International Doctoral Thesis
Tesis Doctoral Internacional

Jairo Hidalgo Migueles

Accelerometer-determined physical activity and its relationship with health

Methodology and application

Medición de actividad física con acelerómetros y su relación con la salud

Metodología y aplicación

PROGRAMA DE DOCTORADO EN BIOMEDICINA

DEPARTAMENTO DE EDUCACIÓN FÍSICA Y DEPORTIVA
FACULTAD DE CIENCIAS DEL DEPORTE
UNIVERSIDAD DE GRANADA
2020

Editor: Universidad de Granada. Tesis Doctorales
Autor: Jairo Hidalgo Migueles
ISBN: 978-84-1306-734-6
URI: http://hdl.handle.net/10481/65421

Accelerometer-determined physical activity and its relationship with health: methodology and application
Design and main layout: Jairo Hidalgo Migueles, Miguel Rodríguez Barahona
Illustrations: Miguel Rodríguez Barahona
Funding: Ministerio de Educación, Cultura y Deporte

Al sólo sé que no sé nada,
a la bendita curiosidad,
al progreso imparable
y a la timidez desafiante.
A mis padres, a mis hermanas, a mis familias.

To the I only know that I know nothing,
to the holy curiosity,
to the unstoppable progress
and to the defiant shyness.
To dad and mom, to my sisters, to my families.

UNIVERSIDAD DE GRANADA

DEPARTAMENTO DE EDUCACIÓN FÍSICA Y DEPORTIVA
FACULTAD DE CIENCIAS DEL DEPORTE
UNIVERSDAD DE GRANADA

FACULTAD DE CIENCIAS DEL DEPORTE

Accelerometer-determined physical activity and its relationship with health

Methodology and application

Medición de actividad física con acelerómetros y su relación con la salud

Metodología y aplicación

JAIRO HIDALGO MIGUELES

Doctoral Thesis Supervisors [Directores de la Tesis Doctoral]

Francisco B. Ortega Porcel
PhD
Prof. Titular de Universidad
Universidad de Granada
Granada (España)

Vincent T. van Hees
PhD
Fundador e investigador
Accelting
Almere (Países Bajos)

Granada, 18 de diciembre de 2020

El doctorando D. Jairo Hidalgo Migueles ha realizado la presente Tesis Doctoral como beneficiario de una beca-contrato con cargo al programa de Formación de Profesorado Universitario (FPU15/02645) del Ministerio de Educación, Cultura y Deporte, por resolución de 05 de agosto de 2016, de la Secretaría de Estado de Educación, Formación Profesional y Universidades.

Table of Contents

Research projects and funding

ActiveBrains

The present International Doctoral Thesis was carried out primarily under the framework of the **ActiveBrains project** (http://profit.ugr.es/activebrains?lang=en), which was funded by the following organizations:

- Spanish Ministry of Economy and Competitiveness and Fondo Europeo de Desarrollo Regional (FEDER) (DEP-2013-47540, DEP2016-79512-R, DEP2017-91544-EXP, BES-2014-068829, FJCI-2014-19563, IJCI-2017-33642, and RYC-2011-09011).
- Spanish Ministry of Education (FPU 14/06837, FPU 15/02645, and FPU 16/02760).
- University of Granada, Plan Propio de Investigación 2016, Excellence actions: Units of Excellence, Unit of Excellence on Exercise and Health (UCEES). Visiting Scholar.
- Horizon 2020 Framework Programme, Grant/ Award Number: 667302.
- Junta de Andalucía, Conserjería de Conocimiento, Investigación y Universidades and European Regional Development Fund (ERDF) (SOMM17/6107/UGR).
- Redes temáticas de investigación cooperativa RETIC (SAMID III) European Regional Development Fund (ERDF) (RD16/0022).
- EXERNET Research Network on Exercise and Health in Special Populations (DEP2005-00046/ACTI).
- Alicia Koplowitz Foundation.

MINISTOP

Additionally, the **MINISTOP project** (**M**obile-based **IN**tervention **I**ntended to **S**top **O**besity in **P**reschoolers) is also present in this International Doctoral Thesis. The MINISTOP project was funded by the following organizations:

- The Swedish Research Council for Health.
- Working life and Welfare.
- Karolinska Institutet.
- Bo and Vera Axson Johnssons' foundation.

Pilot study on accelerometry

Finally, this International Doctoral Thesis also includes data from a non-funded pilot study on accelerometry conducted in the University of Granada.

ABSTRACT

“Simplify, simplify”

Thoreau

CONTENTS

Abstract

Background | Accelerometers are the method of choice for the measurement of physical behaviours (i.e., physical activity [PA], sedentary behaviour [SB], and sleep) in current research. The rapid technological advances and the access to the accelerometers' raw data addresses a series of challenges upon the need for transparent, comparable, and reproducible accelerometer data processing methods. The widely studied associations of physical behaviours with health outcomes should be expanded to understand the behaviours interplay in their relationship with health. Children with overweight or obesity might find in physical behaviours an effective tool to improve their cardiometabolic and brain health.

Objectives | Two main objectives are addressed in this Thesis: (i) to advance the current knowledge on accelerometer data collection and processing methods to study physical behaviours in children and adults with accelerometers; and (ii) to explore the associations of accelerometer-determined physical behaviours with cardiometabolic and brain health in children with overweight or obesity, as well as the effects of the ActiveBrains exercise randomized controlled trial

Methods | The design of the studies included in this Thesis are a systematic review, a software description article, seven cross-sectional studies, a consensus statement article, and a randomized controlled trial. This Thesis encompasses data mainly from the ActiveBrains project, and complementary from the MINISTOP project and a pilot study on accelerometry. ActiGraph GT3X+ accelerometers attached to the right hip and wrists are used to quantify physical behaviours. Gold-standard measures of energy expenditure (i.e., doubly labelled-water), brain grey matter volume (i.e., magnetic resonance imaging), cardiometabolic health (i.e., blood biomarkers), and body composition (i.e., dual-energy x-ray absorptiometry) are included. Analytical approaches used include linear and quadratic regressions, analysis of variance (ANOVA), compositional data analysis, multivariate pattern analysis, and mediation models.

Main findings | In regards to the objective 1, this Thesis: (i) provides accelerometer data collection and processing recommendations based on existing literature; (ii) describes an open-source software to process raw accelerometer data to quantify physical behaviours in which the PhD candidate is co-developer; (iii) finds that open-source acceleration metrics present a higher performance than proprietary activity counts to estimate energy expenditure; (iv) observes that open-source acceleration metrics are more comparable between them than with activity counts and provides cut-points to quantify PA intensity from dominant wrist-worn accelerometer data; (v) demonstrates large discrepancies in the time spent in SB and PA intensities when quantified from different cut-points, suggesting that it is not currently possible to know the prevalence of a population meeting the PA guidelines based on accelerometer data; (vi) proposes step-based metrics (including steps/day and various cadence-based intensity indicators) as a good proxy to some indicators of overall PA (i.e., counts per day, light-moderate-vigorous PA, moderate-to-vigorous PA) in children with overweight or obesity; (vii) provides a comprehensive description and international consensus on the analytical

approaches most-frequently used in the field, and practical recommendations on what analytical approaches are the best-suited to a given research question.
Relative to the objective 2, the current Thesis: (viii) observes that the association of PA and SB with grey matter volume in the hippocampus in children with overweight or obesity might be moderated by weight status (reallocating 20 min/day from SB to moderate-to-vigorous PA was associated with 100 mm^3 more GMV in the right hippocampus in children with obesity type I); (ix) finds that sleep behaviours are associated with grey matter volume in several cortical and subcortical brain regions independently of SB and PA, and this seemed to be relevant for academic achievement in children with overweight or obesity; (x) remarks that a more stable and less fragmented activity-rest pattern (and earlier occurrence of PA) is associated with better academic achievement, executive function, and intelligence in children with overweight or obesity; (xi) demonstrates that a 20-week exercise program improves cardiometabolic health in children with overweight or obesity, while no effect is observed on mental health.

Conclusion | The findings from this International Doctoral Thesis provide valuable recommendations on best-practice accelerometer data collection and processing techniques to measure physical behaviours, as well as consensus recommendations on analytical approaches for the field of PA epidemiology. Moreover, this Doctoral Thesis highlights the value of open-source data processing algorithms and the important role of PA, SB, sleep, and the activity-rest pattern in relation with brain health outcomes in children with overweight or obesity. Finally, this Doctoral has demonstrated that meaningful and positive changes in cardiometabolic health in children with overweight or obesity can be obtained with a 20-week exercise program, which should inform future health programs.

Resumen

Contexto | Los acelerómetros son el método preferido para cuantificar los comportamientos físicos (actividad física [AF], comportamiento sedentario [CS] y sueño) en estudios científicos. Los grandes avances tecnológicos y el acceso a los datos brutos de los acelerómetros permiten afrontar una serie de retos relacionados con la necesidad de métodos de procesamiento de datos transparentes, comparables y reproducibles. Las asociaciones de los comportamientos físicos con la salud están ampliamente estudiadas, y a su vez deben extenderse para comprender mejor las interrelaciones entre los comportamientos físicos en su relación con la salud. Los niños con sobrepeso u obesidad podrían beneficiarse de los comportamientos físicos para mejorar su salud cardio-metabólica y cerebral.

Objetivos | Esta Tesis trata de responder a dos objetivos generales: (i) avanzar el conocimiento actual en cuanto a métodos de recolecta y procesamiento de datos de acelerómetros para estimar comportamientos físicosen niños y adultos; y (ii) explorar las asociaciones de los comportamientos físicos (medidos con acelerómetros) y la salud cardio-metabólica y cerebral en niños con sobrepeso u obesidad, así como los efectos del ensayo aleatorizado controlado ActiveBrains.

Métodos | Los diseños de estudio incluidos en esta tesis son una revisión sistemática, un artículo de descripción de software, siete estudios transversales, un artículo de consenso entre expertos y un ensayo aleatorizado controlado. Esta tesis utiliza datos principalmente del proyecto ActiveBrains, y complementariamente del estudio MINISTOP y de un estudio piloto de acelerometría. En todos ellos, los acelerómetros ActiGraph GT3X+ colocados en la cadera derecha y en las muñecas se utilizan para cuantificar los comportamientos físicos. Además, esta tesis incluye medidas *'gold-standard'* de gasto energético (i.e., agua doblemente marcada), volumen de materia gris cerebral (i.e., resonancia magnética), salud cardio-metabólica (biomarcadores sanguíneos) y composición corporal (i.e., absorciómetro dual de rayos X). Los análisis estadísticos utilizados incluyen regresiones lineales y cuadráticas, análisis de varianza (ANOVA), análisis de datos composicionales, análisis de patrones multivariantes y modelos de mediación.

Hallazgos principales | En cuanto al objetivo 1, esta Tesis: (i) proporciona recomendaciones basadas en la literatura existente sobre los criterios para recoger y procesar datos de acelerómetros; (ii) describe un software de acceso libre (en el que el candidato a PhD es co-desarrollador) para procesar datos de acelerómetros para la cuantificación de comportamientos físicos; (iii) encuentra que las métricas de aceleración *'open-source'* estiman mejor el gasto energético que las cuentas de actividad proporcionadas por la marca de acelerómetros; (iv) observa que las métricas *'open-source'* se comparan mejor entre sí que con las cuenta de actividad, y proporciona puntos de corte para cuantificar la intensidad de la AF a partir de datos de la muñeca dominante; (v) demuestra grandes discrepancias en la cuantificación del tiempo en CS e intensidades de AF cuando se utilizan distintos puntos de corte, sugiriendo que no es posible conocer la prevalencia de personas que cumplen las recomendaciones de AF en una población a partir de datos de acelerómetros; (vi) propone que las métricas relacionadas con los pasos (pasos/día y varios indicadores de cadencia) son un buen indicador de la AF general (i.e., counts al día, AF ligera-moderada-vigorosa, y AF

moderada-vigorosa) desarrollada por niños con sobrepeso u obesidad; (vii) proporciona una descripción comprensiva y un consenso internacional sobre las estrategias de análisis de datos que deben ser utilizadas para una determinada pregunta de investigación.

En cuanto al objetivo 2, la presente Tesis: (viii) observa que la asociación de la AF y CS con el volumen de materia gris en el hipocampo en niños con sobrepeso u obesidad podría estar moderada por el estado ponderal de peso (reemplazar 20 min/día de CS por AF moderada-vigorosa se asoció con 100 mm^3 de más materia gris en el hipocampo derecho en niños con obesidad tipo I); (ix) encuentra que los comportamientos de sueño están asociados con el volumen de materia gris en varias regiones corticales y subcorticales del cerebro, independientemente del CS y AF, y que estas asociaciones parecen ser relevantes para el rendimiento académico de niños con sobrepeso u obesidad; (x) remarca que un patrón de actividad-descanso más estable y menos fragmentado (así como hacer AF más temprano) se asocia con un mejor rendimiento académico, función ejecutiva e inteligencia en niños con sobrepeso u obesidad; (xi) demuestra que un programa de ejercicio de 20 semanas mejora la salud cardio-metabólica en niños con sobrepeso u obesidad, mientras que no se observó ningún efecto en salud mental.

