U.PORTO

FACULDADE DE DESPORTO
UNIVERSIDADE DO PORTO

ACT-PRED: Accelerometry-based Prediction of Energy Expenditure and Physical Activity Intensity Classification in Daily Tasks

Ana Daniela Barbosa Rodrigues de Oliveira

Porto, 2023

ACT-PRED: Accelerometry-based Prediction of Energy Expenditure and Physical Activity Intensity Classification in Daily Tasks

Dissertação apresentado com vista à obtenção do 2º ciclo em Atividade Física, Exercício e Saúde, da Faculdade de Desporto da Universidade do Porto, ao abrigo do Decreto-Lei nº 74/2006, de 24 de março, na redação dada pelo Decreto-Lei nº 65/2018 de 16 de agosto e com as alterações introduzidas pelo Decreto-Lei n.o 27/2021, de 16 de abril.

Orientador: Professor Doutor Hélder Rui Martins Fonseca

Co-orientador: Mestre Lucas de Souza Veras

Ana Daniela Barbosa Rodrigues de Oliveira

Porto, 2023

Ficha de catalogação

Oliveira, A (2023). ACT-PRED: Accelerometry-based Prediction of Energy Expenditure and Physical Activity Intensity Classification in Daily Tasks. Porto: Dissertação com vista à obtenção do 2º ciclo em Atividade Física, Exercício e Saúde, apresentado à Faculdade de Desporto da Universidade do Porto.

Palavras-chave: acelerómetro, acelerometria, atividade física, dispêndio energético, intensidade da atividade física

Acknowledgments

(Agradecimentos)

Em primeiro lugar, gostaria de agradecer à minha família, aos meus pais em particular, pelo apoio incondicional e por acreditarem em mim e me permitirem percorrer os meus sonhos. Sem o suporte deles isto não era possível. A vocês: obrigada será sempre pouco!

Ao meu namorado, Afonso, que me compreende, que me ajuda e facilita a vida em casa para que tudo o resto seja alcançável.

Aos meus amigos, que me acompanham neste caminho com uma palavra de motivação. Às minhas Joanas em especial porque me ajudaram muito neste caminho!

Aos meus colegas de mestrado, que partilharam comigo conhecimento, mas também muito boa disposição. E aos colegas que se tornaram amigos durante a aprendizagem.

A todos os professores que, ao longo deste percurso, se mostraram sempre disponíveis e com uma motivação que me levou a querer mais.

Agradeço também ao professor Hélder Fonseca, por toda a disponibilidade e dedicação não só enquanto orientador, mas como professor na sua totalidade. Obrigada por ter acolhido este desafio comigo.

O maior agradecimento ao Lucas, porque este trabalho nunca poderia ter sido realizado sozinha. Pelas manhãs que acordou cedo para irmos recolher dados, pelos dias que se deitou tarde a ver alguma coisa importante, por todo o tempo de despendeu para me ajudar. O meu maior agradecimento é para ti: obrigada por tudo!

Um obrigada a todos os que não estão referidos individualmente mas que me acompanharam e que, de alguma forma, contribuíram para este processo.

«Aqueles que passam por nós não vão sós.

Deixam um pouco de si, levam um pouco de nós.»

Antoine de Saint-Exupery, O Pequeno Príncipe.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

AC	activity counts
BMI	body mass index
CV	coefficient of variation
DLW	double-labeled water
DXA	dual-energy x-ray absorptiometry
EE	energy expenditure
ENMO	euclidean norm minus one
GRF	ground reaction forces
LMM	linear mixed models
LOOCV	leave-one-out cross-validation
MAD	mean amplitude deviation
MAE	mean absolute error
MAPE	mean absolute percent error
MET	metabolic equivalent
MPA	moderate physical activity
NHANES	national health and nutrition examination survey
PA	physical activity
PAEE	physical activity energy expenditure
PAI	physical activity intensity
R^2	coefficient of determination
RMR	resting metabolic rate
RMSE	root mean square error
ROC	receiver operating characteristic
RQ	Respiratory quotient
SB	sedentary behavior
SD	standard deviation
SE	sensitivity
SEE	standard error of estimation
SP	specificity

VCO_2	carbon dioxide production
VO_2	oxygen uptake
VPA	vigorous physical activity

Resumo

Fundamento: Os acelerómetros são comumente utilizados para estimar tempo sedentário e dispêndio energético (DE) em atividade física. Contudo, os critérios de recolha e processamento de dados provenientes de acelerómetros evoluíram numa multiplicidade de abordagens. Consequentemente, não existe um método universalmente aceite. **Objetivo**: Este estudo teve como propósito desenvolver equações de previsão do DE e estabelecimento de pontos de corte para classificação de tarefas diárias com base em diversas métricas provenientes de acelerómetros posicionados na anca e punhos dominante e não dominante. **Métodos**: 49 adultos (18 a 65 anos) completaram sete atividades distintas enquanto usavam três Actigraph GT3X+ (anca direita, punho dominante e não dominante). O volume de oxigénio consumido (VO_2) e de dióxido de carbono produzido (VCO_2) foram simultaneamente medidos por calorimetria indireta. As métricas de aceleração analisadas foram *activy counts*, *mean amplitude deviation* (MAD) e *euclidean norm minus one* (ENMO). Foram utilizados modelos lineares mistos para previsão do DE e curvas ROC (*receiver operating characteristic*) para determinação dos pontos de corte para classificação de intensidades. Gráficos de Bland-Altman, erros da previsão, estatística *Kappa* e concordância percentual foram calculados usando dados de validação cruzada de exclusão. **Resultados**: Os modelos de predição apresentaram uma equação quadrática que teve como preditores uma das métricas do acelerómetro, idade, sexo e índice de massa corporal. O DE previsto teve boa concordância, com um erro percentual absoluto médio inferior a 16% para a anca aplicando a métrica ENMO. A classificação global dos pontos de corte desenvolvidos foi caracterizada como quase perfeita, com concordância percentual acima de 63%, utilizando o acelerómetro no punho não dominante e a métrica MAD. O erro da previsão e a classificação dos pontos de corte apresentaram variações no posicionamento do acelerómetro bem como na métrica escolhida. **Conclusões**: As atividades sedentárias são melhor previstas com a colocação do acelerómetro no punho não dominante e as atividades vigorosas com a colocação na anca. A colocação na anca parece ser melhor por ser mais central, mas em estudos de longo prazo seria interessante considerar o punho não dominante, pois aumenta a adesão ao protocolo do estudo.

Palavras-chave: acelerómetro ou acelerometria, actigraph, atividade física, dispêndio energético, intensidade da atividade física

Abstract

Background: Accelerometers are commonly used to estimate sedentary time, physical activity, and physical activity energy expenditure. Due to the need to answer unique research questions, data collection and processing criteria have evolved in many different approaches. Consequently, there is no universally accepted method. **Purpose:** This study aims were to develop regression equations to predict energy expenditure (EE) and cut-points to classify daily tasks based on several metrics obtained from hip, dominant, and non-dominant wrist accelerometer placement data. **Methods:** 49 adults (18– 65 yr) completed seven activities (ranging from lying to running) while wearing three Actigraph GT3X+ on the right hip, dominant wrist, and non-dominant wrist. Oxygen consumption (VO_2) and carbon dioxide produced (VCO_2) were measured by indirect calorimetry. Accelerometer metrics analyzed were activity counts, mean amplitude deviation (MAD) and euclidean norm minus one (ENMO). Linear mixed models were used to predict EE and receiver operating characteristic (ROC) curves to determine cut-points to classify sedentary activities (SA) and physical activity intensity. Bland-Altman plots, prediction accuracy, Kappa statistic and percent agreement were calculated using leave-one-out cross-validation data. **Results:** All prediction models presented a quadratic equation that had as predictors one of the accelerometer metrics, age, sex, and body mass index. Predicted EE indicated a good agreement presenting a mean absolute percent error below 16%, on the hip using ENMO. Global classification agreement from developed cut-points was categorized as almost perfect with a percent agreement above 63%, using the accelerometer on the non-dominant wrist and the accelerometer metric MAD. Prediction accuracy and classification agreement presented variations in the placement of the accelerometer as well as the chosen metric. **Conclusions:** SA are better predicted with non-dominant wrist placement and vigorous activities with hip placement. Hip placement seems to be better as it is more central, but in long-term studies, it would be interesting to consider the use of the non-dominant wrist as it increases adherence to the study protocol.

Keywords: accelerometer or accelerometry, actigraph, energy expenditure, physical activity, physical activity intensity

1. General Introduction

General Introduction

In an era characterized by sedentary lifestyles and the rising burden of chronic diseases, the promotion of physical activity (PA) has emerged as a crucial strategy for improving public health and well-being. Scientific research has unequivocally demonstrated the multifaceted benefits of regular PA for enhancing physical health and its positive impact on mental health, cognitive function, and overall quality of life (Rebar et al., 2015; Warburton et al., 2006).

This dissertation delves into the intricate relationship between PA, energy expenditure (EE), and health, with a specific focus on the potential of accelerometry-based methods to assess and predict these parameters in the context of daily life.

PA encompasses a wide spectrum of bodily movements, from structured exercise routines to everyday tasks such as walking, climbing stairs, and household chores. The benefits of PA for maintaining health have been recognized and well-documented, especially in the prevention and treatment of chronic diseases, such as cardiovascular diseases (Lee et al., 2012; Mozaffarian et al., 2008), type 2 diabetes by enhancing insulin sensitivity and glucose regulation (Sigal et al., 2006), and some types of cancer (Alves et al., 2016; Hamasaki, 2016; Kirkham & Davis, 2015; Montoye et al., 2015), including breast and colon cancer, as well as improvements in bone health and musculoskeletal function (Borer, 2005; Friedenreich et al., 2010). Therefore, maintaining a physically active lifestyle is essential for a healthy and high-quality life (Sirichana et al., 2017).

The intensity of PA is a key determinant of its health benefits. Moderate to vigorous-intensity activities have been shown to confer greater cardiovascular and metabolic benefits compared to light-intensity activities (Haskell et al., 2007). However, accurately quantifying the EE associated with various PAs, especially those performed during daily tasks, presents a considerable challenge. Daily living activities can contribute to moderate expenditure, depending on the oxygen consumption required to perform the task. However, quantifying daily PA is challenging due to the variation in intensity and duration of activities throughout the day (Sirichana et al., 2017). Traditional methods, such as indirect calorimetry, have provided valuable insights into EE but are limited by their feasibility for real-world applications and associated costs (Bassett et al., 2010; Hills et al., 2014).

Understanding the energy demands of daily tasks is essential for several reasons. First, it enables the assessment of an individual's adherence to recommended PA guidelines and provides insights into overall activity patterns. This information is pivotal for designing targeted interventions to promote active lifestyles and mitigate the risk of chronic diseases. Secondly, quantifying EE during daily tasks can aid in developing personalized exercise prescriptions, optimizing training regimens for athletes, and tailoring rehabilitation protocols for individuals recovering from injuries (Sallis et al., 2016). Thirdly, the ability to accurately measure physical activity intensity (PAI) during daily tasks contributes to unraveling the complex interplay between activity patterns and health outcomes (Montoye et al., 2015; Strath et al., 2015).

In recent years, technological advancements have paved the way for novel methods of assessing PA and EE in real-world settings (Hernández-Vicente et al., 2022; Ndahimana & Kim, 2017). Accelerometers, small wearable devices that measure movement and acceleration, have gained prominence as tools for capturing activity patterns and intensity levels. These devices offer a unique opportunity to bridge the gap between laboratory-based assessments and the dynamic nature of daily activities (Freedson et al., 1998).

This dissertation seeks to explore the potential of accelerometry-based methods in predicting EE and classifying PAI during daily tasks. This research aims to develop a framework for accurately estimating EE and categorizing PAI across a spectrum of everyday activities. The integration of accelerometry data with advanced modeling techniques holds the promise of enhancing our understanding of the energy demands associated with daily tasks and informing interventions aimed at promoting active lifestyles and preventing chronic diseases.

