.

27. Bartoń K. MuMIn: Multi-Model Inference; 2024. Available from: `https://CRAN.R-project.org/package=MuMIn`.

28. Mills BR. MetBrewer: Color Palettes Inspired by Works at the Metropolitan Museum of Art; 2022. Available from: `https://CRAN.R-project.org/package=MetBrewer`.

29. Wilke CO. cowplot: Streamlined Plot Theme and Plot Annotations for 'ggplot2'; 2024. Available from: `https://CRAN.R-project.org/package=cowplot`.

30. Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez JC, et al. pROC: an open-source package for R and S+ to analyze and compare ROC curves. BMC Bioinformatics. 2011;12:77.

31. Robinson D, Hayes A, Couch S. broom: Convert Statistical Objects into Tidy Tibbles; 2024. Available from: `https://CRAN.R-project.org/package=broom`.

32. Zeileis A, Hothorn T. Diagnostic Checking in Regression Relationships; 2002. Available from: `https://CRAN.R-project.org/doc/Rnews/`.

33. Fox J, Weisberg S. An R Companion to Applied Regression. 3rd ed. Thousand Oaks CA: Sage; 2019. Available from: `https://socialsciences.mcmaster.ca/jfox/Books/Companion/`.

34. Ben-Shachar MS, Lüdecke D, Makowski D. effectsize: Estimation of Effect Size Indices and Standardized Parameters. Journal of Open Source Software. 2020;5(56):2815. doi:10.21105/joss.02815.

35. Christopher H Jackson. Multi-State Models for Panel Data: The msm Package for R. Journal of Statistical Software. 2011;38(8):1–29. doi:10.18637/jss.v038.i08.

36. dos Santos Batista E, Basilio Gomes SR, Bruno de Morais Ferreira A, França LGS, Fontenele Araújo J, Mortatti AL, et al.. From Movement to METs: A Validation of ActTrust® for Energy Expenditure Estimation and Physical Activity Classification in Young Adults; 2025.

37. Miller RG. Simultaneous Statistical Inference. Springer Series in Statistics. New York, NY: Springer New York; 1981.

38. Yandell BS. Practical Data Analysis for Designed Experiments. 1st ed. Routledge; 2017.

39. Davoudi A, Wanigatunga AA, Kheirkhahan M, Corbett DB, Mendoza T, Battula M, et al. Accuracy of samsung gear s smartwatch for activity recognition: Validation study. JMIR mHealth and uHealth. 2019;7(2):e11270.

40. Skotte J, Korshøj M, Kristiansen J, Hanisch C, Holtermann A. Detection of physical activity types using triaxial accelerometers. Journal of physical activity and health. 2014;11(1):76–84.

41. Yuan H, Hill EA, Kyle SD, Doherty A. A systematic review of the performance of actigraphy in measuring sleep stages. Journal of Sleep Research. 2024;33(5):e14143.

42. Sadeh A. The role and validity of actigraphy in sleep medicine: an update. Sleep medicine reviews. 2011;15(4):259–267.

43. Evangelista LA, Eckeli AL, Stelzer FG, Carvalho VD, Soares HF. Validation of a National Actigraph in Children and Adolescents;138:107776. doi:10.1016/j.sleep.2025.107776.

44. Fekedulegn D, Andrew ME, Shi M, Violanti JM, Knox S, Innes KE. Actigraphy-based assessment of sleep parameters. Annals of Work Exposures and Health. 2020;64(4):350–367.