Conclusión | Los resultados de esta Tesis Doctoral Internacional proporcionan recomendaciones sobre las mejores formas de recoger y procesar datos de acelerómetros para medir los comportamientos físicos, así como recomendaciones consensuadas sobre las estrategias de análisis para el ámbito de la epidemiología de la AF. Además, esta Tesis Doctoral subraya el valor del procesamiento de datos y algoritmos *'open-source'* y el papel de la AF, CS, sueño y patrones de actividad-descanso en relación con la salud cerebral en niños con sobrepeso u obesidad. Finalmente, esta Tesis Doctoral ha demostrado que se pueden obtener cambios positivos y significativos en salud cardio-metabólica de niños con sobrepeso u obesidad con un programa de ejercicio de 20 semanas, lo que debe ser tenido en cuenta en futuros programas de salud.

Abbreviations

ADHD | Attention-deficit hyperactivity disorder
ANOVA | Analysis of variance
BC | Bias corrected
BFEN | Band-pass filtering Euclidean Norm
BMI | Body mass index
BMR | Basal metabolic rate
CI | Confidence interval
CPM | Counts per minute
CRF | Cardiorespiratory fitness
CVD | Cardiovascular disease
DFT | Design fluency test
DNMS | Delayed non-match-to-sample
DXA | Dual-energy X-ray absorptiometry
ENMO | Euclidean Norm minus one *G* with negative values rounded to zero
ESM | Electronic supplementary material
FIRST | FMRIB's Integrated Registration and Segmentation Tool
FSL | FMRIB's Software Library
GGT | γ-glutamyl transferase
GMV | Grey matter volume
GPL | General Public License
HDL | High-density lipoprotein cholesterol
HEARTY | Healthy Eating Aerobic and Resistance Training in Youth
HFEN | High-pass filtered Euclidean norm
HFEN$_+$ | HFEN plus the Euclidean norm minus 1*G*
HOMA | Homeostatic model assessment
ICAD | International Children's Accelerometry Database
ICC | Intraclass correlation coefficient
IQ | Intelligence quotient
IS | Interdaily stability
ISCOLE | International Study of Childhood Obesity, Lifestyle and Environment
ITT | Intention-to-treat
IV | Intradaily variability
K-BIT | Kaufman brief intelligence test
L5 | Mean counts per minute of the 5 h with the lowest activity in the day
LCCC | Lin's concordance correlation coefficient
LDL | Low-density lipoprotein cholesterol
LMVPA | Light-moderate-vigorous physical activity
LFE | Low-frequency extension
LFECounts | Activity counts with the low-frequency extension filter in the vector magnitude
LFENMO | Euclidean Norm minus one *G* of the low-pass filtered accelerations with the negative values rounded to zero
LGPL | Lesser General Public License
LPA | Light physical activity
M10 | Mean counts per minute of the 10 h with the highest activity in the day
MAD | Mean amplitude deviation
MAPE | Mean absolute percent error
MetS | Metabolic syndrome
MET | Metabolic equivalent
MIMS | Monitor Independent Motion Summary
MINISTOP | Mobile-based INtervention Intended to STop Obesity in Pre-schoolers
MNI | Montreal Neurological Institute
MPA | Moderate physical activity
MPRAGE | Magnetization-prepared rapid gradient-echo
MRI | Magnetic resonance imaging
MVPA | Moderate-to-vigorous physical activity
MX | Person's most active X min/time (MX)
N_D | Deuterium dilution space
N_O | ^{18}O dilution space
NHANES | National Health And Nutrition Examination Survey
PA | Physical activity
PAEE | Physical activity-related energy expenditure
PRISMA | Preferred reporting items for systematic reviews and meta-analyses
RCT | Randomized controlled trial
rMSE | root mean square error
ROI | Region-of-interest
ROC | Receiver operating characteristic
ROC-AUC | Receiver operating characteristic area under the curve
SB | Sedentary behaviour
SD | Standard deviation
SE | Standard error
SPM | Statistical parametric mapping
T2D | Type 2 diabetes
TEE | Total energy expenditure
TL5 | Timing of the 5 h with the lowest activity in the day
TM10 | Timing of the 10 h with the highest activity in the day
TMT | Trail making test
VA | Vertical axis
VM | Vector magnitude
VACounts | Activity counts in the vertical axis
VMCounts | Activity counts in the vector magnitude
VO_2max | Maximal oxygen consumption
VPA | Vigorous physical activity
VVPA | Very vigorous physical activity
WASO | Wakening after sleep onset
WET | Work energy theory
WHO | World Health Organization
WHS | Women's Health Study
WOF | World Obesity Federation

PREFACE

"Every minute outside and awake is a good minute"

Calvin and Hobbes

CONTENTS

Historical view on the 'Physical Activity and Health' field

Abbreviations in this page:
PA: physical activity

It is by no means easy to set the onset date of an event in history, and the 'Physical Activity and Health' field is not an exemption. Everyone in the world has ever heard the motto '*Mens sana in corpore sano*', or in English 'A sound mind in a sound body', calling for the link between physical and mental health. Since this motto is frequently-quoted in Latin, one may think the Ancient Romans used it to highlight the health benefits of physical activity (PA). Nothing is further from reality.

The 'Mens sana' motto first appeared in a collection of satirical poems written by Juvenal, a Roman poet, in the 2nd century AD [1]. Juvenal's intention was no other than teaching Roman citizens what they should pray for. In the book IV of Juvenal satires, the poet debates about the myriad objects that prayers sought from the gods such as beauty, wealth or long life. Under Juvenal's criterion, it was piacular to bother the gods with such self-interested wishes. Specifically, the Satire X states that praying for a healthy mind and a healthy body is everything prayers should seek from the gods. Nothing to do with PA, neither with exercise or sports as can be seen in **Figure 1**.

orandum est ut sit ***mens sana in corpore sano.***	*It is to be prayed that* ***the mind be sound in a sound body.***
fortem posce animum mortis terrore carentem,	*Ask for a brave soul that lacks the fear of death,*
qui spatium uitae extremum inter munera ponat	*which places the length of life last among nature's blessings,*
naturae, qui ferre queat quoscumque labores,	*which is able to bear whatever kind of sufferings,*
nesciat irasci, cupiat nihil et potiores Herculis aerumnas credat saeuosque labores	*does not know anger, lusts for nothing and believes the hardships and savage labors of Hercules better*
et uenere et cenis et pluma Sardanapalli.	*than the satisfactions, feasts, and feather bed of an Eastern king.*
monstro quod ipse tibi possis dare; semita certe	*I will reveal what you are able to give yourself; For certain, the one footpath of a*
tranquillae per uirtutem patet unica uitae	*tranquil life lies through virtue*
	Satire X by Juvenal

Figure 1
Satire X by Juvenal (2nd century AD)

Then, who and why first linked this motto to physical and mental health? To answer this, we should travel in space (from Rome to Liverpool) and time (from the 2nd century to the 19th century AD). It was John Hulley, who used this motto to promote his football team 'Liverpool Athletic Club'. This motto fitted perfectly in the English boarding schools, which offered not only intellectual education but also a thorough physical training, looking for a complete education. After that, it was Pierre de Coubertin himself, the main developer of the modern Olympic Games, who used the '*Mens Sana*' motto to promote the physical exercise benefits. The motto was recurrent in the

Abbreviations in this page:
PA: physical activity

Olympics advertisement afterward (see an example in the Olympic medal from Amsterdam, 1928; **Figure 2**).

Figure 2
Medal from the Olympic Games in Amsterdam, 1926

Beyond poems and mottos, Ancient physicians, mainly from China and Greece, did believe in the value of PA for health. For example, 'if we could give every individual the right amount of nourishment and exercise, not too little and not too much, we would have found the safest way to health' is a quote attributed to the Ancient Greek physician Hippocrates, considered by many as 'the father of medicine'. However, a diametrically opposed view prevailed until the 20th century when complete bed rest was prescribed for patients with acute myocardial infarction.

In 1953, Morris et al. found an increased risk of myocardial infarction in bus drivers compared with bus conductors [2]. Additionally, mortality rates after the myocardial infarction occurrence were higher in bus drivers (**Figure 3**) [2]. Given the inherent PA levels to those professions (i.e., being bus drivers less active than conductors), PA was considered the main factor that explained this increased risk [2]. This study is considered by many as the origin of the 'Physical Activity and Health' scientific field. Since then, a myriad of observational studies have reported the PA benefits on non-communicable diseases and mortality [3–5].

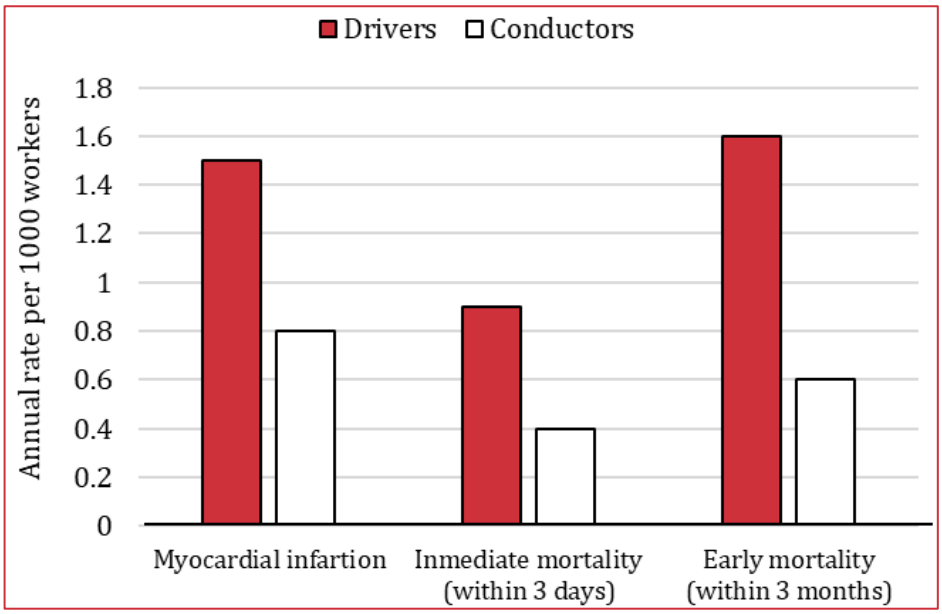


Figure 3
Myocardial infarction incidence and related mortality rates reported in the study by Morris et al. in 1953 [2]

Once the relationship between PA and health was repeatedly observed, the next major question of the field was about the amount: how much PA should be performed to be healthy? The American College of Sport Medicine was an early reference with the publication in 1975 of *Guidelines for Graded Exercise Testing and Exercise Prescription* [6]. This book and its subsequent revised edition had a major influence on the fields of exercise science and clinical and

rehabilitation medicine. The next major development in public health recommendations for PA was led by the American College of Sports Medicine and the Centers for Disease Control and Prevention in 1995. The specific recommendation outlined in this document was to engage in at least 30 min of MPA, which has been highly influential in the modern guidelines. Already in 2008, the US Government launched an update on the public health recommendations for PA [7]. Not much later, in 2010, the World Health Organization (WHO) launched the international version of these guidelines [8]. The PA guidelines were developed mainly based on PA data assessed with self-reports. In brief, 60 min per day of moderate-to-vigorous PA (MVPA) were recommended for children and adolescents, and 150 min per week of MVPA for adults and older adults. Hereinafter, these cut-off points have widely served to investigate how the attainment to PA guidelines is associated with an array of health outcomes. Currently, further efforts from the US Government and the WHO have resulted in updated recommendations on PA, which now encompass evidence based on self-report and modern technologies as tools to asses PA.

Moving on to the last decade, the prestigious scientific journal '*The Lancet*' launched its first version of the 'Physical Activity Series' (https://www.thelancet.com/series/physical-activity) in 2012, at the time of the London Olympic Games. These Series included high quality research on global PA surveillance, PA promotion and associations with health. The global surveillance of PA was carried out with self-reports given the lack of consensus for PA assessment with activity monitors. Hallal et al. described PA data in adolescents from 105 countries, and in adults from 122 countries [9]. Worldwide, they found that around 30% of adults did not reach the recommended 150 min of MVPA per week; and neither around 80% of adolescents the recommended 60 min of MVPA per day [9].

The worldwide physical inactivity health threats were described by Lee et al. [10]. They quantified that 9% of the worldwide premature mortality is attributable to physical inactivity [10]. In this regard, it is important to differentiate between physical inactivity and sedentary behaviour (SB). Time spent in SB has demonstrated to be an independent (of PA) risk factor for cardiovascular disease (CVD) [11], with around 4% of premature mortality worldwide attributable to it [12]. Thus, one may be active and sedentary at the same time if performing the recommended amount of PA, but also engaging too long in SB.

Both PA and SB coexist with sleep time in the 24 h of the day; insufficient sleep is a risk factor for CVD [13]. Interrelationships between PA, SB and sleep, collectively described as physical behaviours [14], have drawn the attention of a number of researchers worldwide, giving birth to a new field that could be described as 'Physical

Abbreviations in this page:
CVD: cardiovascular disease
MPA: moderate physical activity
MVPA: moderate-to-vigorous physical activity
PA: physical activity
SB: sedentary behaviour
WHO: World Health Organization

Abbreviations in this page:
PA: physical activity
SB: sedentary behavior

Behaviour epidemiology' (**Figure 4**). These interrelationships are mainly defined as: the usual high correlations observed between these behaviours; and the fact that increasing the daily time devoted to any of these behaviours (e.g., PA) would reduce the daily time spent in at least one of the remaining behaviours (e.g., SB and/or sleep). Therefore, the 'Physical Activity and Health' field is evolving to the 'Physical Behaviour Epidemiology' field, which englobes more precisely the studies included in this Thesis.