As the global burden of chronic diseases, many of which associated with sedentary behaviors, continues to escalate, it is imperative to prioritize strategies that promote PA and mitigate sedentary behavior (SB). This dissertation embodies a forward-thinking approach that capitalizes on technological innovations to advance our ability to measure, predict, and understand PAI and EE in the context of daily tasks. By shedding light on the intricate relationship between PA and health, this research has the potential to inform public health policies, to guide clinical practice, and to empower individuals to make informed choices that can positively impact their well-being.

2. Literature Review

2.1. Introduction

The association between PA and positive health outcomes has been well-established. Thus, maintaining a physically active lifestyle is essential for a healthy and quality life. Since there is a strong correlation between PA levels and disease/mortality risk (Blond et al., 2020), the American College of Sports Medicine (ACSM) classified PAI according to the level of metabolic equivalents (METs), which represent multiples of resting energy expenditure, into light (<3 METs), moderate (3-6 METs) and vigorous (>6 METs) intensity. At least 150 minutes of moderate-intensity aerobic physical activity per week is recommended to promote health. Although the benefits of regular, moderate-intensity physical activity (MPA) have been shown, objectively quantifying daily PA has proven to be a difficult task due to the variation in intensity and duration of activities throughout the day.

The gold standard method to measure the EE on free living conditions is the double-labeled water (DLW) (Westerterp, 2017). However, the high costs associated as well as analytical requirements limit its feasibility in large studies. Indirect calorimetry is one of the most frequently used methods as a criterion measure for the assessment of EE, but it remains expensive and is mostly limited to laboratory conditions. The use of wearable activity monitors, such as accelerometers, has become increasingly popular during the last decade to estimate EE. Accelerometers are objective measurement tools that enable the measurement of accelerations and that researchers can use to estimate the intensity of a given activity and therefore the quantity of energy that individuals expend performing it as well as to quantify the amount of time spent engaging in activities of different intensities. Because of their technical advantages and accessible price, the use of accelerometers has been growing in recent years (Pedišić & Bauman, 2015), both in physical activity epidemiological research (Ekelund et al., 2019) as well as in consumer grade wearable devices such as sport watches and physical activity trackers (Wright et al., 2017).

2.2. Benefits of Physical Activity for Health

PA is undeniably paramount for overall health and well-being. This is corroborated by a prosperity of scientific evidence (Kaminsky & Montoye, 2014; Warburton & Bredin, 2017) gathered, at least, during the last fifty years. With technological advancement, SB has increased dramatically and is now a global health concern, linked to various chronic diseases and adverse health outcomes (Guthold et al., 2020; Lee et al., 2012; Shiroma et al., 2014). In contrast, engaging in regular PA provides many benefits across physical, mental, and emotional fields. Scientific studies consistently underline the role of PA in reducing the risk of chronic conditions, including cardiovascular disease, type 2 diabetes, obesity, and several types of cancer (Alves et al., 2016; Hamasaki, 2016; Kirkham & Davis, 2015; Lee et al., 2012; Montoye et al., 2015; Mozaffarian et al., 2008; Sigal et al., 2006). The benefits of PA for bone health and for promoting healthy aging have also been thoroughly documented (Bielemann et al., 2013; Friedenreich et al., 2010). PA also plays a role in lowering the risk of premature death (Warburton & Bredin, 2017), emphasizing its significance in increasing life expectancy. These findings collectively accentuate that integrating regular PA into daily life is a fundamental necessity for achieving and maintaining optimal health and vitality. Therefore, to define public health policies and to quantify population-level risk for a variety of diseases, chronic disorders, and functional outcomes, accurate quantification of physical activity energy expenditure (PAEE) patterns is crucial (Liu et al., 2021).

2.3. Daily Tasks and Their Impact on Physical Activity

PA, as defined by the scientific literature, comprehends any bodily movement produced by skeletal muscles that results in EE beyond the resting state (Caspersen et al., 1985). PA includes a broad range of bodily motions, including everyday activities like walking, climbing stairs, and household chores, as well as structured physical exercise. Daily tasks integrated into one's household or labor routine can substantially contribute to increase daily PA. Active living, characterized by walking or cycling for transportation, using stairs, and participating in household chores, offers a practical means of enhancing daily PA (Matthews et al., 2007).

2.4. Measuring Energy Expenditure: Methods and Concerns

Accurate and reliable measurement of EE in daily life is crucial for assessing PA’s impact on health (Strath et al., 2013).

Subjective and objective methods fall into two major groups for evaluating PA. Subjective techniques, like questionnaires, rely on the individual to either keep track of events as they happen or to remember events from the past. These self-reported methods are cost-effective but are subject to recall bias (Migueles et al., 2017) and therefore are highly imprecise (Prince et al., 2008). All wearable monitors that directly measure one or more biological signals, such as heart rate, acceleration, or another measure of PA or EE, as they happen, are considered objective techniques. There are numerous methods available to objectively assess PA. Traditional methods, such as DLW or indirect calorimetry, provide accurate measurements but can be costly and impractical for large-scale studies (Liu et al., 2021; Westerterp & Plasqui, 2004). Wearable devices fall into this objective category, where accelerometers are the most commonly wearable device used to assess PA (Strath et al., 2013). More in-depth information about each method can be found in the following review manuscript (Ndahimana & Kim, 2017).

2.5. Accelerometers: Unveiling the Mechanics of Motion

An accelerometer is a wearable device designed to measure acceleration. This foundational idea in physics denotes the variation of velocity in time. Accelerometers are a vital tool for measuring PA in the context of human movement. Researchers and healthcare providers can analyze activity patterns and evaluate their effects on health by interpreting acceleration as a stand-in for movement (Liu et al., 2021; Strath et al., 2013; Troiano et al., 2014).

In the 1980s, the first accelerometer-based PA monitor to be adopted by researchers was developed (Montoye et al., 1983; Wong et al., 1981). When they were first introduced, accelerometer-based devices were viewed as an innovative and specialized evaluation technique but that also had several drawbacks, such as high device costs, dependability, calibration and validity issues (Chen & Bassett, 2005; Troiano et al., 2014). Nevertheless, the apparent utility of objective PA data gathered over several free-living days proved to be very attractive to many PA researchers by the early 2000s when accelerometer technology became more widely available (Troiano, 2005). Accelerometers are now used for a wide variety of PA measurements ranging from daily physical activity characterization to high level professional sport (Liu et al., 2021; Mendes et al., 2018; Migueles et al., 2017; Troiano et al., 2014).

These devices measure motion and acceleration and offer an objective and unbiased evaluation of PA. They measure accelerations of body movement in either 1 plane (usually vertical), 2 planes (vertical and mediolateral or vertical and anterior-posterior), or 3 planes (vertical, mediolateral, and anterior-posterior) (Chen & Bassett, 2005). Their output can be delivered as measurements of activity counts (AC) or raw acceleration data, each one with positive and negative aspects (Chen & Bassett, 2005; Troiano et al., 2014).

As with any other method of measurement of PA, accelerometers have some advantages and disadvantages. In terms of advantages, they are appropriate for large-scale investigations since they are portable, wearable, and non-invasive. They are capable of capturing the frequency, duration, and intensity of physical movement in a time-stamped manner and allow for continuous monitoring, capturing daily activity, offering a wide-ranging view of an individual's movement pattern (Matthews et al., 2018; Strath et al., 2013). Despite these benefits, accelerometers also have some weaknesses. They might not be able to distinguish between different types of physical activities with accuracy and have some limitations on non-linear activities. For example, their capacity to offer the best information for some sports, like cycling or weightlifting, is limited (Matthew, 2005; Mendes et al., 2018). Differences in fitness level and metabolic rate are also not considered by accelerometers. Additionally, accelerometers can misidentify inactive movements such as vibrations from riding a motorized vehicle as PA (Le Masurier & Tudor-Locke, 2003; Liu et al., 2021). Lastly, different manufacturers use various proprietary undisclosed algorithms to convert raw data into AC, making it difficult to compare the output of different brands and equipment's (Kavanagh & Menz, 2008; Plasqui et al., 2013).

One of the key decisions to be made when designing an accelerometer-based study is the placement of the accelerometer, where consensus has yet to be reached (Troiano et al., 2014). Accelerometers are attached to specific body parts, most frequently the wrist, hip, chest, thigh, and lower back. According to the review published by Liu et al., research using wrist-worn devices has been rising as part of initiatives to lessen participant burden and promote compliance (Liu et al., 2021). The evidence on the effect of placement for estimating EE is divided. Ankle placement may be more accurate for step count estimation, thigh placement may distinguish better postures and types of activity, while wrist-worn seems to increase compliance in research and clinical settings (Crowley et al., 2019; Liu et al., 2021; van Hees et al., 2011). Some studies also found that some body placements, like the hip or wrist, could be comparable (Ellis et al., 2014; Troiano et al., 2014). Also, different placements may respond differently to diverse PA types and intensities. While wrist location may be more effective at identifying daily activities that include more upper body movement, such as housework, the ankle and hip placement may be better at assessing moderate to vigorous physical activity (VPA) (Duncan et al., 2020; Liu et al., 2021). Due to facts such as greater adherence and compliance with the use of accelerometers on the wrist, as well as less burden on participants, the United States National Health and Nutrition Examination Survey (NHANES) changed from waist to wrist placement in the 2011-2012 and 2013-2014 survey cycles (Troiano et al., 2014). Hence, the placement of the accelerometer must consider what is intended to be measured.

2.6. Accelerometers: Calibration & Validation

Firstly, it is necessary to distinguish between two types of calibration for wearable activity monitors. According to Bassett et al. (2012) ‘‘unit calibration’’ is made to reduce inter-instrument variability and to ensure that devices are correctly measuring the direct signals, and ‘‘value calibration’’ denotes the procedure used to convert the direct signals into other established measurement units. The remaining of this section will address value calibration.

Value calibration refers to the process by which researchers obtain data that allow them to convert direct signals from monitors into estimates of EE, time spent in various physical activity intensity categories and categorize activity types (Bassett et al., 2012).

The term “validity” is also important and refers to whether an instrument measures what it is intended to measure. There are different types of validity, with criterion-referenced validity being the most used by researchers in the field of PA since the variables to be measured are objective (Bassett et al., 2012). In the end, the prediction results for EE obtained through accelerometry will be compared with the values measured through indirect calorimetry, which is the reference criterion used in this study.

The output data from accelerometers can either be raw acceleration, often reported in gravitational acceleration units (g), or AC, which are processed data obtained from raw acceleration and that are based on a, frequently undisclosed, manufacturer-specific algorithm (Bassett et al., 2012; Chen & Bassett, 2005; Troiano et al., 2014). Both outputs can represent overall movement, but a calibration procedure is required to convert them into more biologically meaningful data such as metabolic equivalents for instance (Matthew, 2005; Welk, 2005). The majority of calibration studies use the accelerometer output to determine EE and PAI levels (Crouter et al., 2018; Diniz-Sousa et al., 2020; Duncan et al., 2020; Mendes et al., 2018; Migueles et al., 2017; Migueles et al., 2019). Its use is also employed to assess SB (Chastin et al., 2014; Gao et al., 2021; Staudenmayer et al., 2015) and sleep monitoring (Song et al., 2023). Additionally, they can be used to estimate biomechanical parameters, like ground reaction forces (GRF) (Fortune et al., 2014; Neugebauer et al., 2014; Veras et al., 2023), to detect falls (Bourke et al., 2007), to recognize gait patterns (Moe-Nilssen & Helbostad, 2004), analyze posture (Leone et al., 2023), detect the type of PA being performed (Zhang et al., 2012), among other applications.

In the first studies using accelerometers, AC were the sole metric used as this was the only one available at the time. However, as previously mentioned, the AC calculation relies on manufacturer-specific algorithms, which leads different accelerometers to produce different count values even when measuring the same accelerations (Plasqui et al., 2013). Back then, acceleration data was summarized in AC because accelerometers could not store all the raw information. Technological advances have allowed the collection and storage of acceleration data at high frequencies which has thus eliminated the need to summarize it in AC. The use of raw acceleration data enhances comparability among accelerometers from different manufacturers (Mendes et al., 2018). Furthermore, the capability of extracting time and frequency domain elements from the data enables the application of more sophisticated statistical and computational approaches during the calibration process (John et al., 2013; van Hees et al., 2016). Researchers can harness these advantages to gain deeper insights into human movement and behavior across various fields of research. Current suggestions for standardizing the methodology used in accelerometer research also recommend that calibration studies replace arbitrary AC with new metrics based on raw acceleration, such as mean amplitude deviation (MAD) (Vähä-Ypyä et al., 2015) and Euclidean norm minus one (ENMO) (van Hees et al., 2013). These new accelerometer metrics should demonstrate parity or greater precision compared to the long-standing AC units (Bassett et al., 2012). The use of raw data has become more ordinary and the research field would benefit from creating a consensus approach for reporting a set of standardized metrics to promote harmonization across studies (Evenson et al., 2022). In addition to comparing the criterion measurement with the accelerometer output, there are other parameters that calibration studies must take into account, such as defining the sample characteristics, PA protocols, and statistical methodology, since these aspects can influence the validity of the study (Bassett et al., 2012; Welk, 2005).