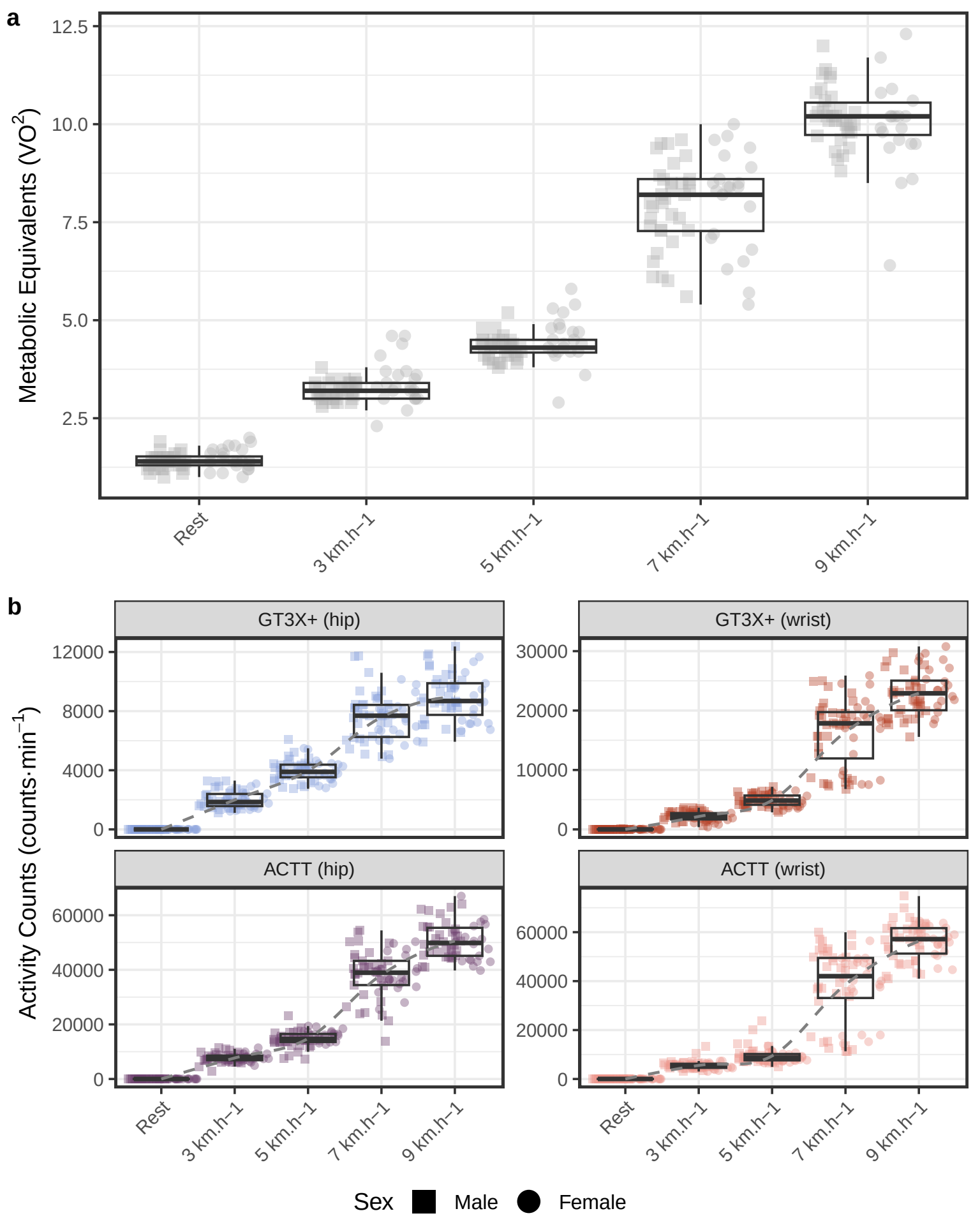
45. Morris CJ, Yang JN, Scheer FA. The impact of the circadian timing system on cardiovascular and metabolic function. Progress in brain research. 2012;199:337–358.