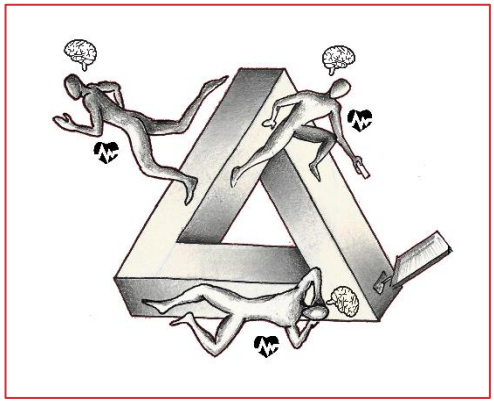


Figure 4
Graphical representation of the 'Physical Behaviour Epidemiology' scientific field.
Designed for this PhD thesis using the tribar as a base.

In the last decade, the use of accelerometers to assess PA, SB, and sleep has become an objective and feasible alternative to self-reports. Accelerometers measure the accelerations of the body placement where they are attached, serving as indicator of body movement (PA). Algorithms can be applied to estimate the PA intensity, volume, type, and other related behaviours, such as SB or sleep. The extremely rapid development in this field allows for a better determination of the physical behaviours' relationships with health in humans. ActiGraph (Pensacola, FL, USA) is one of the pioneer manufacturers of accelerometers, which are still today widely used for the study of PA and health. First ActiGraph (formerly known as the Manufacturing Technology Incorporated [MTI] ActiGraph, and the Computer Science Applications Inc. [CSA]) devices were developed in the early 1990s for applications within the US military forces. Shortly after, accelerometers were applied to sports sciences as an objective way of assessing PA in free-living conditions, being the European Youth Heart Study one of the pioneer studies in using accelerometers in 1998-99 to quantify PA in children and adolescents at population level [15,16].

Key concepts

Core terms

Physical behaviours | In the last few years, the 'Physical Activity and Health' research field has not only focused on PA. The field has included other constructs such as SB and sleep. These behaviours share the 24 hours of every day, which makes them highly dependent on each other. Altogether, these behaviours have been named as 'physical behaviours' and represent a change in the focus of the field from studying the isolated associations of PA with health, to the study of the combined effects and interactions between physical behaviours in relation to health [17]. In this Thesis book, physical behaviours refer collectively to PA, SB and sleep.

Physical activity | PA is any bodily movement produced by skeletal muscles that requires energy expenditure [18]. This term encompasses all types (e.g., walking, running, dancing), intensities (e.g., light [LPA], moderate [MPA], vigorous [VPA]) and domains (e.g., leisure, occupational, transportation, household) of PA. Additionally, PA can also be considered relative to the physiological effect produced (e.g., aerobic, anaerobic, muscle-strengthening). In this thesis, accelerometer-determined PA encompasses all PA types, intensities and domains, with special attention to walking, aerobic PA and muscle-strengthening PA in specific studies.

Sedentary behaviour (SB) | SB refers to any waking behaviour characterized by an energy expenditure ≤ 1.5 metabolic equivalents (METs), while in a sitting, reclining or lying posture [19]. As such, SB presents two major features: low energy expenditure and the body posture. While accelerometers have been widely used in relation to energy expenditure, the usefulness of accelerometers data to estimate body posture is limited [20]. Therefore, the estimation of SB via accelerometers should be interpreted with caution. In this Thesis, only time spent at SB is considered given the limitations of the body posture estimation using accelerometers data. This term is often used wrongly as a synonym of physical inactivity, but it is not. See definition of physical inactivity below.

Sleep | Sleep is defined as an active, repetitive and reversible brain process of reduced perception and responsiveness to environmental stimuli [21]. From a movement perspective, sleep periods are characterized by very low movements during a relatively long period of the day, usually in a lying posture.

Accelerometer | Technically, accelerometers are sensors able to detect accelerations. Accelerometer sensors are included in an array of devices, such as smartphones, activity monitors, pedometers, or cars. Among the applications of accelerometers, it is their usability to estimate physical behaviours. Accelerometers data can be used to estimate PA, SB and sleep by applying algorithms to the data. While accelerometer refers specifically to the sensor, in this Thesis, the term accelerometer is used to refer to the whole device, being used therefore as synonym of activity monitor or tracker among others.

Other related terms

Physical activity intensity | The intensity of the PA is quantified in terms of the energy expenditure that it requires to be performed. A relative measure of the energy expenditure is often used as criterion, namely METs. The METs indicate the relative energy cost of a certain PA in terms of ml of O_2 consumed per kg of weight during one minute (ml/kg/min). On average, adults consume 3.5 ml/kg/min in basal conditions, which is often used as an absolute reference (i.e., 1 MET = 3.5 ml/kg/min). Relative METs can be derived from the

Abbreviations in this page:
LPA: light physical activity
MET: metabolic equivalent
MPA: moderate physical activity
MVPA: moderate-to-vigorous physical activity
PA: physical activity
SB: sedentary behavior
TEE: total energy expenditure
VPA: vigorous physical activity

Abbreviations in this page:
BMI: body mass index
LPA: light physical activity
MET: metabolic equivalent
MPA: moderate physical activity
MVPA: moderate-to-vigorous physical activity
PA: physical activity
PAEE: physical activity-related energy expenditure
SB: sedentary behavior
TEE: total energy expenditure
VPA: vigorous physical activity

measurement of the energy consumption of each individual at basal conditions. In the 'Physical Activity and Health' field, PA intensities have been usually classified into LPA (e.g., 1.5-3 METs), MPA (e.g., 3-6 METs) and VPA (e.g., >6 METs). Traditionally, the 'Physical Activity and Health' field has mainly focused on MVPA. However, given the current interest and discussions about PA of intensities less than MPA, the term PA is used in this Thesis to discuss the full range of intensities (i.e., from LPA to VPA).

Exercise | The term exercise refers to a subcategory of PA, characterized for being planned, structured, repetitive and purposive [18]. Usually, the major aim of exercise is to improve at least one component of physical fitness.

Physical inactivity | Physical inactivity should not be confounded with SB. Physical inactivity stands for an insufficient PA level to meet the current public health recommendations on PA [19]. Oppositely, to be physically active refers to meeting the PA recommendations. As such, a child might be physically active and sedentary at the same time if engaging in at least 60 min of MVPA per day, but concomitantly spending long times in SB.

Physical activity-related energy expenditure | Component of the total energy expenditure (TEE) related to the performance of PA, usually measured in kcal or kJ. Although PAEE is not the biggest component of TEE, it is the most malleable; as such, PAEE is the main focus of researchers and health care professionals dealing with energy balance.

Obesity

Obesity | Obesity defines abnormal or excessive adiposity in the organism, as measured by body fat percentage. Furthermore, its distribution in the body is also determinant of health, being the abdominal or central fat the most harmful. Central fat, measured as waist circumference, is considered a cardiometabolic risk factor for both children and adults. Apart from the excess of adiposity, it is important to consider that most of the current knowledge on the health consequences of obesity is based on body mass index (BMI). Therefore, obesity could also mean an excess of body weight rather than only adiposity.

Overweight | Similar to obesity, overweight refers to an excess of body adiposity, although at a lower level than obesity. Overweight is the weight status above normal-weight and below obesity.

Body mass index (BMI) | BMI is a measure of body composition that encompasses the body weight relative to the height (i.e., kg/m^2). Given its simple measurement, it is widely used as indicator of obesity even if it does not directly measure body fat. There are internationally accepted cut-off points for normal-weight (18.5 – 24.9 kg/m^2), overweight (25 – 29.9 kg/m^2) and obesity (≥ 30 kg/m^2).

Although it has been widely criticized, BMI has demonstrated a better predictive capacity for premature mortality than 'gold-standard' measures of adiposity [22].

Cardiometabolic health

Cardiometabolic health | In this thesis, the term 'physical health' refers to cardiometabolic health, body composition, physical fitness components. Even though physical health could encompass a wider range of health outcomes (e.g., musculoskeletal health), this thesis rather focuses on the abovementioned components.

Body composition | Body composition is considered an important component of cardiometabolic health given their well-established relationships. Body composition refers to a method of studying the body as a composite of its core components: fat mass, lean mass, and water.

Physical fitness | Physical fitness relates to the ability to perform PA. Fitness is a status that can be effectively modified through regular PA, but which also depends on genetic factors [18]. It is a powerful marker of health across the lifespan [23]. In childhood, the fitness components more related to health include cardiorespiratory (CRF), muscular and speed-agility fitness. CRF is the ability to perform PA during a long period of time, usually measured as the maximal oxygen consumption capacity (VO_2max). Muscular fitness measures the muscle strength and it is usually measured separately for upper and lower limbs. Speed-agility fitness is usually measured with the time spent in carrying out a certain circuit which requires speed and agility.

Cardiometabolic health | The cardiometabolic health refers to status of those factors that are important for the cardiovascular and metabolic health. This includes, but it is not limited to: blood pressure, lipid biomarkers (i.e., high-density lipoprotein [HDL], low-density lipoprotein [LDL], total cholesterol and triglycerides), glycemia biomarkers (i.e., insulin and glucose) and waist circumference. Poor levels in these indicators are high-risk factors for the development of type 2 diabetes (T2D) and CVD.

Brain health

Brain health | Brain health includes brain biomarkers (e.g., grey matter volume [GMV]) and factors associated with academic achievement, cognition and mental health [24].

Brain structure | The human brain is the central organ of the human nervous system. The human brain consists of the cerebrum, the brainstem and the cerebellum. In this Thesis book, the main component of the brain structure investigated is the GMV. The grey matter includes vasculature, glial cells and neuronal bodies and contributes to the process of information in the brain.

Abbreviations in this page:
BMI: body mass index
CRF: cardiorespiratory fitness
CVD: cardiovascular disease
GMV: gray matter volume
HDL: high-density lipoprotein
LDL: low-density lipoprotein
PA: physical activity
T2D: type 2 diabetes
V_2O: maximal oxygen consumption

Abbreviations in this page:
GMV: gray matter volume
IQ: intelligence quotient

Academic achievement | Rather than focusing on the school grades, a continuous, standardized, valid, and reliable measure of the academic achievement is used in this thesis. We used the Spanish version of the Woodcock-Muñoz test for academic achievement. Trained examinators assessed children individually in a face-to-face protocol lasting around 100-120 min per child. The battery included tests of reading, language, mathematics, and sciences.

Executive function | Executive functions include a set of cognitive processes that are necessary for the cognitive control of behaviour. In ActiveBrains, the executive function domains measured were cognitive flexibility, inhibition, and working memory.

Intelligence | Although the term intelligence may include a broad set of characteristics hardly measurable, this thesis included an intelligence quotient (IQ) score. Such score was composed of crystallized and fluid intelligence scores from a standardized test. Even if this score is not full representative of intelligence, it is a proxy to what traditionally has been defined as intelligence.

Psychological ill-being | Psychological ill-being represent unpleasant feelings or emotions that represent pre-clinical psychological states and clinically diagnosed mental health disorders. Specifically, psychological ill-being refers to stress, anxiety, depression and negative affect.

Psychological well-being | Psychological well-being is a composite score of those positive affective states and functioning with optimal effectiveness in life. In this Thesis, psychological well-being is composed of positive affect, happiness, optimism.

Psychological self-perceptions | Psychological self-perceptions represent the opinions and feeling of individuals on themselves. In this Thesis, the psychological self-perceptions investigated are self-efficacy, self-concept and self-esteem.

INTRODUCTION

“Life is like riding a bicycle. To keep your balance, you must keep moving.”

Albert Einstein

CONTENTS

Motivation

Research in the 'Physical activity and Health' has been incredibly enhanced by the development of new technologies to better quantify PA and related behaviours (e.g., SB, sleep, among others). Indeed, these advances allows a more objective and accurate quantification of the behaviours, which allow to draw a more precise picture on how these behaviour associate with health benefits. However, the rapid advances have also led to inconsistencies on the collection, treatment, and application of the accelerometer data to answer specific research questions. The field would therefore benefit from studies on the accuracy and comparability of different data collection and processing techniques in the quantification of physical behaviours (i.e., PA, SB, and sleep), as well as from the appropriate data analytical approaches to examine their relationship with health outcomes. These two blocks are approached in this Thesis, which is accordingly organized in two sections.

Abbreviations in this page:
PA: physical activity
SB: sedentary behaviour

Measurement of physical behaviours

There cannot be study of the associations of physical behaviours with health without valid and reliable quantification of physical behaviours. This premise underlies the first section of this Thesis. Interestingly, the first method used to quantify PA in epidemiological research simply consisted in the assumption that the PA level is relative to work [2]. This assumption has been repeatedly confirmed [25–27], being the occupation an indicator of the overall PA, especially in those workers with a low education level [25]. Although the occupation can be an indirect estimator of the PA level, there currently exist alternatives to directly assess PA. Overall, the PA assessment tools can be classified into self-reported, activity monitors, and 'gold-standard' or criterion methods. Although validity and reliability are higher using 'gold-standard' methods, their feasibility is very limited. An introduction and contextualization of currently available PA assessment methods in each of these categories is needed to understand the reasons why this thesis focuses on the assessment of PA with accelerometers.