Accelerometer calibration studies must take certain methodological considerations into account. The sample must be representative of the intended population in terms of age, weight, or body mass index (BMI) (Strath et al., 2012). Thus, it is important that calibration studies research specific populations such as children (Migueles et al., 2019), adults (Ellingson et al., 2017), obese (Diniz-Sousa et al., 2020), and the elderly (Duncan et al., 2020). As already mentioned, the placement of the accelerometer influences the output and therefore distinct calibrations are necessary for each position. Additionally, a varied range of PA, representing activities usually performed by the target population, should also be performed during calibration procedures in intensities ranging from sedentary to vigorous as well as activities that incorporate tasks usually performed by this population (Bassett et al., 2012; Welk, 2005).

2.7. Energy Expenditure: Prediction and Intensity Classification

Technological development has led to an increase in accelerometer use, both from a scientific point of view in a research context and by the general population using wearable devices that promise to estimate the total amount of energy spent during the day or in PA and to categorize activities according to their intensity (Füzéki et al., 2017; Menai et al., 2017). The first research using accelerometers to estimate EE dates back to the 80s, when the portable accelerometer device was created (Wong et al., 1981) and a few years later tested (Montoye et al., 1983). Thenceforth, as technology advanced, novel methodologies and statistical methods have been developed, with the same goal – to increase the accuracy of accelerometer estimates.

Melanson and Freedson (1995) developed the first linear regression model based on accelerometers to predict EE in kilocalories per minute (kcal·min^{-1}). The population in this study were young adults (15 males, 13 females) with normal weight, and they wore three accelerometers at different anatomical regions – hip, ankle, and wrist. The protocol consisted of treadmill activities with three different speeds – slow walking (4.8 km·h^{-1}), fast walking (6.4 km·h^{-1}), and jogging (8.1 km·h^{-1}) – and the criterion measure applied was indirect calorimetry. Different combinations of predictors were used in the forecasting models, and if the EE prediction was made only from a single accelerometer, then the best model used body mass (in kg) and AC measured at the wrist as predictors. This model explained 82% of the variance in EE and resulted in a standard error of estimation (SEE) value of 1.05 kcal·min^{-1} (Melanson & Freedson, 1995). Freedson et al. (1998) continued the work carried out, applying a similar protocol to the previous, but in this research, they used a more evolved model of the accelerometer (CSA 7164, the previous was a CSA 5032), and only one at the right hip. The linear regression was developed with AC and body mass (in kg) as predictors and firmly explained the variation in EE ($R^2 = 0.82$) (Freedson et al., 1998).

However, elaborating studies with activities made only on a treadmill was a very limited option to record what happens in everyday life. Swartz et al. (2000) were among the first researchers to include lifestyle activities such as gardening, carrying weighted items, vacuuming, washing dishes, among others, in their linear regression models of EE estimation. Nevertheless, the different results demonstrated an over- or under-estimation of the regression for individual activities. They significantly underpredicted activities like mowing with a power mower and a manual mower ($P = 0.001$) and overpredicted the energy cost for ironing and caring for children as well as slow walking (Swartz et al., 2000). In the past, it was also demonstrated that the regression equations created for walking and running slightly overstated the EE for walking and light physical activities (LPA), whereas they significantly underestimated the EE for moderate-intensity lifestyle activities. The lifestyle regression equations offer a more accurate estimate for the EE of moderate-intensity activities, but they seriously overstate the energy cost of sedentary and light activities while underestimating the EE for vigorous activities (Bassett, 2000). It was therefore realized that a different approach would be necessary, as a forecast model that included running/walking and tasks of daily living would result in high errors in the estimation of EE. To overcome this limitation, a two-regression model was created by Crouter et al. (2006). For their study, an accelerometer placed at the right hip was used and the study included locomotor activities like walking and running at various speeds along with lifestyle, leisure, and sports activities. A method based on the coefficient of variation (CV) of AC was developed to differentiate between locomotor activities and other activities. If the CV per 10 seconds was less than or equal to 10 a locomotor activity would be assumed, and an exponential curve ($R^2 = 0.701$) would be applied. If the CV was greater than 10, free-living activities would be assumed and a cubic curve ($R^2 = 0.854$) would be applied (Crouter et al., 2006). This new approach was able to successfully increase EE prediction accuracy compared to previous studies and was also refined by the same group of researchers later in 2010 (Crouter et al., 2010).

More recently, several studies have been carried out with the application of accelerometers to estimate and validate EE. Various authors (Ho et al., 2019; Lee & Tse, 2019; Neil-Sztramko et al., 2017) have conducted research on the measurement of EE with activities like walking and running. However, this kind of research fails to facilitate accurate and reliable information on the EE on daily tasks. Other authors (Ellingson et al., 2017; Ellis et al., 2014; Montoye et al., 2015; Sirichana et al., 2017; Staudenmayer et al., 2015; Strath et al., 2015), in addition to running and walking, have used protocols that included daily activities performed in the laboratory, such as supine resting, sitting and standing reading a book/typing/fidgeting, climbing stairs, throwing/catching a ball, stationary biking, walking while carrying groceries, among others. This is similar to what was done in our experimental study.

There is also a need to define cut-off points for different intensities of PA. For this purpose, different approaches are applied by many authors. Hildebrand et al. (2014) and Lee & Tse (2019) used linear regression models, but the first authors applied these with ENMO (Hildebrand et al., 2014; Lee & Tse, 2019). Neil-Sztramko et al. (2017) and Rhudy et al. (2020) identified cut-off points maximizing sensitivity (SE) [true positives/ (true positives + false negatives)] and specificity (SP) (true negatives/(true negatives + false positives)]. Ducan et al.

(2020) also used the approach to maximize SE and SP, but with the use of ENMO (Duncan et al., 2020; Neil-Sztramko et al., 2017; Rhudy et al., 2020).

Our study is a culmination of studies accomplished in the past, as it uses activities that are similar to what individuals perform during their daily lives, such as sitting at the computer or carrying groceries, together with locomotor activities such as walking and running. The EE prediction is based on different metrics to reduce the estimation error. Cut-off points for the different PA intensities will also be estimated based on METs, giving information to the community on how much energy each activity “spends”, thus helping people to make informed changes in their daily physical activity, such as climbing stairs instead of using the elevator or escalators to increase EE and reach the recommended amount of daily physical activity endorsed by the WHO.

2.8. Accelerometer & Energy Expenditure: The Statistical Approach

The statistical methodology used in calibration studies is another important factor. A common technique is to create a regression model to convert the accelerometer output into an outcome variable, such as estimations of EE. More than a dozen regression equations have been created for ActiGraph alone. (Bassett et al., 2012; Strath et al., 2012). Most calibration studies that resort to approaches that use multiple data points for each individual technically violate the independence assumption of multiple regression (Welk, 2005). Implementing the linear mixed model approach, which enables the analysis of repeated data, is one possibility to address this issue. Additionally, mixed models allow testing of quadratic and cubic trends. Mixed models design flexibility and interpretability represent a substantial advancement in the methodology for calibration research (Welk, 2005). Although the application of mixed models has an important role in calibration studies, currently, other methods have also been used more frequently, namely machine-learning techniques (Farrahi et al., 2019; Pfeiffer et al., 2022). In addition to modeling the accelerometer output, these techniques also use statistical summaries of data in time and frequency domains to thoroughly describe the acceleration pattern (Staudenmayer et al., 2015). They increase accelerometer prediction accuracy, specifically for sedentary and non-locomotor activities (Montoye et al., 2017). However, if both machine learning and regression models have similar accuracies, regression models should be prioritized, since they are more simple to apply and interpret (Montoye et al., 2017).

Receiver Operating Characteristics (ROC) curve analysis is another approach that has been used for establishing activity intensity cut-points (Mendes et al., 2018). A ROC curve consists of a graphical representation of the SE and SP of distinct cut-points. The SE is plotted on the y-axis and the SP on the x-axis. Top-left corner coordinates (0, 1) are considered the perfect classification. A well-behaved ROC curve is accounted for when the values rise precipitously on the y-axis, pending to the top-left corner, indicating that the method being tested has an elevated SE and a low false-positive rate, and shows favorable discrimination properties (Welk, 2005).

Another crucial aspect in accelerometer research is to assess the performance of the accelerometer calibration results (Bassett et al., 2012). In this regard, the statistics used for this evaluation tend to be biased, and are frequently too optimistic, when the method's developed for the assessment are conducted on the same sample used for validation. As a result, it is necessary to cross-validate the prediction equations with another sample (Staudenmayer et al., 2012). In ideal conditions, the evaluation is made by using another sample from the target population. However, this approach could be costly, hence a split-sample cross-validation method is commonly used. The division of the study sample in two parts (one for calibration and one for cross-validation) would be a good strategy, however this could be difficult when studying small sample sizes. The leave-one-out cross-validation (LOOCV) method could overcome this issue, in which one participant’s data is detached in a testing dataset, maintaining the remaining participants in the training dataset, and repeating the procedure until each participant is used in the testing dataset (Staudenmayer et al., 2012).

3. Methods

3.1. Participants

A convenience sample of 49 adults were recruited for this study from the population of the University of Porto following an online advertisement. Of these, 28 were males, the average age was 28.4 ± 8.5 years, average height 168.7 ± 8.3 cm and average weight of 69.4 ± 13.5 kg with a body mass index (BMI) of 24.5 ± 3.7 kg·m^{-2}; $\bar{X}$ ± SD). Height was measured using a stadiometer (Seca 213, Hamburg, Germany) and weight with a digital scale (Seca 899, Hamburg, Germany) following standard procedures (CDC, 2017). Inclusion criteria for participation in the study were: age between 18-65 years, no self-reported neurological or musculoskeletal limitations that could compromise the performance of the protocol tasks, familiarization with the activities of the protocol, namely walking and running in a treadmill, and everyday life conditions. Exclusion criteria were: presence of any health condition (recent body weight change >10%; hypo- or hyperthyroidism) or medication that could significantly affect energy expenditure (e.g. stimulants, beta-adrenergic drugs, anti-cholinergic drugs, thyroid hormones), a health condition that could contraindicate the performance of any of the assessments covered in the protocol (e.g. pregnancy, phobias). Through a questionnaire carried out at the time of the assessment, information was collected on lifestyle factors (smoking, alcohol consumption, physical activity/exercise, sleep duration, last meal, fasting time) as well as current medication. Participants were instructed to perform an overnight fast for at least 12 hours (but stay hydrated by drinking water), to not consume alcoholic beverages or coffee in the preceding 12 hours, to not perform vigorous physical activity in the last 24 hours, to not smoke on the day of the assessment, try to get good quality and quantity of sleep, and empty their bladder before the assessment. All participants were informed about the experiments purpose and protocol before giving written informed consent. The study protocol was approved by the local Ethics Committee (CES 192-14).