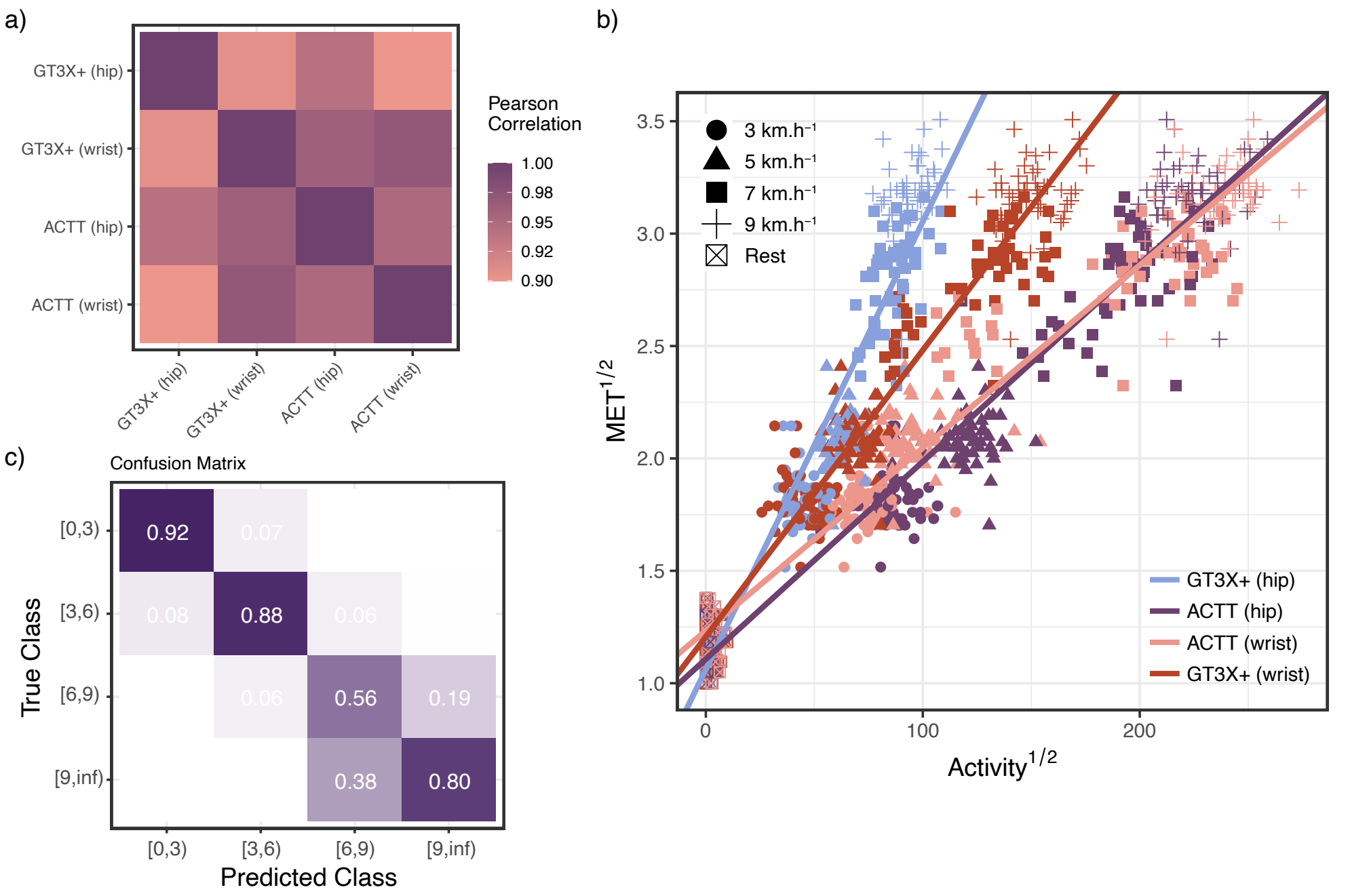
46. Duncan MJ, Murphy L, Oftedal S, Fenwick MJ, Vincent GE, Fenton S. The associations between physical activity, sedentary behaviour, and sleep with mortality and incident cardiovascular disease, cancer, diabetes and mental health in adults: a systematic review and meta-analysis of prospective cohort studies. Journal of Activity, Sedentary and Sleep Behaviors. 2023;2(1):19.