Self-report tools

Self-reports or parental reports (for younger populations) have been used for more than 50 years so far [28]. Among their advantages, self-reports are of low burden for the participants and researchers, cost-effective, versatile, and accepted by the research and the medical communities [29]. However, their limitations are notable, such as: the participants subjectivity which may bias the findings [30]; or the social desirability, which may affect the answers to the questionnaires [31,32]. Traditionally, the advantages of self-reports have outpointed their limitations, so that they have been widely used

Abbreviations in this page:
PA: physical activity
SB: sedentary behavior

in surveillance, epidemiology, cross-sectional or longitudinal studies. Most of the knowledge on the 'Physical Activity and Health' field was based on self-report measures, even public health guidelines on PA were first developed attending to the data described in self-reports.

In the last decades, the 'Physical Activity and Health' field has evolved with the awareness of the limitations of self-reports and the development of objective methods to quantify PA in the daily life. Many observational studies have corroborated or contrasted previous assumptions from self-reported data using activity monitors. However, self-reports are still the method of choice for global surveillance [9] given the lack of a consensus and the high impact of the data processing on the quantification of PA with activity monitors.

Activity monitors or wearables

Activity monitors or wearables are any device which can be placed on a person to monitor certain aspects related to PA (**Figure 5** for an example). Basically, activity monitors include pedometers, accelerometers, heart rate sensors, and multi-sensor systems based on combinations of the previous sensors. Pedometers and accelerometers provide mechanical information on PA (i.e., movement), while heart rate monitors record the physiological response to PA. The 'Physical Activity and Health' field is increasingly using pedometers and accelerometers to measure various components of PA. In regard to heart rate monitors, they are less feasible for long recordings (i.e., longer than 7 days). Therefore, pedometers and accelerometers are more-frequently used, and heart rate monitors are mainly used as a criterion method to validate or calibrate the previous ones in terms of PA intensity assessment.

Figure 5
ActiGraph GT3X+ accelerometer. This is one of the models most-frequently used in research.

There is not a defined gold-standard among activity monitors [33]. The choice of which activity monitor to use depends on a number of factors, such as the specific component of PA of interest, the target population of its cost-effectivity, among others. In brief, pedometers provide an adequate solution to measure the main form of PA performed by the general population [34,35]. Furthermore, they provide an easily interpretable and understandable measure (i.e., steps) that can be used to quantify PA in cross-sectional, longitudinal or intervention studies, as well as a target to achieve the recommended level of PA. Otherwise, modern accelerometers (see example in **Figure 5**) record the accelerations produced by the body attachment site where they are placed on throughout the day. Acceleration information is much richer than steps, since there are available algorithms to estimate not only steps, but also other activity types (such as

sitting or cycling [36]), time spent in different PA intensities [37] or SB, or even certain characteristics of sleep [38–40]. Therefore, accelerometers are being widely used in the 'physical behaviour epidemiology' field. The accelerometer data processing methods are being developed drastically fast, with no clear consensus on the best practices to collect and process accelerometer data.

Abbreviations in this page:
PA: physical activity
SB: sedentary behavior

'Gold-standard' or criterion tools

Finally, the 'gold-standard' or criterion methods are mainly used as reference in validation or calibration studies. They are not feasibly applicable in large populations during long periods, so they cannot be used to measure the habitual PA performance in the daily life. However, they represent the ground truth when evaluating the validity of the abovementioned activity monitors or self-reports.

These methods encompass measures of the oxygen consumption or energy expenditure through direct or indirect calorimetry or doubly labelled water, as well as activity type classification through direct observation, among others. They are characterized by a high accuracy and reliability and a very low feasibility for their application in large populations or for long-term recordings. Therefore, they are not used for lifestyle PA assessment, but they are used as criterion methods to validate either self-reports or activity monitors for the assessment of PA, SB or other related behaviours.

Why does this Thesis focus on activity monitors?

Figure 6 shows a broad categorization of the methods to assess PA abovementioned in terms of accuracy and feasibility. Activity monitors are placed at a fair trade of feasibility and accuracy. Accelerometers are the cornerstone of this thesis because:

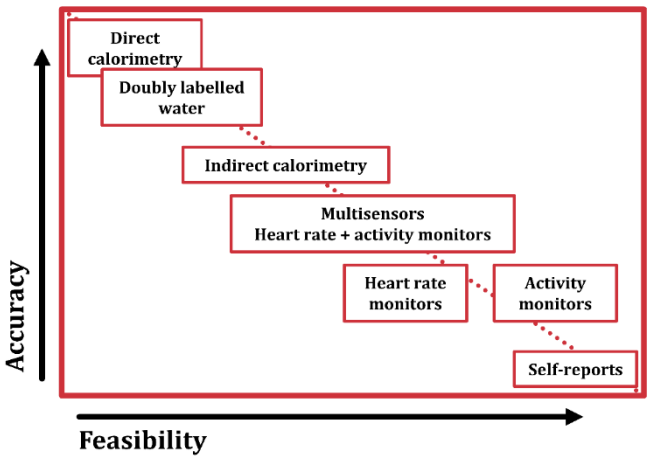


Figure 6
Broad comparison of PA assessment methods in terms of accuracy and feasibility.

1. Their application in research and clinic settings, and their commercialization has substantially increased [41].
2. Their drastic development requires research to improve comparability across data collection and processing protocols.
3. Accelerometers can substantially increase validity and reliability for the PA assessment compared to self-reports.
4. Beyond PA, accelerometers data can be used to estimate SB and sleep characteristics, which open a wide range of opportunities in the physical behaviour epidemiology field.

Abbreviations in this page:
BMI: body mass index
CVD: cardiovascular disease
PA: physical activity
T2D: type 2 diabetes
WHO: World Health Organization
WOF: World Obesity Federation

Physical behaviours and health in childhood obesity

Obesity is a devastating condition that shortens the lifespan by affecting the function of many organ systems [42]. Obesity increases the risk of T2D and CVD [43], two of the top-10 leading causes of death worldwide [44]. Not only that, obesity itself is estimated to cause more than 2.6 million deaths per year. Obesity may onset in childhood, having dramatic health consequences also at early ages. Children with obesity present poor cardiometabolic [45,46] and mental health [47], and high risk for T2D and CVD later in life [48].

The health consequences and the high prevalence of childhood obesity worldwide place it as one of the most serious global public health challenges of the 21st century. The WHO and the World Obesity Federation (WOF) estimated more than 340 million children and adolescents with overweight or obesity worldwide in 2016. The rate of children with overweight or obesity has tripled from 1960 to date. The BMI increase has plateau in high-income countries in recent years, although at very high and worrisome values [49].

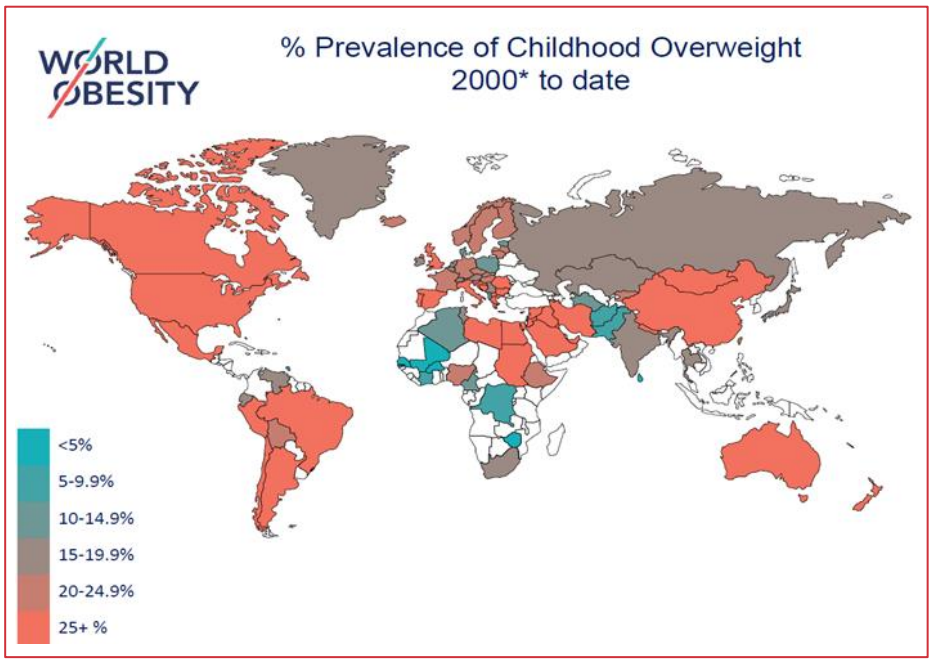


Figure 6
Worldwide prevalence of childhood overweight (including obesity) as defined by the WOF.

The high prevalence and the dramatic consequences of obesity throughout the lifespan require of cost-effective strategies to lower the obesity rates, and the risk of comorbidities in the populations living with obesity. Best practice interventions should target the lifestyle factors at early ages, and PA is key in the management of obesity in children [43].

Cardiometabolic health

The T2D burden in youths is alarmingly increasing [50], which is in parallel with the worldwide childhood obesity rise [49]. Almost every child diagnosed with T2D presents obesity [50]. Additionally, childhood obesity also is a high-risk factor for future T2D and CVD [48]. Physical fitness has demonstrated to be a powerful marker of

health throughout the lifespan, starting at your ages. Thus, best practice prevention of T2D and CVD onsets in childhood obesity, and PA (to efficiently improve physical fitness) should be the cornerstone of prevention strategies given its proven multi-organ benefits [51]. In this Thesis book, 'cardiometabolic health' includes an array of health outcomes related to cardiometabolic health, body composition, and physical fitness.

Extensive cross-sectional and longitudinal observational studies establish a link between PA and cardiometabolic health, body composition, and physical fitness in children [51]. Specifically, there is evidence enough to conclude that PA is favourable for bone health, adiposity, and physical fitness [51], but more research is needed to understand the PA associations with cardiometabolic health. Specifically, studies investigating the effects of PA on cardiometabolic health, body composition and physical fitness in children with obesity are limited. It seems that PA can be effective to improve the lipid profile [52] and the blood pressure [53] in children with obesity.

In this thesis, the effects of a PA program on cardiometabolic health, body composition, and physical fitness in children with overweight or obesity is addressed, with a detailed examination of the PA and SB profiles of the children.

Brain health

'Brain health' is defined as a composite of factors related with cognition, brain, and mind [24]. Improving brain health during childhood is important to enhance brain development, achieve academic goals and improve cognition [51]. Within the brain, GMV is a measure of the volume of tissue in the brain region being examined. It represents all tissue properties contained in grey matter including vasculature, glial cells, and neuronal cell bodies. Further, greater GMV in the developing brain is positively associated with brain health outcomes such as academic achievement or cognition [54], being crucial characteristic to success in school and in general life.

Lifestyle behaviours such as PA, SB and sleep are associated with academic achievement and cognition in children [51]. A recent Position Stand of the American College of Sports Medicine concluded that there is evidence supporting that PA can benefit cognition, and particularly executive function (i.e., cognitive flexibility, inhibition and working memory) [55]. These relationships may be partly explained by the associations of these lifestyle behaviours with GMV [54]. Specifically, GMV in the hippocampus is crucial for short- and long-term memory [56,57]. GMV in the hippocampus can be amplified by PA and sleep [58]. However, most of the studies investigating the associations of PA and sleep in children with brain health outcomes have used self-reported PA. Investigations of these

Abbreviations in this page:
CVD: cardiovascular disease
GMV: grey matter volume
PA: physical activity
SB: sedentary behavior
T2D: type 2 diabetes

Abbreviations in this page:
GMV: grey matter volume
PA: physical activity
RCT: randomized controlled trial
SB: sedentary behavior

associations with objective methods would confirm previous literature and accurately quantify the magnitude of the association.

Less attention has been paid to the circadian rhythm of PA and sleep in relation to brain health, namely the activity-rest pattern. The activity-rest pattern defines the timing and stability of these behaviours throughout the day and across days. Unstable and fragmented activity-rest patterns are associated with obesity [59,60], lower CRF [60], and depression [59] in childhood.

In this Thesis, the associations of accelerometer-determined PA, SB and sleep with brain health will be investigated in children with overweight or obesity. Specifically, we will examine the GMV, academic achievement, executive function, and intelligence quotient (IQ). As discussed in the Preface, it is important to investigate the associations of PA, SB and sleep with health considering the co-dependencies of these behaviours (see **Figure 4**). The analyses performed in this thesis book have been selected to appropriately approach this co-dependency.

Thesis structure

This Thesis main topic of study is physical behaviours, including PA, SB and sleep, although the main focus is on PA. Two sections are separately presented in this thesis book based on the nature of the objectives they respond to. The **Section I** addresses those aims related to the objective measurement of PA, SB and sleep with accelerometers. Seven studies on methodological issues in which the accelerometer data were the object of study are presented. I have had the opportunity to work with data from the ActiveBrains and the MINISTOP (Mobile-based INtervention Intended to STop Obesity in Pre-schoolers) projects to answer these questions. The ActiveBrains project is a randomized controlled trial (RCT) aimed to test the effects of an exercise program on academic achievement, cognition, physical and mental health outcomes in children with overweight or obesity. The MINISTOP project is a mobile-based intervention intended to stop obesity in pre-schoolers. Additionally, I led a pilot study on how the body attachment site affects the accelerometer data recording in adults. The **Section II** focuses on the associations of PA, SB and sleep with physical and mental health outcomes in children with overweight or obesity. For this section, I analysed data from the ActiveBrains project on PA, SB, sleep, physical and brain health in children with overweight or obesity.