3.2. Protocol

The data collection was conducted at the Research Centre in Physical Activity, Health, and Leisure (CIAFEL) of the Faculty of Sport of the University of Porto (FADE/UP). The protocol was divided into three parts. In the first part, a whole body composition assessment was performed using Dual-Energy X-ray Absorptiometry (DXA), lasting approximately seven minutes. Secondly, the participant laid down in a resting supine position for 30 minutes to assess resting energy expenditure through indirect calorimetry (Quark CPET, Cosmed, Rome, Italy). During this measurement, the subject was left alone in a calm room with light off at 22ºC ambient temperature. To ensure an accurate measurement of RMR, participants were asked to fast for at least 12 hours, not smoke or drink coffee on the day of the assessment and avoid intense physical activities the day before. The data set used in the analysis of this study corresponds to the 10 minutes where the variation in RMR and RQ was lowest. The values of VO_2 and VCO_2 were used to calculate EE with the Weir's equation (Weir, 1949). In the third part, several pre-established activities were performed in the order listed below for five minutes each, with one-minute resting intervals between activities. The tasks were as follows: (1) sitting and typing on a computer; (2) standing and transferring water from one basin to another, (3) carrying bags with weights, organizing the weights, dusting shelves, sweeping the floor, and picking up trash, storing and arranging a set of sheets and books; (4) walking at 4 km/h; (5) walking at 6 km/h; (6) running at 9 km/h; (7) climbing stairs. All walking and running tasks were performed on a treadmill (quasar 4.0; h/p/cosmos, München, Germany). A member of the research team made sure that the accelerometer monitor was placed correctly before the beginning of each task. Accelerometer and respiratory gases data were continuously gathered throughout the 7 activities.

3.3. Measurements

While completing the activities, participants wore three activity monitors (AM) from ActiGraph (GT9X Link, 100Hz; ±16g range; ActiGraph, Pensacola, USA). The AM placement for these activities were as follows: i) right hip (along the anterior axillary line, at the level of the iliac crest); ii) non-dominant wrist (at the level of the styloid process of the ulna), and (iii) dominant wrist. During the protocol, energy expenditure was measured using a portable indirect calorimetry system (Cosmed K5, Cosmed, Rome, Italy) worn on the back. Participants wore a face mask covering the nose and mouth to continuously capture gas exchange while oxygen consumption (VO_2) and carbon dioxide produced (VCO_2) were measured breath by breath. Heart rate was also assessed using a heart rate strap monitor placed on the participant's chest at the level of the xiphoid process.

3.4. Data processing and statistical analyses

The data collected from both the accelerometers and the indirect calorimetry system were subsequently processed using MATLAB (R2022b, MathWorks, USA). Data from all devices was synchronized by using the timestamp from the CosmedK5. After the synchronization, data was divided into the tasks mentioned above (from 1 to 7). Then, from the 5-minute interval, the first and last minutes were removed from all tasks, except task 3. The first and last minutes were removed to obtain greater data stability. This was not done in task 3 as it was a circuit with multiple small tasks, and crucial data would have been removed. Afterwards, different accelerometer metrics such as AC, MAD (Vähä-Ypyä et al., 2015), and ENMO (van Hees et al., 2013) were calculated. All these metrics were computed based on the resultant vector raw acceleration, stored using 5-second epochs. Finally, the average from 5-second epochs on each task was used for statistical analysis.

Data processing and statistical analyses were conducted using R statistical software (R version 4.2.2, R Foundation for Statistical Computing, Vienna, Austria). To develop EE prediction models, linear mixed models (LMM) were applied. Diverse LMMs were tested with data derived from the hip and wrist (dominant and non-dominant) worn accelerometers. Different models were developed for each accelerometer placement and metric. As predictors for the models developed, the accelerometer metric (AC, MAD, ENMO), sex, age, body mass, BMI ($kg.m^{-2}$) and BMI categories were tested. BMI categories were defined as follows: underweight BMI< 18.5; normal weight 18.5 ≤ BMI < 24.9; overweight 25 ≤ BMI < 29.9; obesity class I 30 ≤ BMI < 34.9, obesity class II 35 ≤ BMI < 39.9 and obese class III BMI ≥ 40.

Linear, quadratic, exponential and logarithmic models were also tested on each accelerometer metric. Final models were chosen according to -2log-likelihood statistics. The conditional coefficient of determination (R^2) was also calculated (Nakagawa & Schielzeth, 2013). Variables included in the final models were the linear and quadratic term of all the accelerometer metrics, age, sex and BMI category.

The cut-points to categorize PAI created through AC, MAD and ENMO were obtained by applying ROC curves, for the hip and both wrists. The indicators used to summarize the cutoff points were SE, SP, and area under the curve. Activities were classified as: SA if ≤1.5 MET; light physical activity (LPA) if between 1.6–2.9 MET; moderate physical activity (MPA) if between 3.0–6.0 MET and vigorous physical activity (VPA) if >6 MET (Health & Services, 2018). LPA limits were provided with SA and MPA cut points. Model validation was performed using the leave-one-out cross-validation (LOOCV) method (Staudenmayer et al., 2012). Dataset obtained from LOOCV were used in the following validation analyses. For the EE prediction models, accuracy was assessed by mean absolute error (MAE), mean absolute percent error (MAPE), and root mean square error (RMSE). Also, Bland-Altman plots were used to visualize the agreement between the actual EE values and the values predicted using the models. Furthermore, to determine if the models systematically under or overestimated the predicted values, one-sample *t*-tests were conducted to ascertain if the biases notably deviated from zero. Additionally, linear regressions were used to see if the differences between the actual and predicted values were influenced by their magnitude. For the cut-points, the percent agreement and Kappa statistic (both individual and global agreement) were used to measure the classification agreement of SA and PAI obtained from indirect calorimetry and those obtained from cut-points being the agreement qualitatively classified as fair if between .21–.40, moderate if between .41–.60, substantial if between .61–.80, and almost perfect if between .81–1.00 (Landis & Koch, 1977).

4. Results

Results

Accelerometer-based energy expenditure prediction equations were developed based on linear mixed models. The EE prediction equations developed for all accelerometer metrics from hip, dominant wrist, and non-dominant wrist placement, their R^2 and accuracy indices are detailed in Table 1. The R^2 values for the EE models ranged from 0.76 to 0.89, showing that the models could explain at least 76% of the EE variance, except for the dominant wrist when using AC which could only explain 47% of the EE variance.

Regarding the accuracy indices, MAE ranged from 0.92 to 1.99 kcal·min^{-1}. MAPE showed higher variation with values ranging from approximately 16% to 59%. The MAPE for the dominant wrist using AC negatively stood out with and error of 157%. The variation of RMSE ranged from 1.25 to 2.56 kcal·min^{-1}. Based on the MAPE value, we observed that the equation that presented the best estimate for EE was the one using the ENMO metric with the accelerometer placed at the hip (MAPE = 15.98%). The EE prediction equation based on the accelerometer placed at the non-dominant wrist and using the MAD metric presented the lowest value of MAE (0.89 kcal·min^{-1}) and RMSE (1.25 kcal·min^{-1}).

Bland-Altman plots are shown in Figure 1 for all accelerometer metrics and placements. None of the prediction models showed a bias significantly different than zero (all $p > 0.05$), which means that there was no consistent under or overestimation of the EE values by the prediction models developed. Regarding the dispersion of the values, in all analysis, the majority of the data points lied within the plot's limits of agreement. However, all prediction models presented a proportional bias, showing that the values magnitude influenced the prediction error. Nevertheless, by assessing the R^2 of the linear regressions developed to measure the proportional bias, this effect appeared to be very small on most models. The maximum R^2 found was 0.06, apart from the equation that predicted EE based on AC resulting from the accelerometer placed at the dominant wrist, which yielded an R^2 of 0.21.

Table 1. Regression equations, R^2 and accuracy indices.

Acc		Regression Equations for Energy Expenditure (kcal.min^{-1})	R^2	MAE	MAPE	RMSE
Placement	Metrics					
Hip	AC	EE = 0.01348355·(AC) - 2.097862 x 10^{6}·(AC^{2}) + 0.0001885934·(age) - 0.8150812·(sex) + 0.090859·(BMI) - 1.472863	0.84	1.09	38.42%	1.46
	MAD	EE = 0.01769149·(MAD) - 7.621356 x 10^{6}·(MAD^{2}) + 0.005449842·(age) - 0.834207·(sex) + 0.12028·(BMI) -1.844528	0.82	1.21	41.37%	1.51
	ENMO	EE = 0.03276712·(ENMO) - 2.683138 x 10^{5}·($ENMO^{2}$) + 0.004803521·(age) - 0.967626·(sex) + 0.1246731·(BMI) - 2.471343	0.80	1.29	**15.98%**	1.61
Dominant Wrist	AC	EE = 0.005442986·(AC) -3.927977 x 10^{7}·(AC^{2}) - 0.005583778·(age) - 0.9528074·(sex) + 0.1168803·(BMI) – 1.414901	0.47	1.99	157.20%	2.56
	MAD	EE = 0.02598534·(MAD) -1.529447 x 10^{5}·(MAD^{2}) - 0.006868966·(age) - 0.6853687·(sex) + 0.1426291·(BMI) - 3.602786	0.87	0.99	42.84%	1.31
	ENMO	EE = 0.02972061·(ENMO) -2.049393 x 10^{5}·($ENMO^{2}$) + 0.01004525·(age) - 0.8735255·(sex) + 0.1397076·(BMI) - 3.222438	0.81	1.16	58.68%	1.58
Non-dominant Wrist	AC	EE = 0.009100688·(AC) – 1.870809 x 10^{6}·(AC^{2}) + 0.006571667·(age) - 0.9831532·(sex) + 0.1225404·(BMI) – 2.918789	0.76	1.29	53.69%	1.76
	MAD	EE = 0.02226611·(MAD) -1.175238 x 10^{5}·(MAD^{2}) - 0.0046701·(age) - 0.6884789·(sex) + 0.1504716·(BMI) - 3.462773	**0.89**	0.92	30.20%	**1.25**
	ENMO	EE = 0.02785467(ENMO) -1.867554 x 10^{5}($ENMO^{2}$) + 0.01147104·(age) - 0.9438452·(sex) + 0.1312051·(BMI) - 3.135008	0.85	1.07	37.40%	1.45

Abbreviations: AC, activity counts; BMI, body mass index; ENMO, euclidean norm minus one; kcal, kilocalorie; MAD, mean amplitude deviation; MAE, mean absolute error; MAPE, mean absolute percent error; R^2, coefficient of determination; RMSE, root mean square error