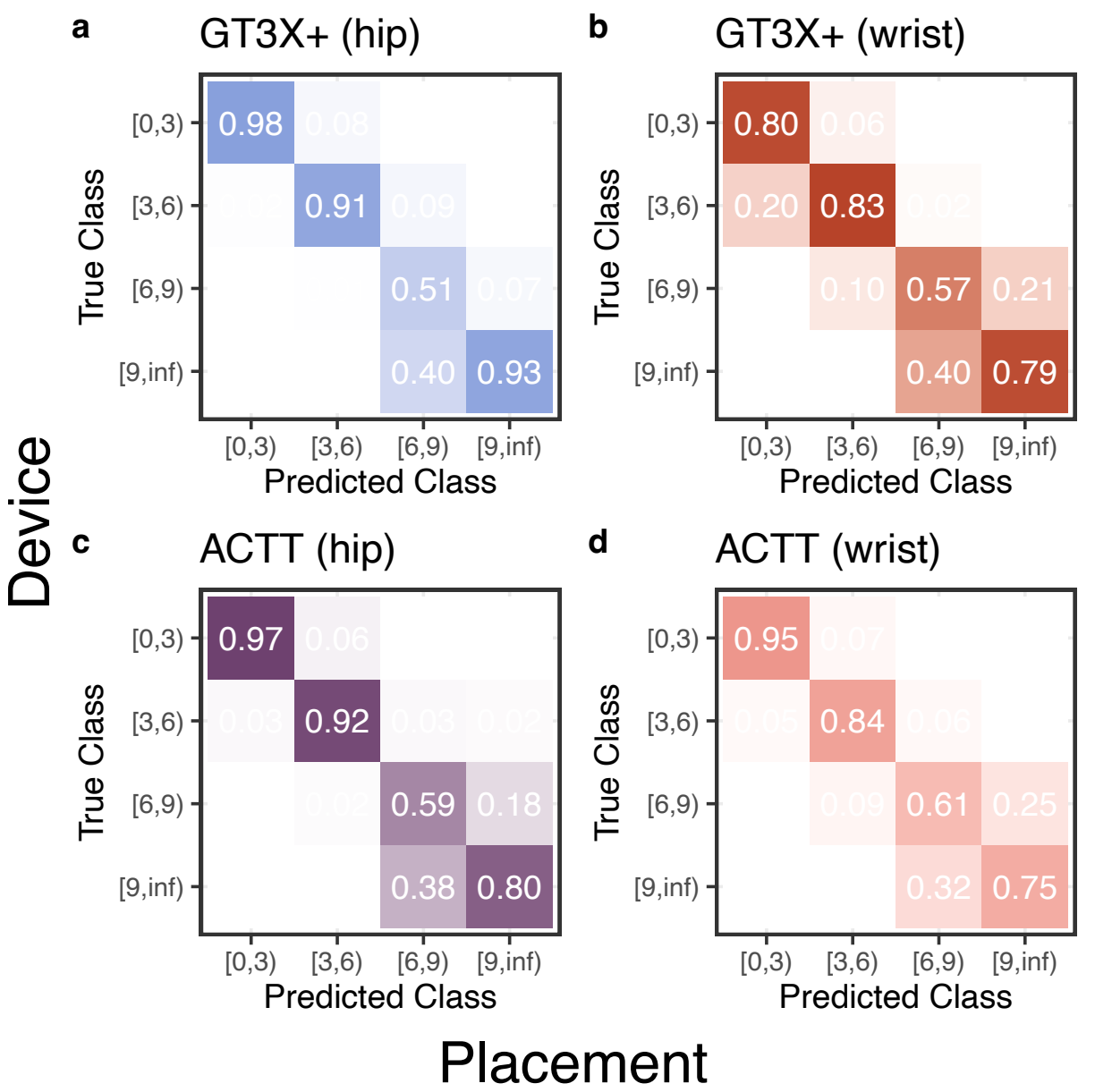
47. Burchartz A, Anedda B, Auerswald T, Giurgiu M, Hill H, Ketelhut SI, et al. Assessing physical behavior through accelerometry–state of the science, best practices and future directions. Psychology of Sport and Exercise. 2020;49:101703.

48. Burtscher J, Millet GP, Burtscher M. Pushing the limits of strength training. American journal of preventive medicine. 2023;64(1):145–146.

49. Panel CC, Watson NF, Badr MS, Belenky G, Bliwise DL, Buxton OM, et al. Recommended amount of sleep for a healthy adult: a joint consensus statement of the American Academy of Sleep Medicine and Sleep Research Society. Journal of Clinical Sleep Medicine. 2015;11(6):591–592.


a
Metabolic Equivalents (VO2)
12.5
10.0
7.5
5.0
2.5
Rest
3 km.h−1
5 km.h−1
7 km.h−1
9 km.h−1
b
Activity Counts (counts·min−1)
GT3X+ (hip)
GT3X+ (wrist)
ACTT (hip)
ACTT (wrist)
12000
8000
4000
0
30000
20000
10000
60000
40000
Sex
Male
Female


a)
GT3X+ (hip)
GT3X+ (wrist)
ACTT (hip)
ACTT (wrist)
Pearson Correlation
1.00
0.98
0.95
0.92
0.90
b)
3 km.h⁻¹
5 km.h⁻¹
7 km.h⁻¹
9 km.h⁻¹
Rest
GT3X+ (hip)
ACTT (hip)
ACTT (wrist)
GT3X+ (wrist)
MET^1/2
Activity^1/2
c)
Confusion Matrix
True Class
Predicted Class
[0,3)
[3,6)
[6,9)
[9,inf)
0.92
0.07
0.08
0.88
0.06
0.06
0.56
0.19
0.38
0.80


a
GT3X+ (hip)
0.98
0.08
0.02
0.91
0.09
0.51
0.07
0.40
0.93
b
GT3X+ (wrist)
0.80
0.06
0.20
0.83
0.02
0.10
0.57
0.21
0.40
0.79
c
ACTT (hip)
0.97
0.06
0.03
0.92
0.03
0.02
0.02
0.59
0.18
0.38
0.80
d
ACTT (wrist)
0.95
0.07
0.05
0.84
0.06
0.09
0.61
0.25
0.32
0.75
True Class
[0,3)
[3,6)
[6,9)
[9,inf)
Predicted Class
Device
Placement