Gaps addressed in this Thesis

We identified a number of gaps in this current scientific knowledge that are addressed by the studies included in this Thesis book. The **Table 1** briefly describes these gaps, together with the contributions of this thesis book.

	Gap	Contribution
SECTION I	There is a lack of recommended procedures for accelerometer data collection and processing decisions to obtain valid, reliable and reproducible estimates of PA, SB, sleep.	Systematic review of the accelerometer data collection and processing protocols to assess PA and related behaviours (**Study I**). Practical recommendations based on the previous evidence (**Study I**)
	Comparability of the PA, SB and sleep estimates derived from accelerometers using different (yet frequently used) protocols is under explored	Co-development of an open-source software which allows consistent accelerometer raw data processing in several brands (**Study II**)
		Cross-sectional studies on the comparability of SB and PA outputs using different body attachment sites, data processing decisions and cut-points (**Studies III and IV**)
	The assessment of PAEE using wrist-worn accelerometer data is limited	Study on wrist accelerometer data processing for the estimation of energy expenditure in pre-schoolers (**Study V**)
	Cut-points for the assessment of SB/PA from dominant wrist-worn accelerometers are very limited	Development of cut-points for the assessment of SB and PA from dominant-wrist in adults (**Study IV**)
	The promotion of walking is cost-effective and beneficial for older adults. However, studies are lacking in children with overweight or obesity.	Cross-sectional study on how walking relates to overall PA in children with overweight or obesity. Implications for PA promotion (**Study VI**)
	There are not clear recommendations on the choice of analytical approaches to investigate the associations of physical behaviours with health.	Consensus study on the implications of: (1) descriptors to capture physical behaviours from accelerometer data, (2) statistical methods to analyse accelerometer-determined physical behaviours (**Study VII**)
SECTION II	Information on the association of PA and sleep with brain health in children with overweight or obesity is scarce	Cross-sectional studies on PA, SB and sleep associations with GMV, academic achievement and cognition in children with overweight or obesity (**Studies VIII, IX and X**)
	The PA benefits on physical health have been widely investigated. However, there is still the need to further investigate the effect of PA program holistically on body composition, physical fitness and cardiometabolic health in children with overweight or obesity	RCT to explore the effects of a 20-week long, multi-gamed based, concurrent aerobic and strength exercise program on cardiometabolic health (including body composition, physical fitness and cardiometabolic health) in children with overweight or obesity (**Study XI**)
	There has not been investigated whether changes on the daily time spent in physical behaviours due to an exercise program can explain the groups' and the individuals' responses to the exercise	Investigation of the changes in the daily time spent in physical behaviours produced during the exercise program, and whether these changes could explain differential responses to the exercise (**Study XI**)
	The chronic effects of an exercise program on mental health in children with overweight or obesity are under reported	RCT to explore the effects of a 20-week long, multi-gamed based, concurrent aerobic and strength exercise program on mental health (including psychological ill-being and well-being) in children with overweight or obesity (**Study XI**)

Table 1
Overall view of the gaps identified and the contributions of this Thesis book

Abbreviations in this page:
GMV: grey matter volume
PA: physical activity
PAEE: physical activity-related energy expenditure
RCT: randomized controlled trial
SB: sedentary behavior

AIMS

"If you want something, go get it. Period."

The pursuit of happiness

CONTENTS

Section I | Measurement of physical behaviours

Section I focuses on examining the influence of accelerometer data collection and processing decisions on the assessment of physical behaviours (i.e., PA, SB and sleep) and related outcomes (e.g., PAEE, steps). Three aims were addressed in this section with a systematic review, a software description study, four observational studies, and an expert consensus statement.

Aim I (Study I) – State-of-the-art and practical considerations

To provide practical considerations for the assessment of SB, PA, PAEE, steps, or sleep, with accelerometers based on an in-depth review of the accelerometer data collection and processing decisions made in the existing literature. This is approached via two specific objectives addressed in Study I:

1. To compile and classify existing studies assessing SB, PA, PAEE, or sleep using the ActiGraph GT3X+.
2. To review data collection and processing criteria when using GT3X+ and provide age-specific practical considerations based on the validation/calibration studies identified.

Aim II (Studies II-VI) – Data collection and processing

To investigate the accelerometer data collection and processing criteria influence in the estimation of energy expenditure, as well as the comparability of acceleration metrics, PA, SB, and steps when derived from differing data collection protocols (e.g., different devices or body attachment sites), or data processing decisions (e.g., different acceleration metrics, cut-points, or algorithms). The specific aims addressed in Studies II-VI are:

1. To provide a one-stop overview of the GGIR package, the papers underpinning the theory of GGIR, and how research contributes to the continued growth of GGIR (software co-developed by the PhD candidate).
2. To assess the capacity of different acceleration metrics from wrist accelerations to estimate TEE and PAEE assessed with doubly-labelled water.
3. To study the comparability between different acceleration metrics across right hip, dominant wrist, and non-dominant wrist attachment sites during different periods of the day (i.e., 24 hours, waking, and sleeping hours).
4. To use previously established cut-points for accelerations measured at the non-dominant wrist [61,62] to develop and cross-validate cut-points in a separate sample for accelerations measured at the dominant wrist in a sample of young adults.

Abbreviations in this page:
ENMO: Euclidean norm minus 1 *G*
LFENMO: ENMO of the low-pass filtered raw accelerations
MAD: mean amplitude deviation
PA: physical activity
PAEE: physical activity-related energy expenditure
SB: sedentary behaviour
TEE: total energy expenditure
VACounts: activity counts in the vertical axis
VMCounts: activity counts in the vector magnitude

Acceleration metrics: in this Thesis, the term acceleration metrics is used indistinctally from accelerometer signal aggregation. It refers to those metrics that are derived from processing the accelerometer raw signal to remove the gravitational acceleration and noise. Acceleration metrics include ENMO, MAD, LFENMO, VMCounts, VACounts, or Activity Index, among others.

Abbreviations in this page:
PA: physical activity
SB: sedentary behaviour

5. To examine how cut-points relative to different attachment sites and acceleration metrics affect the final estimations of SB and PA in children with overweight or obesity.
6. To investigate the proportion of overall PA that is explained by ambulatory activity in children with overweight or obesity.
7. To study step-based patterns relative to PA guidelines achievement in children with overweight or obesity.

Aim III (Study VII) – Consensus on data analytical approaches

To provide a comprehensive description, discussion, and consensus on the analytical approaches most-frequently used in the 'Physical activity and Health' field and their implications for the study of associations with health outcomes. Study VII approaches this through two specific aims:

1. To provide a comprehensive description and discussion on the most-frequently used analytical approaches (i.e., from descriptors to statistical modelling) currently used in the field, highlighting their strengths and limitations and providing practical recommendations on their use.
2. To identify current gaps and future research directions around the analysis and use of accelerometer data in the 'physical behaviour epidemiology' field.

Section II | Physical behaviours and health in childhood obesity

Section II is dedicated to study the association of accelerometer-determined physical behaviours with cardiometabolic and brain health outcomes in children with overweight or obesity. Physical behaviours are always considered as modifiable exposures in this section. Three cross-sectional studies and a RCT were carried out.

Abbreviations in this page:
GMV: grey matter volume
IQ: intelligence quotient
PA: physical activity
PAEE: physical activity-related energy expenditure
RCT: randomized controlled trial
ROI: region-of-interest
SB: sedentary behaviour

Aim IV (Studies VIII-X) – Cross-sectional studies on brain health

To study the cross-sectional associations of PA, SB and sleep with GMV, and their implications for academic achievement, executive function, and IQ in children with overweight or obesity. Specific aims in Studies VIII-X include:

1. To investigate associations of objectively-measured SB and PA with GMV in the hippocampus.
2. To explore the association of sleep behaviours with GMV in the different brain regions, with a special focus on the hippocampus as a region-of-interest (ROI) in children with overweight or obesity.
3. To study the associations of the activity-rest pattern indicators with academic achievement, executive function and IQ in children with overweight or obesity.
4. To investigate whether the GMV in those regions associated with the physical behaviours were also related to academic achievement, executive function, and IQ in children with overweight or obesity.

Aim V (Study XI) – RCT on cardiometabolic and mental health

To investigate the effects of a 20-week multigame-based PA program on cardiometabolic and mental health in children with overweight or obesity.

1. To investigate the effects of a 20-week exercise program on cardiometabolic and mental health in children with overweight or obesity.
2. To examine the within-individual variability in the effects observed.
3. To explore the exercise program fidelity with an in-depth look at the physical behaviour changes occurring during the exercise implementation.

METHODS

"Measure what is measurable, and make measurable what is not so"

Galileo Galilei

CONTENTS

Research projects

This Thesis is mainly derived from the **ActiveBrains** project (6 out of the 11 studies and a systematic review were conducted under the umbrella of the ActiveBrains project), and is complemented with the **MINISTOP** project and a **pilot study** on accelerometer data collection and processing methodology. Additionally, we conducted a software description article of the GGIR software (in which the PhD candidate is co-developer); and we led an **international expert consensus statement** on data analytical approaches for accelerometer-measured physical behaviours in epidemiological studies.

Abbreviations in this page:
ADHD: attention-deficit hyperactivity disorder
BMI: body mass index
PA: physical activity
RCT: randomized controlled trial
SB: sedentary behaviour
WOF: World Obesity Federation

The ActiveBrains project (Study I, V-VI, and VIII-XI)

Study design

The ActiveBrains project is a two-arm (1:1) RCT (NCT02295072) aimed to examine the effects of a 20-week PA program on brain structure and function, cognitive performance, academic achievement, and physical and mental health indicators in children with overweight or obesity (**Figure 7**) [63]. Children from Granada (Spain) were recruited in the Endocrinology Unit of San Cecilio and Materno Infantil Hospitals, and additionally via schools, radio, and television advertising. A total of 100 children were targeted upon statistical power analysis based on the primary outcome of the project (i.e., brain imaging – hippocampus). Finally, 110 children were enrolled and randomized after the baseline assessment into exercise or control (wait-list) group. The assessments and intervention were phased in three waves for logistical reasons. The ActiveBrains project was approved by the Ethics Committee on Human Research of the University of Granada. The Studies V, VI, VIII, IX, and X use the baseline data from the ActiveBrains project (cross-sectional studies), while the Study XI has a RCT design.

Figure 7
The ActiveBrains project logo

Inclusion and exclusion criteria

Eligible children were required: (i) to be **8.0 – 11.9 years old**; (ii) to present **overweight or obesity** based on the sex-and-age specific international BMI standards (WOF) [64,65]; (iii) **not to have any physical disabilities or neurological disorder** that limits exercising; (iv) **not to use medication** that influence central nervous system functioning; (v) to be **right-handed** as measured by the Edinburgh inventory (given the brain differences -primary outcome- in left- and right-handed individuals); (vi) **not to have attention-deficit hyperactivity disorder** (ADHD) as measured with the ADHD rating scale; (vii) to be pre-pubertal according to Tanner stages, and in the case of girls, **not to have started menstruation**.

Abbreviations in this page:
MVPA: moderate-to-vigorous physical activity
PA: physical activity
SB: sedentary behaviour
VMCounts: activity counts in the vector magnitude

Physical activity intervention

The PA program was based on the public health guidelines on PA for children (http://www.health.gov/paguidelines/). Specifically, we offered 90-min sessions from Monday to Friday, recommended children to attend to a minimum of 3 sessions/week during 20 weeks, and recommended to attend as many sessions as possible. Each session included a 10-min warm-up, 60 min of MVPA based on multi-games, 15-20 min of muscle- and bone-strengthening activities, and 5-10 min of cool-down stretching and relaxing activities. Each child wore a heart rate monitor in every session with 5 individualized intensity zones based on their previously-measured maximum heart rate (60-69%, 70-79%, 80-84%, 85-89%, and 90-100%). Our aim was to reach as much time as possible above their 80% of maximum heart rate. No dietary intervention was conducted, yet all participants (both control and exercise groups) received a booklet with healthy PA and dietary information.

Measures

A brief description of the measures from the **ActiveBrains project** which are used in this thesis book is presented. More details on all the evaluations conducted can be found elsewhere [63]. Likewise, a more detailed definition of the data collection for each specific study is presented in the results section of this thesis book.

Physical behaviours (explanatory/exposure variable)

This thesis is not focused on a specific primary outcome, it is centred around physical behaviours (**PA, SB, and sleep**) as lifestyle modifiable exposure (or explanatory variable) and their relationship with an array of physical and brain health outcomes instead.

For the physical behaviour assessment, participants were required to wear two accelerometers (**ActiGraph GT3X+**, Pensacola, FL, USA) for seven days (24 hours protocol) on the right hip and the non-dominant wrist. Children were instructed to only remove accelerometers for water activities (e.g., shower, swimming, etc.), and always at the same time. Concomitantly, the participants (assisted by their parents when needed) reported the time they went to bed and woke-up in a diary log every day.