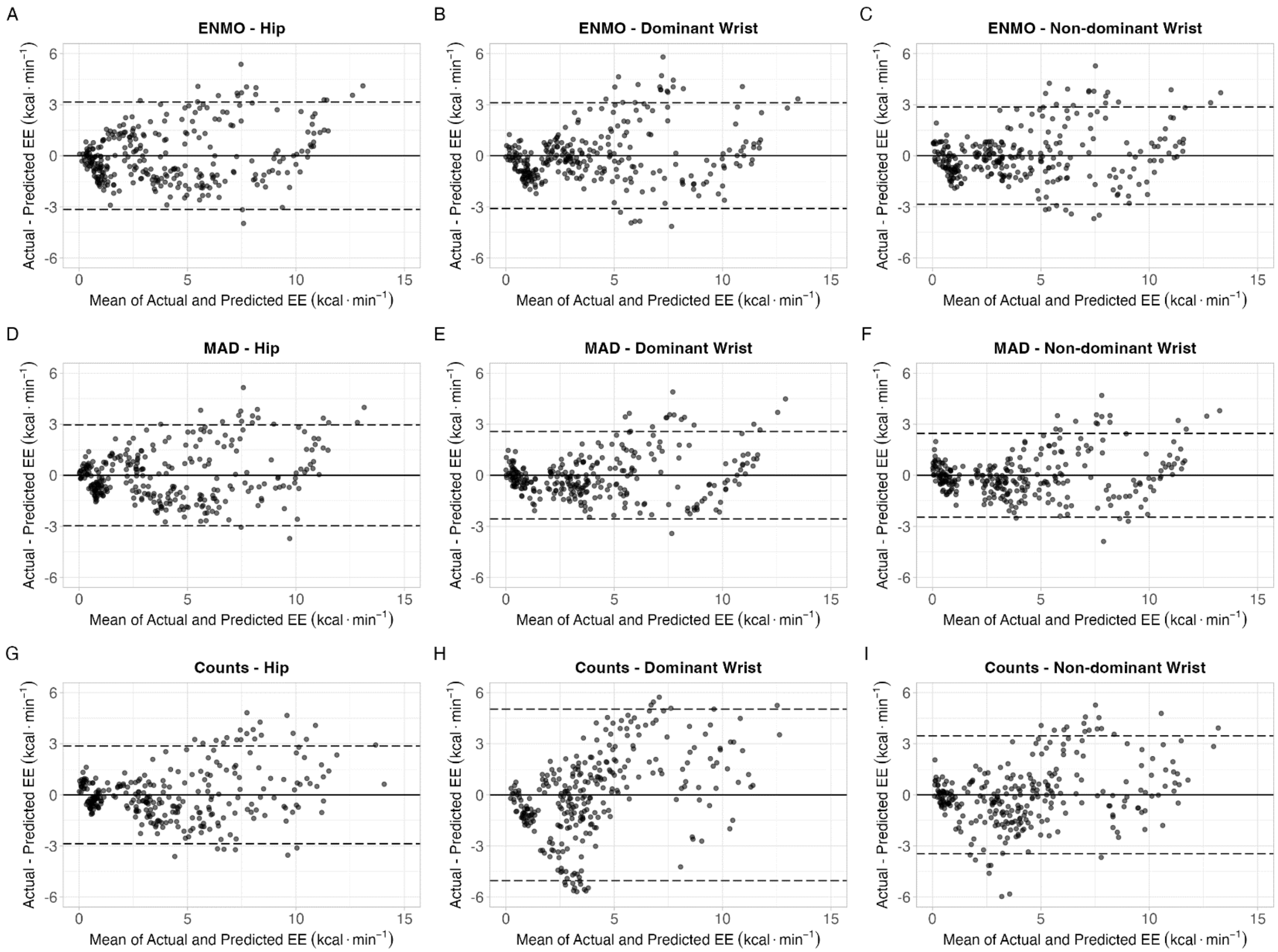

Figure 1. Bland-Altman plots of measured and predicted EE in all accelerometer metrics derived from the hip, dominant wrist and non-dominant wrist placement, with the different metrics explored – ENMO, MAD and AC. Subplots in each row represent the same accelerometer metric (panels A, B and C: ENMO; panels D, E and F: MAF; panels G, H and I: counts) while each column depicts a different placement (panels A, D and G: hip; panels B, E and H: dominant wrist; panels C, F and I: non-dominant wrist) .

Table 2 presents the cut-points developed for all accelerometer metrics from the hip, dominant wrist, and non-dominant wrist placement with their respective SE, SP, and area under the ROC curve. In the analyzes for the accelerometer placed at the hip and for the sedentary classification, SE values were shown to range from 0.92 to 0.98, SP from 0.80 to 0.91 and AUC from 0.91 to 0.96. For moderate intensity, the values were somewhat lower with SE ranging from 0.78 to 0.90, SP between 0.52 to 0.56, and AUC from 0.64 to 0.66. In the vigorous intensity, SE ranged from 0.91 to 0.96, SP was 0.78 and AUC was 0.94 for all metrics at the hip. With the accelerometer placement at the dominant wrist, we observed values from the sedentary classification had a SE ranging between 0.67 to 0.95, SP between 0.8 and 0.91, and AUC with a higher variation ranging between 0.36 and 0.76. For the moderate intensity classification, SE varied between 0.74 and 0.80, SP from 0.47 to 0.58, and AUC from 0.12 to 0.34. For the vigorous intensity classification, SE ranged from 0.71 to 0.94, SP was roughly 0.86, and AUC ranged between 0.52 to 0.66. Results for the accelerometer placed at the non-dominant wrist show that values, for the sedentary classification, ranged from 0.85 and 1 for SE, were roughly 0.89 for SP, and ranged between 0.65 to 0.77 for AUC. For the moderate intensity classification, SE varied between 0.83 and 0.88, SP between 0.53 and 0.55, and AUC between 0.31 and 0.45. For the vigorous intensity classification, SE ranged from 0.78 to 0.94, SP from 0.84 to 0.91, and AUC from 0.59 to 0.65. Individual classification agreement analyses have shown that SA classifications were mostly categorized as having a substantial agreement, MPA as slight or fair, and VPA as substantial or moderate. Percent agreement varied from 41.1% to 64%, with the higher result belonging to the non-dominant wrist.

Table 3 presents results of the mean and standard deviation for VO_2, MET and acceleration metrics for the different tasks included in the protocol. Tasks were also classified as sedentary, light, moderate or vigorous intensity according to the MET captured by indirect calorimetry. This table enables to briefly analyze the extent to which the results obtained through acceleration data are in accordance with the directly measured METs obtained through indirect calorimetry.

Table 2. Proposed cut-points and their classification agreement.

		Cut-points	ROC			Kappa		
			SE	SP	AUC	Individual agreement	Global agreement	Percent agreement
Hip	**AC**						0.81	50.8%
	Sedentary	10	0.98	0.90	0.96	0.75		
	Moderate	215	0.78	0.56	0.66	0.00		
	Vigorous	373	0.91	0.78	0.94	0.49		
	MAD						0.84	58.6%
	Sedentary	6	0.95	0.91	0.96	0.74		
	Moderate	64	0.90	0.52	0.66	0.22		
	Vigorous	245	0.96	0.78	0.94	0.57		
	ENMO						0.79	50.8%
	Sedentary	45	0.92	0.80	0.91	0.56		
	Moderate	58	0.82	0.54	0.64	0.15		
	Vigorous	147	0.94	0.78	0.94	0.54		
Dominant Wrist	**AC**						0.42	41.1%
	Sedentary	393	0.67	0.80	0.75	0.36		
	Moderate	633	0.74	0.47	0.55	0.12		
	Vigorous	774	0.71	0.87	0.87	0.52		
	MAD						0.86	64.0%
	Sedentary	77	0.95	0.91	0.96	0.75		
	Moderate	168	0.80	0.57	0.64	0.34		
	Vigorous	247	0.94	0.86	0.97	0.66		
	ENMO						0.83	58.2%
	Sedentary	54	0.95	0.91	0.95	0.76		
	Moderate	104	0.77	0.58	0.66	0.21		
	Vigorous	193	0.84	0.86	0.95	0.61		
Non-dominant Wrist	**AC**						0.79	64.0%
	Sedentary	282	0.85	0.88	0.91	0.65		
	Moderate	419	0.84	0.53	0.62	0.45		
	Vigorous	687	0.78	0.91	0.93	0.64		
	MAD						0.86	63.3%
	Sedentary	82	1.00	0.89	0.95	0.77		
	Moderate	160	0.83	0.55	0.63	0.31		
	Vigorous	261	0.94	0.84	0.97	0.65		
	ENMO						0.85	62.8%
	Sedentary	47	0.98	0.89	0.93	0.75		
	Moderate	105	0.88	0.54	0.65	0.36		
	Vigorous	215	0.85	0.85	0.95	0.59		

Abbreviations: AC, activity counts; AUC, area under the curve; ENMO, euclidean norm minus one; MAD, mean amplitude deviation; ROC, receiver operating characteristic curve; SE, sensitivity; SP, specificity.

Table 3. Mean ± SD of VO_2, METs, PAI classification and accelerometer output for each activity.

Activity		**Sitting typing**	**Standing**	**House chores**	**Slow walking** (4 km·h^{-1})	**Fast walking** (6 km·h^{-1})	**Running** (9 km·h^{-1})	**Stairs climbing**
VO_2 (mL·min^{-1}·kg^{-1})		5.61 ± 0.84	6.13 ± 1.19	13.20 ± 2.26	13.29 ± 1.42	19.77 ± 2.03	34.53 ± 4.13	27.64 ± 4.84
MET		1.30 ± 0.22	1.43 ± 0.29	3.07 ± 0.58	3.10 ± 0.51	4.62 ± 0.75	8.03 ± 1.06	6.49 ± 1.21
PAI Classification		Sedentary	Sedentary	Moderate	Moderate	Moderate	Vigorous	Vigorous
Hip	AC	1.31 ± 1.92	3.30 ± 3.52	156.12 ± 36.60	266.69 ± 44.69	466.66 ± 71.85	766.81 ± 123.95	396.18 ± 43.17
	MAD	6.13 ± 1.19	4.52 ± 1.24	58.54 ± 15.04	197.86 ± 23.84	379.16 ± 46.58	760.77 ± 106.57	285.64 ± 64.51
	ENMO	35.54 ± 13.63	25.74 ± 8.19	44.12 ± 10.16	124.36 ± 14.43	234.92 ± 36.66	513.11 ± 70.23	168.90 ± 42.67
Dominant Wrist	AC	25.74 ± 8.19	753.32 ± 292.85	645.64 ± 101.05	312.88 ± 58.41	586.59 116.92	1635.23 ± 358.55	570.59 ± 141.61
	MAD	31.54 ± 8.59	59.47 ± 22.14	158.81 ± 30.56	172.94 ± 22.08	272.68 ± 61.72	850.18 ± 136.58	291.18 ± 84.60
	ENMO	40.55 ± 8.90	40.71 ± 17.89	96.73 ± 23.76	108.07 ± 19.35	244.71 ± 66.38	699.81 ± 182.32	189.90 ± 82.51
Non-dominant Wrist	AC	28.54 ± 18.37	217.76 ± 219.62	554.09 ± 110.15	342.76 ± 63.22	642.17 ± 121.80	1656.50 ± 290.95	604.92 ± 137.85
	MAD	18.29 ± 6.49	29.45 ± 16.79	150.20 ± 29.12	180.71 ± 28.09	291.27 ± 61.31	863.15 ± 143.14	311.78 ± 79.78
	AC	30.35 ± 7.34	22.33 ± 13.57	96.68 ± 19.74	123.14 ± 18.33	259.63 ± 61.31	705.85 ± 149.89	216.33 ± 78.29

SA	LPA	MPA	VPA

5. Discussion

Discussion

The aim of this study was to develop regression equations to predict EE and to define cut-points to classify SA and PAI in healthy individuals based on several metrics obtained from hip and both wrists accelerometer placement data. Our results showed that the regression equations and cut-points developed for the diverse placements were different and that the accuracy varied substantially according to the accelerometer metric used.

In regression equations, the use of AC showed the worst results, regardless of the accelerometer placement chosen. The position of the accelerometer at the right hip and the use of ENMO presented propitious results for MAPE (15.98%), even thought RMSE (1.61 kcal·min^{-1}) was somewhat higher than the limit considered as accurate (< 1.30 kcal·min^{-1}) (Lyden et al., 2011).

Concerning cut-points, the results showed that the hip placement of the accelerometer and the use of MAD resulted in better accuracy for both global and percent agreement. The same results were obtained for the dominant wrist. However, the non-dominant wrist showed almost perfect results for global agreement. Despite not having the highest percent agreement (63.3%), it was close to AC results (64%), with a better global agreement. Hip placement had similar results for all metrics, with a positive highlight for MAD. Dominant wrist cut-points results showed that the use of AC were associated with a worst performance. Observing cut-points resulting from the non-dominant wrist data, metric accuracy was similar for global and percent agreement. In general, cut-points presented better results for the non-dominant wrist.

Regarding PAI, moderate activities showed the worst results, no matter which accelerometer metric or placement was selected, being therefore necessary to improve this aspect. Our outcomes suggest that raw acceleration metrics provide notable calibration results that can be adaptable to diverse accelerometer range types, hence they should be implemented instead of AC since this metric depends on the accelerometer brand and or model while metrics based on raw data are, theoretically, the same for all equipment’s. Nevertheless, there is no consensus about what metric should be used as reference (Mendes et al., 2018). So, our study explored two raw acceleration metrics that are more frequently used in literature, so that future investigations could more easily compare our findings with the metric considered most adequate.

Analyzing table 3, it is possible to conclude that sedentary tasks are better predicted with accelerometers placed at the non-dominant wrist, vigorous tasks with accelerometers placed at the hip while moderate tasks are highly dependent of movement pattern. Positioning the accelerometer on the right hip seems to be better as it is more central, but in long-term studies, it would be interesting to consider the non-dominant wrist as it has been shown to increase adherence to the study protocol (Liu et al., 2021). Including more moderate tasks could also be a noteworthy approach to improve both the regression equations and the cut-points accuracy for different intensities.

In the earliest research done in this field, Melanson and Freedson (1995) found that AC were significantly correlated with EE (R^2 = 0.66 – 0.81) regardless of the location of the accelerometer (wrist, hip, ankle). A few years later, Freedson et al. (1998) conducted a study including only locomotion activities performed on the treadmill and the results obtained for the prediction equations were significant (R^2 = 0.82 and SEE = 1.40 $kcal \cdot min^{-1}$). These findings were similar to those obtained in our study, with however an important distinction regarding the fact that, in our study, daily tasks used had unpredictable movement patterns, which can lead to estimation error increase.

One of the first studies that included daily tasks in the development of equations for estimation of EE was carried out more than 20 years ago by Swartz et. al. (2000). These studies had relatively low-grade results for both the wrist (R^2 = 0.033, SEE = 1.38 METs) and the hip (R^2 = 0.32, SEE = 1.16 METs). Later, the research by Hildebrand et al. (2014) used raw acceleration and applied the ENMO metric. The R^2 for monitors placed at the wrist was 0.75 and at the hip 0.81, with R^2 being lower for the wrist compared with the one obtained from the hip placement. However, they do not mention the value of MAE, MAPE, RMSE or any other metric for evaluating the forecast error, making it impossible to determine whether the results could be more accurate or not. The same applies to the cut-points, as this study only mentioned cut-points for 3 and 6 METs but did not disclose any accuracy index to validate the mentioned thresholds. In 2017, Neil-Sztramko et al. (2017) did a more in-depth research showing a positive relationship between AC and MET ($R^2 = 0.47$), even though the magnitude of the effect was not considered very strong. Regarding the cut-points, the discrimination for SA was almost perfect (SE = 1, SP = 0.92, AUC = 0.96) using a threshold of 1514 cpm. Discrimination of MPA (SE = 0.95, SP = 0.84, AUC =0.95) and VPA (SE = 0.89, SP = 0.56, AUC = 0.79) were acceptable using thresholds of 2199 and 4712 cpm, respectively. In this case, the results for SA were similar to our findings, of MPA accuracy were superior to ours, but for VPA the results displayed a worse accuracy compared to ours. However, this study did not include daily tasks which, as was already mentioned previously, is a critical aspect in cut-points given the moderate intensity of these activities. . In turn, Lee et al. (2019), as in our study, also used a quadratic term in the prediction model and the results for both wrists were similar ($R^2 \approx 0.80$). Our models also follow this value. However, the threshold of the cut-points were slightly different because the protocol implemented by these authors only included tasks performed on a treadmill.

Studies with prediction models developed through machine learning also present very positive accuracy results. One of such examples is the study by Ellois et al. (2014) which obtained for the hip and wrist accelerometers average accuracies of 70.2% and 80.2% respectively. Or the study by Montoye et al. (2015) with correlations above 0.80 for EE prediction. The biggest problem with these studies is the lack of sharing and the elaboration of a closed work that does not allow replicating the study to validate the data.

Considering the available evidence in the literature, the results obtained from our study, for both the prediction equations and the development of cutoff points, present very promising results. The results obtained allowed us to understand why the dominant wrist is not frequently used in the literature as an ideal placement for accelerometers since it leads to a lower accuracy, especially in more sedentary tasks. This could happen because, for example, in tasks such as sitting at the computer writing or standing at the counter washing dishes, there is substantial movement of the dominant upper limbs, even though the rest of the body is mostly still and, therefore, global EE is very low. For sedentary tasks, the non-dominant wrist seems to have a better accuracy, since it better reflects the low movement of the rest of the body, consequently reflecting more accurately global EE. This limitation could possibly be generalized to other activities that were not included in the protocol, such as walking with hands in pockets or carrying heavy objects at a slow speed, or activities of daily living such as doing laundry or loading a washing machine. Nevertheless, the prediction errors obtained with these activities could decrease if more intensities were assessed. Moderate intensity activities in turn, were shown to have the highest prediction errors, which may have happened due to the specificity of the selected activities in our protocol and not necessarily due to the calibration method. In locomotion activities such as walking at different speeds, running, and even climbing stairs, the hip accelerometer placement seems to better predict the intensity of the tasks. This may occur due to the accelerometer being placed closer to the center of mass, which better represents global body movement.

This study is not free from limitations, starting with tasks being carried out in the laboratory, a controlled environment, which does not always reflect the type, intensity or pattern of the activities that occur in a natural environment. The criterion method used was based on indirect calorimetry which, although being a robust criterion method, does not assess the anaerobic energy expenditure component of the activity, therefore introducing bias to the obtained results. Also, the assessment of resting metabolic rate could have induced stress in some participants due to the need to wear a face mask for a long period of time, which could have thereby affected the accuracy of the measurement. In activities carried out on the treadmill such as walking and running, a constant pace for all participants was imposed. However, in open activities such as standing while transferring water, sweeping, arranging books and climbing stairs, no instructions were given regarding rhythm and participants were able to choose their own pace. This could have increased heterogeneity of the intensities at which these tasks were performed. Several of the study participants were females that were most likely in different phases of their menstrual cycle, which also influences body temperature and metabolic rate. The phase of the menstrual cycle in which the participants were in, was however not assessed before the protocol. When elaborating the protocol, house chores (task 3) was assumed to be a task usually performed at a light intensity. However, we found that the measured value was 3 METs, which is classified as a moderate activity. However, if we compare it with the 2011 Compendium of PA (Ainsworth et al., 2011) values for the multiple household tasks, the reference value is 2.3 METs, which would be classified as a light intensity. This may have happened due to the use of a face mask, which may generate a certain amount of stress on the individual. The sample used in our study was young, healthy and with a normal BMI. Therefore, these results may not be generalized to other ages, weights, or health conditions. There is therefore a need to carry out similar studies with larger samples, where several heterogeneous conditions can be included.

Comparing the placement of accelerometers at different locations was an asset of our study, especially comparing both wrists, as it allowed us to appreciate that a prediction made by an accelerometer carried out on the dominant or non-dominant side could be quite different. There are several studies in the literature for the prediction of EE, as well as for the classification of cut-points, however they focus mainly on locomotor activities conducted on a treadmill. In this study we included several daily tasks such as typing on the computer, carrying groceries, or even organizing books, so that more accurate equations could be conceived for predicting EE of daily tasks. The same applies to cut-points, as the inclusion of household tasks varies the intensity from other studies, as there is more variety in movement patterns. The use of indirect calorimetry using a portable equipment is a great advantage, as it allows participants to not be limited and to be as mobile as they would be in daily conditions. Using raw acceleration data to apply different metrics opens new possibilities, as it allows comparison between different brands and devices.

6. Conclusions and future perspectives

Conclusions and Future Perspectives

It is well documented that PA has an important impact on physical and mental health, cognitive function, well-being and overall quality of life. PA encompasses a wide spectrum of bodily movements, from structured exercise routines to everyday tasks. The intensity of PA is a crucial variable for its health benefits, however it is challenging to quantify daily PA with sufficient robustness, since activities throughout the day have many variations in their intensity, duration and pattern. The use of accelerometers has become increasingly popular to measure and estimate PAI, both in a research context and by the general population through the use of wearable devices. Accelerometers have demonstrated to be objective and reliable tools for monitoring daily life PA, delivering unbiassed data concerning the intensity, frequency, and duration of motor behaviors. Nevertheless, the accelerometers output needs to be translated into biologically meaningful information through a calibration process. Nowadays, the main uses of accelerometers are related to the estimation and determination of metabolic parameters, such as EE and PAI levels. The calibration of accelerometers has changed over time as a result of methodological and technological advancements, implementing new research designs and methodologies that have improved prediction validity and accuracy. This dissertation addressed the problem of accelerometer calibration to predict EE and the classification of PAI in the general population while performing daily tasks. Our results showed that hip and non-dominant wrist accelerometer placement data allowed and accurate prediction of EE and to correctly classify SA and PAI. Although AC, MAD and ENMO metrics could be used to obtain accurate predictions, ENMO, with the accelerometer placed at the right hip, was shown to be superior with a MAPE of just 15.98%, which was more than twice as accurate when compared to other acceleration metrics used. The findings of this study will assist researchers in being more careful and accurate when choosing the location for the accelerometer as well as will assist in the selection of the metric that could provide the best accuracy. Since this study was conducted under controlled laboratory conditions, future studies in this area should focus on extending the prediction of cardiometabolic variables to daily activities performed in a real outdoor context. Additionally, other free-living activities could be included in the calibration process, which would increase its external validity compared to a laboratory-based calibration.

7. References

References

Ainsworth, B. E., Haskell, W. L., Herrmann, S. D., Meckes, N., Bassett, D. R., Jr., Tudor-Locke, C., Greer, J. L., Vezina, J., Whitt-Glover, M. C., & Leon, A. S. (2011). 2011 Compendium of Physical Activities: A Second Update of Codes and MET values. *Med Sci Sports Exerc*, *43*(8), 1575-1581. https://doi.org/10.1249/MSS.0b013e31821ece12

Alves, A. J., Viana, J. L., Cavalcante, S. L., Oliveira, N. L., Duarte, J. A., Mota, J., Oliveira, J., & Ribeiro, F. (2016). Physical activity in primary and secondary prevention of cardiovascular disease: Overview updated. *World J Cardiol*, *8*(10), 575-583. https://doi.org/10.4330/wjc.v8.i10.575

Bassett, D. R. (2000). Validity of four motion sensors in measuring moderate intensity physical activity. *Medicine & Science in Sports & Exercise*, *32*(9), S471-S480. https://journals.lww.com/acsm-msse/fulltext/2000/09001/validity_of_four_motion_sensors_in_measuring.6.aspx

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

Bassett, D. R., Jr., Wyatt, H. R., Thompson, H., Peters, J. C., & Hill, J. O. (2010). Pedometer-measured physical activity and health behaviors in U.S. adults. *Med Sci Sports Exerc*, *42*(10), 1819-1825. https://doi.org/10.1249/MSS.0b013e3181dc2e54

Bielemann, R. M., Martinez-Mesa, J., & Gigante, D. P. (2013). Physical activity during life course and bone mass: a systematic review of methods and findings from cohort studies with young adults. *BMC Musculoskelet Disord*, *14*, 77. https://doi.org/10.1186/1471-2474-14-77

Blond, K., Brinkløv, C. F., Ried-Larsen, M., Crippa, A., & Grøntved, A. (2020). Association of high amounts of physical activity with mortality risk: a systematic review and meta-analysis. *Br J Sports Med*, *54*(20), 1195-1201. https://doi.org/10.1136/bjsports-2018-100393

Borer, K. T. (2005). Physical activity in the prevention and amelioration of osteoporosis in women: interaction of mechanical, hormonal and dietary factors. *Sports Med*, *35*(9), 779-830. https://doi.org/10.2165/00007256-200535090-00004

Bourke, A. K., O'Brien, J. V., & Lyons, G. M. (2007). Evaluation of a threshold-based tri-axial accelerometer fall detection algorithm. *Gait Posture*, *26*(2), 194-199. https://doi.org/10.1016/j.gaitpost.2006.09.012

Caspersen, C. J., Powell, K. E., & Christenson, G. M. (1985). Physical activity, exercise, and physical fitness: definitions and distinctions for health-related research. *Public Health Rep*, *100*(2), 126-131.

[Record #136 is using a reference type undefined in this output style.]