GGIR vignette (scan or click here)

ActiGraph GT3X+ is a triaxial accelerometer with a dynamic range of ±6 *g*'s. Both hip- and wrist-worn accelerometers were initialized to capture and store accelerations at 100 Hz. The raw accelerations were then downloaded and converted to ".csv" format using the ActiLife v.6.13.3 (ActiGraph, Pensacola, FL, USA). Raw ".csv" files were imported to R software (v. 3.1.2, https://www.cran.r-project.org/) and processed using the GGIR package (v. 1.5-12, https://cran.r-project.org/web/packages/GGIR/). Raw data were also processed in the ActiLife software to obtain activity counts in the vector magnitude (VMCounts, Euclidean Norm of the counts recorded in the three axes)

and in the vertical axis (VACounts) using the normal filter developed by ActiGraph. Step counts were also derived from the ActiLife software using the hip-worn accelerometer data. The processing methods in GGIR involved: (i) Auto-calibration of the data according to the local gravity [66]; (ii) detection of the non-wear time based on the raw acceleration of the three axes [67]. Briefly, each 15-min block was classified as non-wear time if the standard deviation (SD) of 2 out of the 3 axes was lower than 13 m*g* during the surrounding 60-min moving window, or if the value range for 2 out of the 3 axes was lower than 50 m*g*; (iii) detection of sustained abnormal high accelerations, i.e., higher than 5.5 *g*; (iv) calculation of the Euclidean Norm Minus One *G* (ENMO) with negative values rounded to zero; (v) importation of the VMCounts and VACounts calculated in the ActiLife software to the GGIR software; (vi) imputation of detected non-wear time and abnormal high accelerations by means of the acceleration for the rest of the recording period during the same time interval than the affected periods; (vii) identification of waking and sleeping hours using an automatized algorithm guided by the times reported by the participants [40]. Waking and sleeping hours were detected using data from the non-dominant wrist and detected times were then matched to the right hip data for each participant; and, (viii) Estimation of PA intensities and SB using different age-appropriate cut-points [61,68–70].

Abbreviations in this page:
DFT: design fluency test
ENMO: Euclidean Norm minus 1*G*
GMV: grey matter volume
MRI: magnetic resonance imaging
PA: physical activity
SD: standard deviation
TMT: trail making test
VACounts: activity counts in the vertical axis
VMCounts: activity counts in the vector magnitude

Brain health outcomes

The detailed methods of each outcome are defined in the results section. In brief, as indicators of brain health, the ActiveBrains project collected a complete array of outcomes. Magnetic resonance imaging (MRI) scans were used to assess the brain structure, specifically the measures of GMV and total brain volume used in this thesis, among other outcomes. All images were collected with a 3.0 Tesla Siemens Magnetom Tim Trio scanner (Siemens Medical Solutions, Erlangen, Germany) with a 32-channel. MRI is considered a gold-standard measure of the brain volumes.

Academic achievement was assessed with the Spanish version of the Woodcock-Johnson III battery, which is a valid and reliable (internal consistency reliability coefficient > 0.9) measure of academic achievement in children [71]. Reading, language, mathematics, and sciences skills were assessed in individual tests lasting 100-120 min per child. Executive functions include a set of cognitive processes that are necessary for the cognitive control of behaviour. In ActiveBrains, the executive function domains measured were cognitive flexibility, inhibition, and working memory. Cognitive flexibility was assessed with the second and fourth conditions of the design fluency test (DFT); and with the third and fourth conditions of the trail making test (TMT). Both the DFT and the TMT are valid and reliable for

Abbreviations in this page:
DFT: design fluency test
DNMS: delayed non-match-to-sample
DXA: dual-energy X-ray absorptiometry
GGT: γ-glutamyl transferase
HDL: high-density lipoprotein
IQ: intelligence quotient
K-BIT: Kaufman brief intelligence test
LDL: low-density lipoprotein
TMT: trail making test

measuring cognitive flexibility in children [72,73]. Inhibition was assessed with the Stroop test, which is a valid and reliable indicator [72–75]. Performance time for condition 3 (i.e., inhibiting reading by naming colour) minus condition 1 (i.e., colour naming) was used. To assess working memory, we used a modified version of the Delayed non-match-to-sample (DNMS) computerized task, which has been previously validated [76]. Response accuracy for the high-load condition was used. Finally, IQ was assessed with the Spanish version of the Kaufman Brief Intelligence Test (K-BIT), which has been previously validated (coefficient α of 0.86-0.93) [77]. Crystallized and fluid intelligence components were assessed and summed to obtain the overall IQ score.

Additionally, the psychological ill-being, well-being, and self-perceptions components of mental health were assessed with valid self-reported questionnaires. Psychological ill-being included measures of stress (Children's Daily Stress Inventory, scored from 0 to 30) [78], anxiety (State-Trait Anxiety Inventory for Children, scored from 20 to 60) [79], depression (Children's Depression Inventory, scored from 0 to 54) [80], and negative affect (Positive and Negative Affect Schedule for Children, scored from 10 to 30) [81]. Otherwise, psychological well-being included positive affect (Positive and Negative Affect Schedule for Children, scored from 10 to 30) [81], happiness (Subjective Happiness Scale, scored from 4 to 28) [82], and optimism (Life Orientation Test-Revised, scored from 6 to 30) [83]. And self-perceptions measures consisted of self-efficacy (General Self-Efficacy, scored from 10 to 40) [84], self-concept (Five-Factor Self-concept questionnaire, scored from 30 to 300) [85], and self-esteem (Rosenberg Self-Esteem Scale, scored from 10 to 40) [86].

Physical health outcomes

The cardiometabolic health was assessed via indicators of the four core components usually accepted in metabolic syndrome [87]. Abdominal obesity was represented by the waist circumference. Dyslipidaemia indicators included fasting LDL and HDL cholesterol, triglycerides and γ-glutamyl transferase (GGT), all derived from blood samples. Insulin resistance biomarkers were fasting insulin and glucose derived from the blood samples. Lastly, systolic and diastolic blood pressure were assessed in two different days and the lowest values were used in the analyses.

For body composition, body weight and height were measured with a scale and a stadiometer (SECA, Hamburg, Germany) with participants barefoot and wearing light underclothes. BMI was calculated as kg/m^2. Whole-body fat mass and lean mass were measured via dual-energy X-ray absorptiometry (DXA, Discovery Horizon® DXA system, Hologic, Canada ULC). Fat/lean mass indices were calculated as fat/lean mass in kilograms divided by height in meters

squared (kg/m^2). Visceral adipose tissue was also derived from the DXA measurement.

Physical fitness components (i.e., CRF, speed-agility, and muscular fitness) were assessed with the feasible, reliable, and valid tests for children included in the ALPHA fitness battery [88–90]. Specifically, CRF was assessed with the 20 m shuttle-run test. The number of completed laps and the VO_2max in ml/kg/min were recorded [91]. Speed-agility fitness was assessed with the 4x10 m shuttle run test (seconds to complete the test). Muscular fitness with the handgrip strength and the standing long jump tests [63].

Confounders

The main confounders used in this thesis are sex, biological maturation, and parental education. Biological maturation was assessed with the peak height velocity from height and sitting height measurements using the Moore's equations [92]. The peak height velocity provides a continuous, accurate, and discriminant measure of maturational status [92]. Peak height velocity represents the difference (in years) between the chronological age and the age at peak height velocity (calculated from standing and sitting height). Parental education level was categorized as both of them, one of them, or neither of them reached university-level education.

The MINISTOP study (Study III)

Study design

The Mobile-based intervention intended to stop obesity in preschoolers (MINISTOP) study is a RCT (NCT02021786) that aimed to evaluate the effectiveness of a mobile-phone-based intervention to promote better body composition, dietary habits, and PA in healthy preschool-aged children [93–95] Specifically, this thesis includes data from a nested study within this trial that was conducted to evaluate the capacity of wrist-worn accelerometers (ActiGraph GT3X+, Pensacola, FL, USA) to predict free-living PAEE. Ethical approval was obtained from the Research and Ethics Committee (Stockholm, Sweden) and both parents provided informed consent.

Inclusion and exclusion criteria

To be included, parents must: (i) have a four-year-old child and live in the county of Östergötland (Sweden); (ii) have the possibility to have their child measured at baseline at 4.5 years ± 2 months of age; and (iii) be able to speak and read Swedish sufficiently well (at least one parent). Children diagnosed with neurological or endocrine diseases and children who have a parent suffering from a serious physical or psychological disease were excluded from the study. The nested validation study included forty parents and their child.

Abbreviations in this page:
BMI: body mass index
CRF: cardiorespiratory fitness
DXA: dual-energy X-ray absorptiometry
PA: physical activity
PAEE: physical activity-related energy expenditure
RCT: randomized controlled trial
VO_2max: maximal oxygen consumption

Abbreviations in this page:
BMR: basal metabolic rate
ENMO: Euclidean Norm minus 1*G*
N_D: deuterium dilution space
N_O: ^{18}O dilution space
PA: physical activity
PAEE: physical activity-related energy expenditure
TEE: total energy expenditure
VMCounts: activity counts in the vector magnitude

Measures

Physical activity (explanatory variable)

PA was monitored with the GT3X+ accelerometer placed on the non-dominant wrist. Participants were instructed to wear the accelerometer for the first seven days of the 14-day doubly-labelled water period (24-hours per day). Devices were initialized to collect data at 50 Hz, as this sampling frequency is sufficient to capture wrist daily motion [96,97]. Auto-calibration of the acceleration signal, non-wear time detection and treatment was similar to the one performed in the ActiveBrains project (see section above). We derived several acceleration aggregation metrics from the ActiLife software (e.g., VMCounts) and from the GGIR software (e.g., ENMO). See more information on the acceleration metrics derived in Results, Study V.

Energy expenditure (outcome)

The children's TEE and total body water were measured with doubly-labelled water during 14 days. Each child was given an accurately weighed dose of stable isotopes using ^2H_2O (enrichment 99.9%) and $H_2{}^{18}O$ (enrichment 20%): 0.14 g ^2H_2O and 0.35 g $H_2{}^{18}O$ per kg of body weight. Urine samples were stored in glass vials with an internal aluminium-lined screw cap sealing at +4 ºC until sample collection was finished, after which they were stored at -20 ºC until analysis. ^2H and ^{18}O enrichments of dose and urine samples were analysed (both pre and post dosing) using a Finnigan MAT Delta Plus Isotope-Ratio Mass Spectrometer (ThermoFinnigan, Gothenburg, Sweden). The quotient between the ^2H dilution space (N_D) and the ^{18}O dilution space (N_O) was 1.039 ± 0.008 for the 39 children. CO_2 production was calculated according to the method by Davies et al. [98] assuming that 27.1% of the total water losses were fractionated. The Weir equation was applied to derive TEE from carbon dioxide production [99] assuming a food quotient of 0.85 [100]. Over the 14-day measurement period, no major change in body weight was observed (n=39; 0.07 ± 0.32 kg). We applied prediction equations based on weight [101] to estimate basal metabolic rate (BMR). Thereafter, PAEE was calculated as TEE multiplied by 0.9 minus BMR. This includes a reduction in TEE by 10% to adjust for energy expended due to dietary induced thermogenesis.

Anthropometry and body composition (confounders)

Body weight and height were measured with an electronic scale and stadiometer. Fat-free mass (kg) was calculated from total body water assuming that fat-free mass contains 76.4% water [102]. Fat mass (kg) was calculated as the difference between body weight and fat-free mass.

Pilot study on accelerometry (Study IV)

Study design

We recruited a convenience sample composed of students and research personnel from the University of Granada, Spain. The study was carried out in two waves of 45 (23 women, 18-41 years old) and 36 (10 women, 22-30 years old) young adults, respectively. Wave 1 data were used to compare different acceleration metrics across body attachment sites (i.e., right hip, dominant wrist, and non-dominant wrist). Wave 2 was used for cross-validation purposes and they only wore two accelerometers on dominant wrist and non-dominant wrist. This study was conducted according to the Declaration of Helsinki and approved by the Ethics Committee on Human Research of the University of Granada.

Abbreviations in this page:
BMR: basal metabolic rate
BMI: body mass index
ENMO: Euclidean Norm minus 1*G*
LFENMO: ENMO of the low-pass filtered raw accelerations
MAD: mean amplitude deviation
TEE: total energy expenditure
VMCounts: activity counts in the vector magnitude

Measures

Physical activity (accelerometer data collection)

Participants wore the ActiGraph GT3X+ recording accelerations at 100 Hz. Wave 1 participants wore three accelerometers: right hip and both wrists. Wave 2 participants wore two accelerometers, one on each wrist. Raw accelerations were then downloaded (".gt3x" files) and converted to ".csv" format using ActiLife v.6.13.3 (ActiGraph, Pensacola, FL, USA). All participants wore devices for seven complete days, and were instructed to remove them all together for shower and water-based activities (e.g., swimming). Participants were also encouraged to wear devices as much as possible, including sleeping periods.

Auto-calibration of the acceleration signal, non-wear time detection and treatment was similar to the one performed in the ActiveBrains project (see section above). We derived several acceleration metrics from the ActiLife software (e.g., VMCounts) and from the GGIR software (e.g., ENMO, LFENMO, MAD). See more information on the acceleration metrics derived in Results, Study IV. These acceleration metrics have been often used the PA studies.

Anthropometry (confounders)

Participants reported their sex and age upon signing the informed consent. Additionally, participants' body weight and height were measured to the nearest 0.1 kg and 0.1 cm using an electronic scale (SECA 861, Hamburg, Germany) and a precision stadiometer (SECA 225, Hamburg, Germany). BMI was calculated as kg/m^2.