Chastin, S. F., Mandrichenko, O., Helbostadt, J. L., & Skelton, D. A. (2014). Associations between objectively-measured sedentary behavior and physical activity with bone mineral density in adults and older adults, the NHANES study. *Bone*, *64*, 254-262. https://doi.org/10.1016/j.bone.2014.04.009

Chen, K. Y., & Bassett, D. R., Jr. (2005). The technology of accelerometry-based activity monitors: current and future. *Med Sci Sports Exerc*, *37*(11 Suppl), S490-500. https://doi.org/10.1249/01.mss.0000185571.49104.82

Crouter, S. E., Clowers, K. G., & Bassett, D. R., Jr. (2006). A novel method for using accelerometer data to predict energy expenditure. *J Appl Physiol (1985)*, *100*(4), 1324-1331. https://doi.org/10.1152/japplphysiol.00818.2005

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

Crouter, S. E., Oody, J. F., & Bassett, D. R., Jr. (2018). Estimating physical activity in youth using an ankle accelerometer. *J Sports Sci*, *36*(19), 2265-2271. https://doi.org/10.1080/02640414.2018.1449091

Crowley, P., Skotte, J., Stamatakis, E., Hamer, M., Aadahl, M., Stevens, M. L., Rangul, V., Mork, P. J., & Holtermann, A. (2019). Comparison of physical behavior estimates from three different thigh-worn accelerometers brands: a proof-of-concept for the Prospective Physical Activity, Sitting, and Sleep consortium (ProPASS). *Int J Behav Nutr Phys Act*, *16*(1), 65. https://doi.org/10.1186/s12966-019-0835-0

Diniz-Sousa, F., Veras, L., Ribeiro, J. C., Boppre, G., Devezas, V., Santos-Sousa, H., Preto, J., Machado, L., Vilas-Boas, J. P., Oliveira, J., & Fonseca, H. (2020). Accelerometry calibration in people with class II-III obesity: Energy expenditure prediction and physical activity intensity identification. *Gait Posture*, *76*, 104-109. https://doi.org/10.1016/j.gaitpost.2019.11.008

Duncan, M. J., Rowlands, A., Lawson, C., Leddington Wright, S., Hill, M., Morris, M., Eyre, E., & Tallis, J. (2020). Using accelerometry to classify physical activity intensity in older adults: What is the optimal wear-site? *Eur J Sport Sci*, *20*(8), 1131-1139. https://doi.org/10.1080/17461391.2019.1694078

Ekelund, U., Tarp, J., Steene-Johannessen, J., Hansen, B. H., Jefferis, B., Fagerland, M. W., Whincup, P., Diaz, K. M., Hooker, S. P., Chernofsky, A., Larson, M. G., Spartano, N., Vasan, R. S., Dohrn, I. M., Hagströmer, M., Edwardson, C., Yates, T., Shiroma, E., Anderssen, S. A., & Lee, I. M. (2019). Dose-response associations between accelerometry measured physical activity and sedentary time and all cause mortality: systematic review and harmonised meta-analysis. *Bmj*, *366*, l4570. https://doi.org/10.1136/bmj.l4570

Ellingson, L. D., Hibbing, P. R., Kim, Y., Frey-Law, L. A., Saint-Maurice, P. F., & Welk, G. J. (2017). Lab-based validation of different data processing methods for wrist-worn ActiGraph accelerometers in young adults. *Physiol Meas*, *38*(6), 1045-1060. https://doi.org/10.1088/1361-6579/aa6d00

Ellis, K., Kerr, J., Godbole, S., Lanckriet, G., Wing, D., & Marshall, S. (2014). A random forest classifier for the prediction of energy expenditure and type of physical activity from wrist and hip accelerometers. *Physiol Meas*, *35*(11), 2191-2203. https://doi.org/10.1088/0967-3334/35/11/2191

Evenson, K. R., Scherer, E., Peter, K. M., Cuthbertson, C. C., & Eckman, S. (2022). Historical development of accelerometry measures and methods for physical activity and sedentary behavior research worldwide: A scoping review of observational studies of adults. *PLoS One*, *17*(11), e0276890. https://doi.org/10.1371/journal.pone.0276890

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

Fortune, E., Morrow, M. M., & Kaufman, K. R. (2014). Assessment of gait kinetics using triaxial accelerometers. *J Appl Biomech*, *30*(5), 668-674. https://doi.org/10.1123/jab.2014-0037

Freedson, P. S., Melanson, E., & Sirard, J. (1998). Calibration of the Computer Science and Applications, Inc. accelerometer. *Med Sci Sports Exerc*, *30*(5), 777-781. https://doi.org/10.1097/00005768-199805000-00021

Friedenreich, C. M., Neilson, H. K., & Lynch, B. M. (2010). State of the epidemiological evidence on physical activity and cancer prevention. *Eur J Cancer*, *46*(14), 2593-2604. https://doi.org/10.1016/j.ejca.2010.07.028

Füzéki, E., Engeroff, T., & Banzer, W. (2017). Health Benefits of Light-Intensity Physical Activity: A Systematic Review of Accelerometer Data of the National Health and Nutrition Examination Survey (NHANES). *Sports Med*, *47*(9), 1769-1793. https://doi.org/10.1007/s40279-017-0724-0

Gao, Z., Liu, W., McDonough, D. J., Zeng, N., & Lee, J. E. (2021). The Dilemma of Analyzing Physical Activity and Sedentary Behavior with Wrist Accelerometer Data: Challenges and Opportunities. *J Clin Med*, *10*(24). https://doi.org/10.3390/jcm10245951

Guthold, R., Stevens, G. A., Riley, L. M., & Bull, F. C. (2020). Global trends in insufficient physical activity among adolescents: a pooled analysis of 298 population-based surveys with 1·6 million participants. *Lancet Child Adolesc Health*, *4*(1), 23-35. https://doi.org/10.1016/s2352-4642(19)30323-2

Hamasaki, H. (2016). Daily physical activity and type 2 diabetes: A review. *World J Diabetes*, *7*(12), 243-251. https://doi.org/10.4239/wjd.v7.i12.243

Haskell, W. L., Lee, I. M., Pate, R. R., Powell, K. E., Blair, S. N., Franklin, B. A., Macera, C. A., Heath, G. W., Thompson, P. D., & Bauman, A. (2007). Physical activity and public health: updated recommendation for adults from the American College of Sports Medicine and the American Heart Association. *Med Sci Sports Exerc*, *39*(8), 1423-1434. https://doi.org/10.1249/mss.0b013e3180616b27

Health, U. D. o., & Services, H. (2018). 2018 Physical activity guidelines advisory committee scientific report.

Hernández-Vicente, A., Marín-Puyalto, J., Pueyo, E., Vicente-Rodríguez, G., & Garatachea, N. (2022). Physical Activity in Centenarians beyond Cut-Point-Based Accelerometer Metrics. *International Journal of Environmental Research and Public Health*, *19*(18), 11384. https://www.mdpi.com/1660-4601/19/18/11384

Hildebrand, M., VT, V. A. N. H., Hansen, B. H., & Ekelund, U. (2014). Age group comparability of raw accelerometer output from wrist- and hip-worn monitors. *Med Sci Sports Exerc*, *46*(9), 1816-1824. https://doi.org/10.1249/mss.0000000000000289

Hills, A. P., Mokhtar, N., & Byrne, N. M. (2014). Assessment of physical activity and energy expenditure: an overview of objective measures. *Front Nutr*, *1*, 5. https://doi.org/10.3389/fnut.2014.00005

Ho, C. S., Chang, C. H., Lin, K. C., Huang, C. C., & Hsu, Y. J. (2019). Correction of estimation bias of predictive equations of energy expenditure based on wrist/waist-mounted accelerometers. *PeerJ*, *7*, e7973. https://doi.org/10.7717/peerj.7973

John, D., Sasaki, J., Staudenmayer, J., Mavilia, M., & Freedson, P. S. (2013). Comparison of raw acceleration from the GENEA and ActiGraph™ GT3X+ activity monitors. *Sensors (Basel)*, *13*(11), 14754-14763. https://doi.org/10.3390/s131114754

Kaminsky, L. A., & Montoye, A. H. (2014). Physical activity and health: what is the best dose? *J Am Heart Assoc*, *3*(5), e001430. https://doi.org/10.1161/jaha.114.001430

Kavanagh, J. J., & Menz, H. B. (2008). Accelerometry: a technique for quantifying movement patterns during walking. *Gait Posture*, *28*(1), 1-15. https://doi.org/10.1016/j.gaitpost.2007.10.010

Kirkham, A. A., & Davis, M. K. (2015). Exercise Prevention of Cardiovascular Disease in Breast Cancer Survivors. *J Oncol*, *2015*, 917606. https://doi.org/10.1155/2015/917606

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*(1), 159-174.

Le Masurier, G. C., & Tudor-Locke, C. (2003). Comparison of pedometer and accelerometer accuracy under controlled conditions. *Med Sci Sports Exerc*, *35*(5), 867-871. https://doi.org/10.1249/01.Mss.0000064996.63632.10

Lee, I. M., Shiroma, E. J., Lobelo, F., Puska, P., Blair, S. N., & Katzmarzyk, P. T. (2012). Effect of physical inactivity on major non-communicable diseases worldwide: an analysis of burden of disease and life expectancy. *Lancet*, *380*(9838), 219-229. https://doi.org/10.1016/s0140-6736(12)61031-9

Lee, P., & Tse, C. Y. (2019). Calibration of wrist-worn ActiWatch 2 and ActiGraph wGT3X for assessment of physical activity in young adults. *Gait Posture*, *68*, 141-149. https://doi.org/10.1016/j.gaitpost.2018.11.023

Leone, A., Rescio, G., Caroppo, A., Siciliano, P., & Manni, A. (2023). Human Postures Recognition by Accelerometer Sensor and ML Architecture Integrated in Embedded Platforms: Benchmarking and Performance Evaluation. *Sensors (Basel)*, *23*(2). https://doi.org/10.3390/s23021039

Liu, F., Wanigatunga, A. A., & Schrack, J. A. (2021). Assessment of Physical Activity in Adults Using Wrist Accelerometers. *Epidemiologic Reviews*, *43*(1), 65-93. https://doi.org/10.1093/epirev/mxab004

Matthew, C. E. (2005). Calibration of accelerometer output for adults. *Med Sci Sports Exerc*, *37*(11 Suppl), S512-522. https://doi.org/10.1249/01.mss.0000185659.11982.3d

Matthews, C. E., Jurj, A. L., Shu, X. O., Li, H. L., Yang, G., Li, Q., Gao, Y. T., & Zheng, W. (2007). Influence of exercise, walking, cycling, and overall nonexercise physical activity on mortality in Chinese women. *Am J Epidemiol*, *165*(12), 1343-1350. https://doi.org/10.1093/aje/kwm088

Matthews, C. E., Kozey Keadle, S., Moore, S. C., Schoeller, D. S., Carroll, R. J., Troiano, R. P., & Sampson, J. N. (2018). Measurement of Active and Sedentary Behavior in Context of Large Epidemiologic Studies. *Med Sci Sports Exerc*, *50*(2), 266-276. https://doi.org/10.1249/mss.0000000000001428

Melanson, E. L., Jr., & Freedson, P. S. (1995). Validity of the Computer Science and Applications, Inc. (CSA) activity monitor. *Med Sci Sports Exerc*, *27*(6), 934-940.

Menai, M., van Hees, V. T., Elbaz, A., Kivimaki, M., Singh-Manoux, A., & Sabia, S. (2017). Accelerometer assessed moderate-to-vigorous physical activity and successful ageing: results from the Whitehall II study. *Sci Rep*, *8*, 45772. https://doi.org/10.1038/srep45772

Mendes, M. d. A., Silva, I. C. M. d., Ramires, V. V., Reichert, F. F., Martins, R. C., & Tomasi, E. (2018). Calibration of raw accelerometer data to measure physical activity: A systematic review. *Gait Posture*, *61*, 98-110. https://doi.org/10.1016/j.gaitpost.2017.12.028

Migueles, J. H., Cadenas-Sanchez, C., Ekelund, U., Delisle Nyström, C., Mora-Gonzalez, J., Löf, M., Labayen, I., Ruiz, J. R., & Ortega, F. B. (2017). Accelerometer Data Collection and Processing Criteria to Assess Physical Activity and Other Outcomes: A Systematic Review and Practical Considerations. *Sports Med*, *47*(9), 1821-1845. https://doi.org/10.1007/s40279-017-0716-0

Migueles, J. H., Delisle Nyström, C., Henriksson, P., Cadenas-Sanchez, C., Ortega, F. B., & Löf, M. (2019). Accelerometer Data Processing and Energy Expenditure Estimation in Preschoolers. *Med Sci Sports Exerc*, *51*(3), 590-598. https://doi.org/10.1249/mss.0000000000001797

Moe-Nilssen, R., & Helbostad, J. L. (2004). Estimation of gait cycle characteristics by trunk accelerometry. *J Biomech*, *37*(1), 121-126. https://doi.org/10.1016/s0021-9290(03)00233-1

Montoye, A. H., Mudd, L. M., Biswas, S., & Pfeiffer, K. A. (2015). Energy Expenditure Prediction Using Raw Accelerometer Data in Simulated Free Living. *Med Sci Sports Exerc*, *47*(8), 1735-1746. https://doi.org/10.1249/mss.0000000000000597

Montoye, A. H. K., Begum, M., Henning, Z., & Pfeiffer, K. A. (2017). Comparison of linear and non-linear models for predicting energy expenditure from raw accelerometer data. *Physiol Meas*, *38*(2), 343-357. https://doi.org/10.1088/1361-6579/38/2/343

Montoye, H. J., Washburn, R., Servais, S., Ertl, A., Webster, J. G., & Nagle, F. J. (1983). Estimation of energy expenditure by a portable accelerometer. *Med Sci Sports Exerc*, *15*(5), 403-407.

Mozaffarian, D., Furberg, C. D., Psaty, B. M., & Siscovick, D. (2008). Physical activity and incidence of atrial fibrillation in older adults: the cardiovascular health study. *Circulation*, *118*(8), 800-807. https://doi.org/10.1161/circulationaha.108.785626

Nakagawa, S., & Schielzeth, H. (2013). A general and simple method for obtaining R2 from generalized linear mixed-effects models. *Methods in Ecology and Evolution*, *4*(2), 133-142. https://doi.org/https://doi.org/10.1111/j.2041-210x.2012.00261.x

Ndahimana, D., & Kim, E.-K. (2017). Measurement Methods for Physical Activity and Energy Expenditure: a Review. *Clin Nutr Res*, *6*(2), 68-80. https://doi.org/10.7762/cnr.2017.6.2.68

Neil-Sztramko, S. E., Rafn, B. S., Gotay, C. C., & Campbell, K. L. (2017). Determining activity count cut-points for measurement of physical activity using the Actiwatch2 accelerometer. *Physiol Behav*, *173*, 95-100. https://doi.org/10.1016/j.physbeh.2017.01.026

Neugebauer, J. M., Collins, K. H., & Hawkins, D. A. (2014). Ground reaction force estimates from ActiGraph GT3X+ hip accelerations. *PLoS One*, *9*(6), e99023. https://doi.org/10.1371/journal.pone.0099023

Pedišić, Ž., & Bauman, A. (2015). Accelerometer-based measures in physical activity surveillance: current practices and issues. *Br J Sports Med*, *49*(4), 219-223. https://doi.org/10.1136/bjsports-2013-093407

Pfeiffer, K. A., Clevenger, K. A., Kaplan, A., Van Camp, C. A., Strath, S. J., & Montoye, A. H. K. (2022). Accessibility and use of novel methods for predicting physical activity and energy expenditure using accelerometry: a scoping review. *Physiol Meas*, *43*(9). https://doi.org/10.1088/1361-6579/ac89ca

Plasqui, G., Bonomi, A. G., & Westerterp, K. R. (2013). Daily physical activity assessment with accelerometers: new insights and validation studies. *Obes Rev*, *14*(6), 451-462. https://doi.org/10.1111/obr.12021

Prince, S. A., Adamo, K. B., Hamel, M. E., Hardt, J., Connor Gorber, S., & Tremblay, M. (2008). A comparison of direct versus self-report measures for assessing physical activity in adults: a systematic review. *Int J Behav Nutr Phys Act*, *5*, 56. https://doi.org/10.1186/1479-5868-5-56

Rebar, A. L., Stanton, R., Geard, D., Short, C., Duncan, M. J., & Vandelanotte, C. (2015). A meta-meta-analysis of the effect of physical activity on depression and anxiety in non-clinical adult populations. *Health Psychol Rev*, *9*(3), 366-378. https://doi.org/10.1080/17437199.2015.1022901

Rhudy, M. B., Dreisbach, S. B., Moran, M. D., Ruggiero, M. J., & Veerabhadrappa, P. (2020). Cut points of the Actigraph GT9X for moderate and vigorous intensity physical activity at four different wear locations. *J Sports Sci*, *38*(5), 503-510. https://doi.org/10.1080/02640414.2019.1707956

Sallis, J. F., Bull, F., Guthold, R., Heath, G. W., Inoue, S., Kelly, P., Oyeyemi, A. L., Perez, L. G., Richards, J., & Hallal, P. C. (2016). Progress in physical activity over the Olympic quadrennium. *Lancet*, *388*(10051), 1325-1336. https://doi.org/10.1016/s0140-6736(16)30581-5

Shiroma, E. J., Sesso, H. D., Moorthy, M. V., Buring, J. E., & Lee, I. M. (2014). Do moderate-intensity and vigorous-intensity physical activities reduce mortality rates to the same extent? *J Am Heart Assoc*, *3*(5), e000802. https://doi.org/10.1161/jaha.114.000802

Sigal, R. J., Kenny, G. P., Wasserman, D. H., Castaneda-Sceppa, C., & White, R. D. (2006). Physical activity/exercise and type 2 diabetes: a consensus statement from the American Diabetes Association. *Diabetes Care*, *29*(6), 1433-1438. https://doi.org/10.2337/dc06-9910

Sirichana, W., Dolezal, B. A., Neufeld, E. V., Wang, X., & Cooper, C. B. (2017). Wrist-worn triaxial accelerometry predicts the energy expenditure of nonvigorous daily physical activities. *J Sci Med Sport*, *20*(8), 761-765. https://doi.org/10.1016/j.jsams.2017.01.233

Song, T. A., Chowdhury, S. R., Malekzadeh, M., Harrison, S., Hoge, T. B., Redline, S., Stone, K. L., Saxena, R., Purcell, S. M., & Dutta, J. (2023). AI-Driven sleep staging from actigraphy and heart rate. *PLoS One*, *18*(5), e0285703. https://doi.org/10.1371/journal.pone.0285703

Staudenmayer, J., He, S., Hickey, A., Sasaki, J., & Freedson, P. (2015). Methods to estimate aspects of physical activity and sedentary behavior from high-frequency wrist accelerometer measurements. *J Appl Physiol (1985)*, *119*(4), 396-403. https://doi.org/10.1152/japplphysiol.00026.2015

Staudenmayer, J., Zhu, W., & Catellier, D. J. (2012). Statistical considerations in the analysis of accelerometry-based activity monitor data. *Med Sci Sports Exerc*, *44*(1 Suppl 1), S61-67. https://doi.org/10.1249/MSS.0b013e3182399e0f

Strath, S. J., Kaminsky, L. A., Ainsworth, B. E., Ekelund, U., Freedson, P. S., Gary, R. A., Richardson, C. R., Smith, D. T., & Swartz, A. M. (2013). Guide to the assessment of physical activity: Clinical and research applications: a scientific statement from the American Heart Association. *Circulation*, *128*(20), 2259-2279. https://doi.org/10.1161/01.cir.0000435708.67487.da

Strath, S. J., Kate, R. J., Keenan, K. G., Welch, W. A., & Swartz, A. M. (2015). Ngram time series model to predict activity type and energy cost from wrist, hip and ankle accelerometers: implications of age. *Physiological Measurement*, *36*(11), 2335. https://doi.org/10.1088/0967-3334/36/11/2335

Strath, S. J., Pfeiffer, K. A., & Whitt-Glover, M. C. (2012). Accelerometer use with children, older adults, and adults with functional limitations [Article]. *Medicine and Science in Sports and Exercise*, *44*(SUPPL. 1), S77-S85. https://doi.org/10.1249/MSS.0b013e3182399eb1

Swartz, A. M., Strath, S. J., Bassett, D. R., Jr., O'Brien, W. L., King, G. A., & Ainsworth, B. E. (2000). Estimation of energy expenditure using CSA accelerometers at hip and wrist sites. *Med Sci Sports Exerc*, *32*(9 Suppl), S450-456. https://doi.org/10.1097/00005768-200009001-00003

Troiano, R. P. (2005). A timely meeting: objective measurement of physical activity. *Med Sci Sports Exerc*, *37*(11 Suppl), S487-489. https://doi.org/10.1249/01.mss.0000185473.32846.c3

Troiano, R. P., McClain, J. J., Brychta, R. J., & Chen, K. Y. (2014). Evolution of accelerometer methods for physical activity research. *Br J Sports Med*, *48*(13), 1019-1023. https://doi.org/10.1136/bjsports-2014-093546

Vähä-Ypyä, H., Vasankari, T., Husu, P., Suni, J., & 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/https://doi.org/10.1111/cpf.12127

van Hees, V. T., Gorzelniak, L., Dean León, E. C., Eder, M., Pias, M., Taherian, S., Ekelund, U., Renström, F., Franks, P. W., Horsch, A., & Brage, S. (2013). Separating Movement and Gravity Components in an Acceleration Signal and Implications for the Assessment of Human Daily Physical

Activity. *PLoS One*, *8*(4), e61691. https://doi.org/10.1371/journal.pone.0061691

van Hees, V. T., Renström, F., Wright, A., Gradmark, A., Catt, M., Chen, K. Y., Löf, M., Bluck, L., Pomeroy, J., Wareham, N. J., Ekelund, U., Brage, S., & Franks, P. W. (2011). Estimation of daily energy expenditure in pregnant and non-pregnant women using a wrist-worn tri-axial accelerometer. *PLoS One*, *6*(7), e22922. https://doi.org/10.1371/journal.pone.0022922

van Hees, V. T., Thaler-Kall, K., Wolf, K. H., Brønd, J. C., Bonomi, A., Schulze, M., Vigl, M., Morseth, B., Hopstock, L. A., Gorzelniak, L., Schulz, H., Brage, S., & Horsch, A. (2016). Challenges and Opportunities for Harmonizing Research Methodology: Raw Accelerometry. *Methods Inf Med*, *55*(6), 525-532. https://doi.org/10.3414/me15-05-0013

Veras, L., Diniz-Sousa, F., Boppre, G., Resende-Coelho, A., Moutinho-Ribeiro, E., Devezas, V., Santos-Sousa, H., Preto, J., Vilas-Boas, J. P., Machado, L., Oliveira, J., & Fonseca, H. (2023). Mechanical loading prediction through accelerometry data during walking and running. *European Journal of Sport Science*, *23*(8), 1518-1527. https://doi.org/10.1080/17461391.2022.2102437

Warburton, D. E., Nicol, C. W., & Bredin, S. S. (2006). Health benefits of physical activity: the evidence. *Cmaj*, *174*(6), 801-809. https://doi.org/10.1503/cmaj.051351

Warburton, D. E. R., & Bredin, S. S. D. (2017). Health benefits of physical activity: a systematic review of current systematic reviews. *Curr Opin Cardiol*, *32*(5), 541-556. https://doi.org/10.1097/hco.0000000000000437

Weir, J. B. d. V. (1949). New methods for calculating metabolic rate with special reference to protein metabolism. *The Journal of Physiology*, *109*(1-2), 1-9. https://doi.org/https://doi.org/10.1113/jphysiol.1949.sp004363

Welk, G. J. (2005). Principles of design and analyses for the calibration of accelerometry-based activity monitors. *Med Sci Sports Exerc*, *37*(11 Suppl), S501-511. https://doi.org/10.1249/01.mss.0000185660.38335.de

Westerterp, K. R. (2017). Doubly labelled water assessment of energy expenditure: principle, practice, and promise. *Eur J Appl Physiol*, *117*(7), 1277-1285. https://doi.org/10.1007/s00421-017-3641-x

Westerterp, K. R., & Plasqui, G. (2004). Physical activity and human energy expenditure. *Curr Opin Clin Nutr Metab Care*, *7*(6), 607-613. https://doi.org/10.1097/00075197-200411000-00004

Wong, T. C., Webster, J. G., Montoye, H. J., & Washburn, R. (1981). Portable accelerometer device for measuring human energy expenditure. *IEEE Trans Biomed Eng*, *28*(6), 467-471. https://doi.org/10.1109/tbme.1981.324820

Wright, S. P., Hall Brown, T. S., Collier, S. R., & Sandberg, K. (2017). How consumer physical activity monitors could transform human physiology research. *Am J Physiol Regul Integr Comp Physiol*, *312*(3), R358-r367. https://doi.org/10.1152/ajpregu.00349.2016

Zhang, S., Murray, P., Zillmer, R., Eston, R. G., Catt, M., & Rowlands, A. V. (2012). Activity classification using the GENEA: optimum sampling frequency and number of axes. *Med Sci Sports Exerc*, *44*(11), 2228-2234. https://doi.org/10.1249/MSS.0b013e31825e19fd