Summary of methods by study

Table 2
Summary of the methods used in each study

	Study	Project	Design	N Age Sex	Explanatory or exposure variable/s	Explained or outcome variable/s
SECTION I	I	-	Systematic review	235 studies	-	
	II	-	Descriptive	-	-	
	III	MINISTOP	Cross-sectional	39 15.5±0.1 y/o 18 girls	PA	TEE PAEE
	IV	Pilot	Cross-sectional	78 25.8±3.9 y/o 33 women	PA	PA
	V	ActiveBrains	Cross-sectional	104 10.1±1.1 y/o 43 girls	PA SB	PA SB
	VI	ActiveBrains	Cross-sectional	105 10.1±1.1 y/o 43 girls	Steps	PA
	VII	-	Consensus	-	-	-
SECTION II	VIII	ActiveBrains	Cross-sectional	93 10.0±1.1 y/o 37 girls	PA, SB	GMV
	IX	ActiveBrains	Cross-sectional	95 10.0±1.1 y/o 37 girls	Activity-rest pattern	GMV Academic achievement Executive function IQ
	X	ActiveBrains	Cross-sectional	96 10.0±1.1 y/o 38 girls	Sleep	GMV Academic achievement Executive function IQ
	XI	ActiveBrains	RCT	98 10.0±1.1 y/o 41 girls	20-week PA program	Cardiometabolic health Body composition Physical fitness Psychological ill-being Psychological wellbeing Self-perceptions

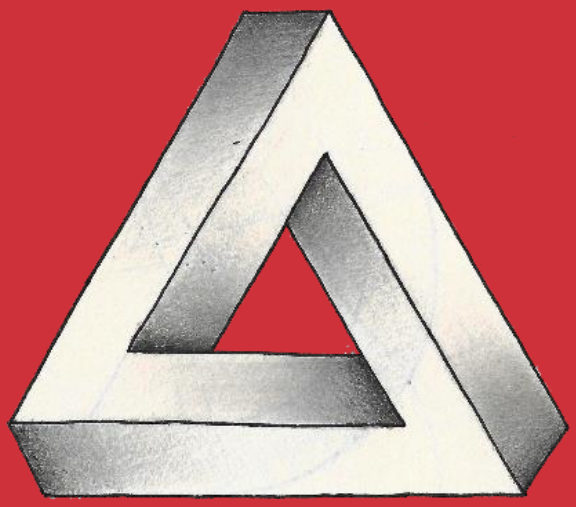
Abbreviations in this page:
GMV: grey matter volume
IQ: intelligence quotient
PA: physical activity
PAEE: physical activity-related energy expenditure
RCT: randomized controlled trial
SB: sedentary behaviour
TEE: total energy expenditure

RESULTS AND DISCUSSION

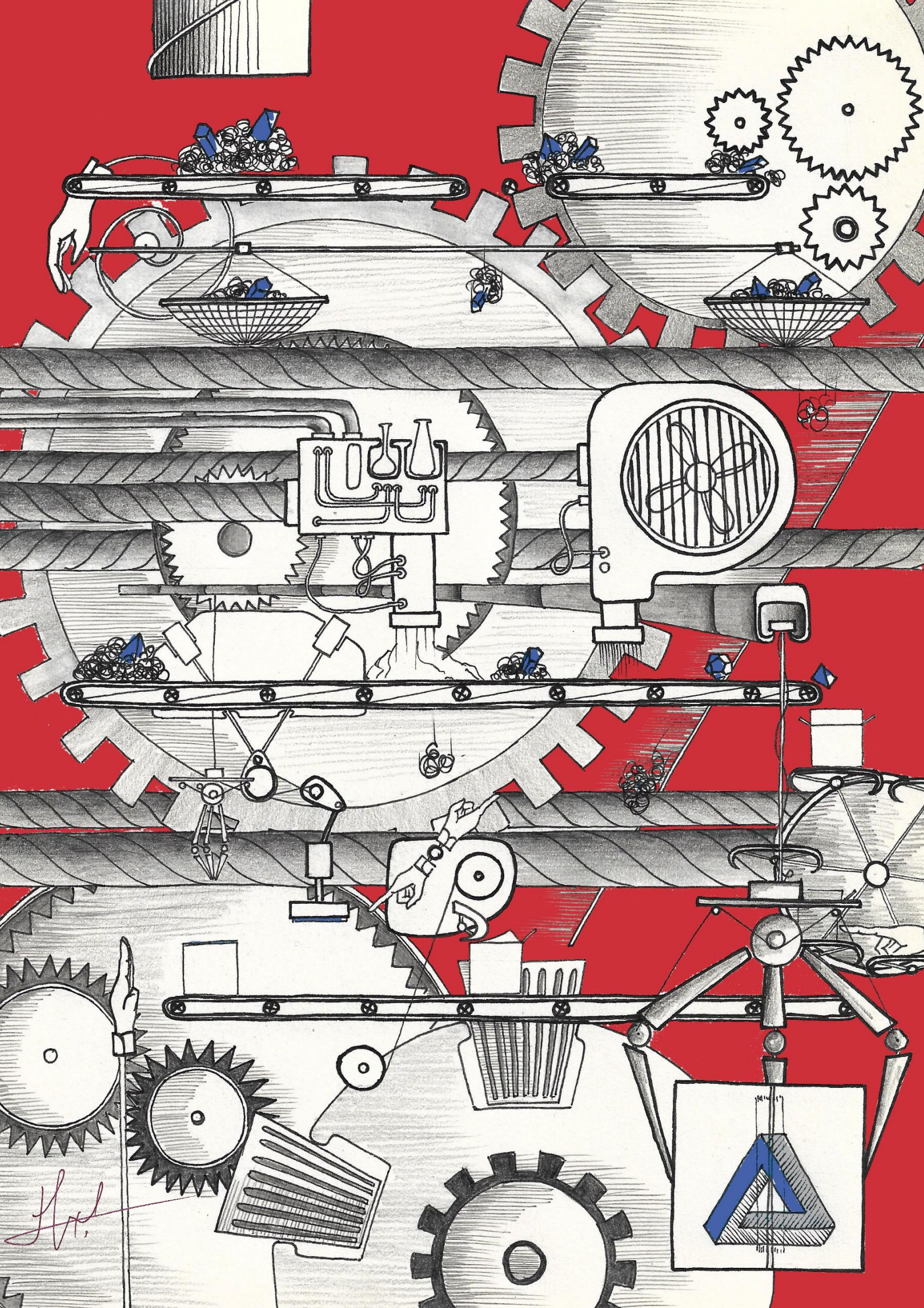
"I put my heart and soul into my work, and I have lost my mind in the process"

Vincent van Gogh

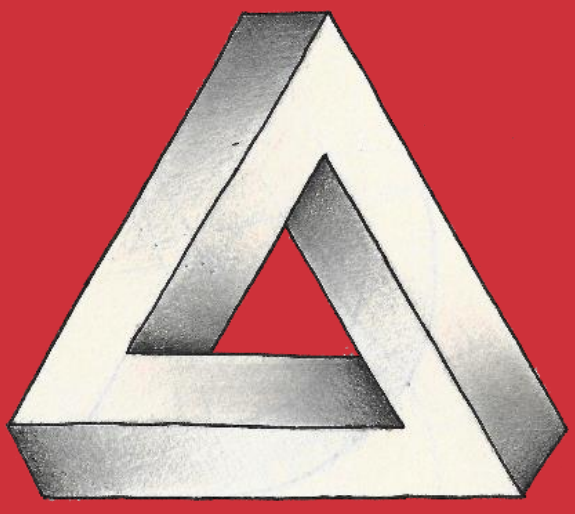
CONTENTS



SECTION I



Study I



Accelerometer data collection and processing criteria to assess physical activity and other outcomes: a systematic review and practical considerations

Migueles JH, Cadenas-Sanchez C, Ekelund U, Nyström CD, Mora-Gonzalez J, Löf M, Labayen I, Ruiz JR, Ortega FB.
Sports Med. 2017 Sep;47(9):1821-1845
DOI: 10.1007/s40279-017-0716-0

Contents

Abstract

Background | Accelerometers are widely used to measure SB, PA, PAEE, and sleep-related behaviours, with the ActiGraph being the most frequently used brand by researchers. However, data collection and processing criteria have evolved in a myriad of ways out of the need to answer unique research questions; as a result, there is no consensus.

Aims | The purpose of this review was to: (i) compile and classify existing studies assessing SB, PA, PAEE, or sleep using the ActiGraph GT3X/+ through data collection and processing criteria to improve data comparability; and (ii) review data collection and processing criteria when using GT3X/+ and provide age-specific practical considerations based on the validation/calibration studies identified.

Methods | Two independent researchers conducted the search in PubMed and Web of Science. We included all original studies in which the GT3X/+ was used in laboratory, controlled, or free-living conditions published from 1 January 2010 to the 31 December 2015.

Results | The present systematic review provides key information about the following data collection and processing criteria: placement, sampling frequency, filter, epoch length, non-wear time, what constitutes a valid day and a valid week, cut-points for SB and PA intensity classification, and algorithms to estimate PAEE and sleep-related behaviours. The information is organized by age group, since criteria are usually age-specific.

Conclusion | This review will help researchers and practitioners to make better decisions before (i.e., device placement and sampling frequency) and after (i.e., data processing criteria) data collection using the GT3X/+ accelerometer, in order to obtain more valid and comparable data.

PROSPERO registration number: CRD42016039991

Abbreviations in this page:
PA: physical activity
PAEE: physical activity-related energy expenditure
SB: sedentary behavior

Key Points

Question
What accelerometer data collection and processing criteria are recommendable to estimate SB and PA outcomes in each age group?

Findings
High discrepancies and lack of harmonization was noted. Practical considerations for every step of the accelerometer data collection and processing are provided based on existing literature.

Meaning
The tabulated data generated will facilitate comparisons between studies using the ActiGraph GT3X and aid in the selection of the most appropriate method to use for each specific research purpose.

Introduction

Health benefits of PA across a person's lifespan have been widely reported [3,4,103]. The use of accelerometers to assess time in SB and PA [104–107] has become an objective and feasible alternative to self-report methods such as questionnaires, which are characterized by their poor reliability and validity, especially in younger populations [30–32]. Accelerometers are wearable devices which measure accelerations of the body segment to which the monitor is attached. The signal is usually filtered and pre-processed by the monitor to obtain activity counts, i.e., accelerations due to body movement. The amount and intensity of daily time in SB and PA, may be obtained by classifying activity counts accumulated in a specific time interval (epoch length) with a set of cut-points, i.e., intensity thresholds for PA intensity classification [37,68,70,108,109]. PAEE or sleep-related behaviours may also be estimated by applying algorithms to objectively-determined activity counts [38,39,110–113]. New methods to estimate these variables from raw acceleration signals (gravity units) instead of activity counts have been developed recently [61,114,115].

Abbreviations in this page:
PA: physical activity
PAEE: physical activity-related energy expenditure
SB: sedentary behavior
VMCounts: activity counts in the vector magnitude

Among the commercially available brands, the ActiGraph (Pensacola, FL, USA) accelerometers are the most frequently used by researchers, accounting for >50% of published studies [41]. This review only considered the latest generation of ActiGraph devices, i.e., GT3X, GT3X+ and wGT3X-BT (hereinafter referred to as GT3X/+). The continuous change in the features of these devices makes it difficult to compare data between studies.

The first ActiGraph accelerometers available were uniaxial (i.e., they could only detect vertical axis [VA] accelerations) and consequently cut-points and algorithms were developed to assess time in SB, PA intensity, PAEE and sleep-related behaviours from VA accelerations [37,39,113]. In mid-2009, ActiGraph released the triaxial GT3X which detected accelerations in three axes (i.e., vertical, mediolateral and antero-posterior axes). The transition from uniaxial to triaxial devices implied new calibration processes, and the algorithms developed for the VA were not applicable to VMCounts (i.e., the square root of the sum of squared activity counts from the 3 axes) [107,108,110,112,116–118].

Due to the extremely fast development in this field, there is an overwhelming amount of data collection and processing criteria decisions, and there is no consensus about which approaches to use. Consequently, it is difficult for researchers and practitioners to make the right decisions about which criteria should be used in each situation. This is important as the chosen criteria have a huge impact on the outcome. In order to address this problem, some studies have

Abbreviations in this page:
PA: physical activity
PAEE: physical activity-related energy expenditure
SB: sedentary behavior

compared certain GT3X/+ outcomes estimated by different cut-points and algorithms [104,119–121] in an attempt to recommend which decisions are the most accurate, yet this information is still scarce.

It is important to note that algorithms validated in a specific age group, might not be valid for other age groups due to different PA patterns, so whenever possible, data collection and processing criteria should be age-specific. Accelerometer methods can be grouped into two categories: data collection protocols, which are decisions that need to be made *a priori* such as device placement or sampling frequency; and data processing criteria, which involves decisions that can be made *a posteriori* such as filters, epoch length, non-wear time definition, cut-points and algorithms. The present review will address all of these criteria separately and specifically by age group. In this review we aimed to: 1) compile and classify existing studies assessing SB, PA, PAEE, or sleep using the ActiGraph GT3X/+ by data collection and processing criteria to improve data comparability and 2) review data collection and processing criteria when using GT3X/+ and provide age-specific practical considerations based on the validation/calibration studies identified. Both objectives were approached separately for the following age groups: pre-schoolers, children/adolescents, adults and older adults. Although there is a large amount of information included in this review, we believe that it is useful for readers to have a single article that summarizes the most important accelerometer methods for each age group separately. This will allow readers to go directly to a specific criterion for the age group they are interested in, (e.g., PAEE in pre-schoolers). In this review, we provide a section with examples of how the information presented can be used in practical terms, as well as a table with practical considerations.

Methods

Study design

The present review focuses on 11 key methodological issues related to GT3X/+ data collection and processing criteria: 1) device placement, 2) sampling frequency, 3) filter, 4) epoch length, 5) non-wear time definition, 6) what constitutes a valid day and valid week, 7) registration period, 8) time in SB and PA intensity classification, 9) PAEE algorithms, 10) sleep algorithms, and 11) step counting. Available information was classified into two different types of studies: 1) any cross-sectional, longitudinal or intervention study, which used the GT3X/+ device and met the inclusion criteria indicated below (objective 1); and 2) studies focused on validation, calibration or comparison of functions related to data collection or processing criteria (objective 2). Therefore, the practical considerations provided

for each age group are based on the results from the validation/calibration studies (see **Table 3**), while the rest of studies were only used to describe the most-frequently used decisions.

Table 3
Summary of practical considerations by age group

Age group / criterion	Pre-schoolers	Children and adolescents	Adults	Older adults
Placement	Hip[a] and wrist	Hip[a] and wrist	Hip[a] and wrist	Hip[a] and wrist
Sampling frequency	90-100 Hz	90-100 Hz	90-100 Hz	90-100 Hz
Filter[b]	Normal	Normal	Normal	LFE
Epoch length[b]	1-15 s	1-15 s	60 s[c]	60 s[c]
Non-wear time definition	Not clear[c]	Not clear[c]	Not clear[c]	Choi et al. [122] algorithm
Valid day[d]	≥ 10 h	≥ 10 h	≥ 10 h	≥ 10 h
Valid week	≥ 4 days	≥ 4 days	≥ 4 days	≥ 4 days
Registration period	24 h	24 h	24 h	24 h
SB/PA intensity classification[e, f]				
Dominant wrist	No data found	Crouter et al. [123]	Staudenmayer et al. [115]	No data found
Non-dominant wrist	Johansson et al. [124] (2-3 y)	Hildebrand et al. [61] Chandler et al. [70]	Hildebrand et al. [61]	No data found
Hip	Costa et al. [125] (2-3 y) Jimmy et al. [126] (4-6 y)	Hänggi et al. [108] (7-11 y) Romanzini et al. [69] (12-19 y)	Sasaki et al. [107]	Aguilar-Farias et al. [116] Santos-Lozano et al. [117]
PAEE algorithm[e]				
Non-dominant wrist	No data found	Hildebrand et al. [61]	Ellis et al. [127]	No data found
Hip	Butte et al. [128] (2-3 y)	Crouter et al. [112] (7-11 y)	Hildebrand et al. [61]	Santos-Lozano et al. [117]
Sleep algorithm[e]	No data found	Sadeh et al. [38]	Sadeh et al. [38] (20-30 y) Cole et al. [39] (> 30 y)	Cole et al. [39]

Note: These recommendations should be considered with caution. We strongly recommend reading section 4 for an understanding of the specific considerations for each age group
[a] There are no algorithms currently available to estimate sleep-related behaviours from data obtained from hip-worn devices
[b] Criterion that could highly affect the output. In these cases, when estimations of PA, PAEE or sleep are the variables of interest, the same criterion as selected in the validation study is recommended. If acceleration metrics are the variables of interest (e.g., counts), the recommendation is provided in this table.

Abbreviations in this page:
ESM: electronic supplementary material
LFE: low-frequency extension
PA: physical activity
PAEE: physical activity-related energy expenditure
PRISMA: preferred reporting items for systematic reviews and meta-analyses
SB: sedentary behaviour

Furthermore, we provide a summary of all data extracted from the validation/calibration papers included in this review by age group in **Electronic Supplementary Material (ESM) S1**. Inclusion and exclusion criteria and analytical methods were specified in advance and registered in the PROSPERO international database of systematic reviews (CRD42016039991) [129]. The study is conducted according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) Statement [130].

ESM 1 (scan or click here)

Search strategy

We searched PubMed and Web of Science, for studies using the ActiGraph GT3X/+ model and classified the studies into the following age groups: pre-schoolers (2-5 years), children (6-11 years), adolescents (12-18 years), adults (19-59 years) and older adults (≥60 years). We combined (using the Boolean operator "OR") the following search terms: GT3X, GT3X+ and ActiGraph. Although, we wanted

to limit the search to GT3X/+, the word ActiGraph was entered in the search because we found that some studies specified the brand (i.e., ActiGraph) instead of the model (i.e., GT3X/+) in the title/abstract/keywords. Since the GT3X/+ models were launched in mid-2009, we limited the dates of the search from the 1st January 2010 to the 31st December 2015 and conducted the final search on January 3rd 2016. We contacted authors of those studies where the data processing and collection information was unavailable in the published article. In a final step, we extended the search to the IEEE (Institute of Electrical and Electronics Engineers) Xplore database, in case we had missed any relevant studies.

Inclusion and exclusion criteria

We included all original studies (cross-sectional, longitudinal or intervention studies) in which the GT3X/+ was used in a laboratory, or under controlled or free-living conditions. Protocol studies, reviews, editorials and abstract or congress communications were excluded, as well as studies conducted in people with mobility problems or in periods of life in which mobility could have been markedly altered (e.g., pregnancy).

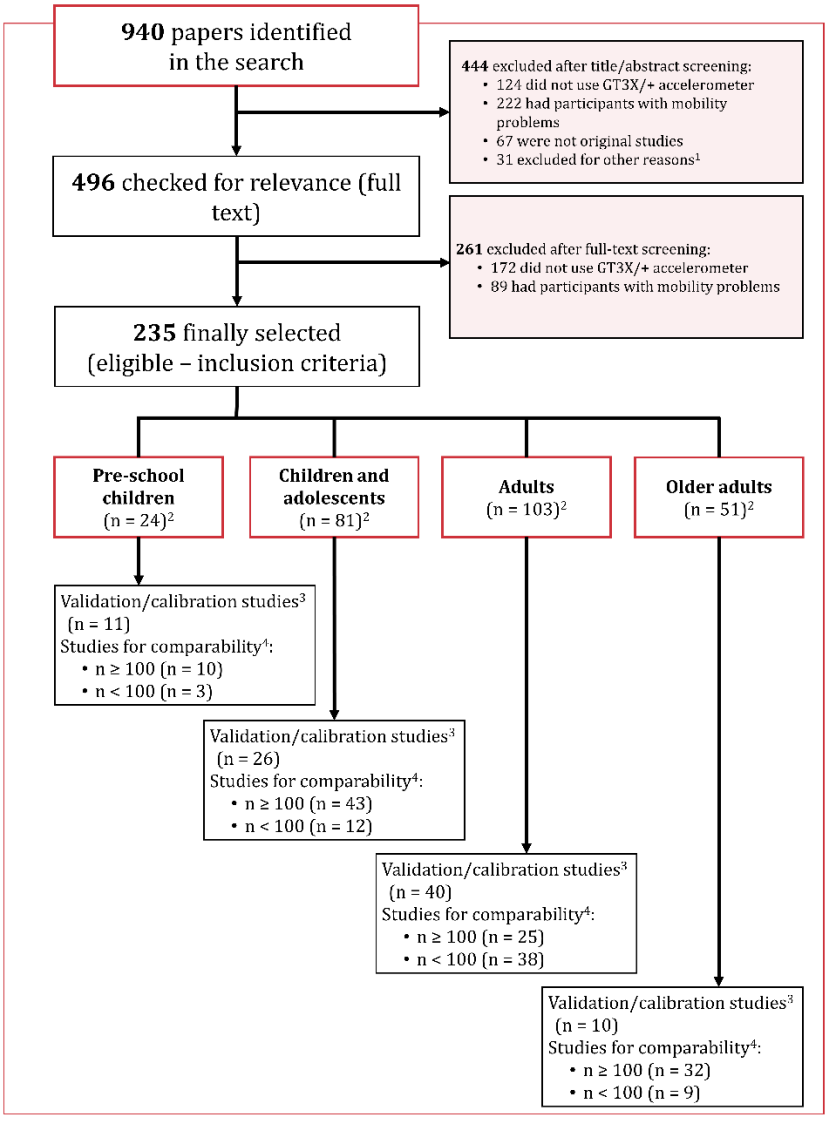


Figure 8
Flowchart of the literature search and study selection process.
[1]Studies using accelerometers of other purposes (e.g., accelerometers attached to dogs).
[2]Studies that included two age ranges were counted in both age groups.
[3]Studies focused on validation, calibration or comparison of functions related to data collection or processing criteria.
[4]All cross-sectional, longitudinal, or intervention studies, which used the GT3X/+ device and met the inclusion criteria.

Two authors (JHM and CCS) independently read the articles and checked whether they met the inclusion/exclusion criteria. They obtained 76% agreement on the papers selected for the review before consensus and 100% agreement after discrepancies were resolved in a consensus meeting. Risk of bias assessment was also conducted independently by JHM and CCS in order to assess the quality of studies (see **ESM 2**).

Abbreviations in this page:
ESM: electronic supplementary material
LFE: Low-frequency extension filter

Results

A total of 940 articles were identified (**Figure 8**), of which 444 were excluded after reading the title and abstract and 261 articles were additionally excluded after reading the full text and did not meet the inclusion/exclusion criteria stated above. Finally, a total of 235 studies were considered eligible for the current systematic review. Of them, 78 were validation/calibration studies. Methods and results of these validation/calibration studies are summarized in **ESM 1**. Detailed information about the methods and results for the rest of studies (i.e., those using GT3X/+ that were not validation/calibration studies) included in this review is available upon request.

ESM 2 (scan or click here)

Table 4
Summary of the criteria used for data collection protocols and data processing from articles reviewed by age group (see **ESM 1** for the criteria used by each of the studies listed in this table)

Reference	Pre-schoolers (n=24) n (%)	Children and adolescents (n=81) n (%)	Adults (n=103) n (%)	Older adults (n=51) n (%)
Placement				
Hip	22 (92)	73 (90)	87 (84)	44 (86)
Non-dominant wrist	2 (8)	6 (7)	8 (8)	5 (10)
Dominant wrist	0 (0)	1 (1)	6 (6)	5 (10)
Others[a]	0 (0)	2 (2)	21 (20)	2 (4)
Not reported	0 (0)	5 (6)	6 (6)	0 (0)
Sampling Frequency				
30 Hz.	16 (67)	53 (65)	70 (68)	39 (76)
40 Hz.	0 (0)	0 (0)	2 (2)	0 (0)
50 Hz.	0 (0)	0 (0)	2 (2)	0 (0)
60 Hz.	2 (8)	2 (2)	6 (6)	2 (4)
70 Hz.	0 (0)	0 (0)	1 (1)	0 (0)
80 Hz.	1 (4)	6 (7)	9 (9)	5 (10)
90 Hz.	0 (0)	0 (0)	2 (2)	0 (0)
100 Hz.	1 (4)	4 (5)	6 (6)	1 (2)
Not reported	4 (17)	18 (23)	15 (15)	5 (10)
Filter				
Normal	8 (34)	14 (17)	25 (24)	6 (12)
LFE	2 (8)	11 (14)	15 (15)	6 (12)
Not reported	14 (58)	53 (65)	67 (65)	40 (80)
Epoch length				
1 s	1 (4)	8 (10)	15 (15)	6 (12)
2 s	0 (0)	1 (1)	2 (2)	0 (0)
3 s	0 (0)	1 (1)	0 (0)	0 (0)
5 s	6 (25)	6 (7)	1 (1)	0 (0)
10 s	0 (0)	16 (20)	6 (6)	1 (2)
15 s	13 (54)	28 (35)	8 (8)	3 (6)
20 s	0 (0)	0 (0)	0 (0)	1 (2)
30 s	1 (4)	3 (4)	0 (0)	0 (0)
45 s	0 (0)	1 (1)	0 (0)	0 (0)
60 s	6 (25)	17 (21)	52 (50)	38 (74)
Not reported	0 (0)	5 (6)	16 (16)	3 (6)
Non-wear time definition[b]				
10-0-0	3 (13)	6 (7)	7 (7)	4 (8)
20-0-0	3 (13)	21 (26)	3 (3)	0 (0)
20-0-2	0 (0)	0 (0)	1 (1)	0 (0)
30-0-0	2 (8)	5 (6)	1 (1)	0 (0)
30-0-1	0 (0)	3 (4)	0 (0)	0 (0)
60-0-0	3 (13)	5 (6)	16 (16)	3 (6)
60-0-2	0 (0)	7 (9)	15 (15)	12 (24)

ESM 1 (scan or click here)

Table 4 - continued
Summary of the criteria used for data collection protocols and data processing from articles reviewed by age group (see **ESM 1** for the criteria used by each of the studies listed in this table)

Abbreviations in this page:
CPM: counts per minute
ESM: electronic supplementary material
PA: physical activity
VA: vertical axis
VM: vector magnitude

Reference	Pre-schoolers (n=24) n (%)	Children and adolescents (n=81) n (%)	Adults (n=103) n (%)	Older adults (n=51) n (%)
60-30-2	0 (0)	0 (0)	0 (0)	1 (2)
90-0-0	0 (0)	1 (1)	5 (5)	1 (2)
90-0-2	0 (0)	0 (0)	3 (3)	2 (4)
90-30-2	0 (0)	0 (0)	2 (2)	14 (27)
120-0-0	0 (0)	0 (0)	1 (1)	0 (0)
180-0-0	0 (0)	0 (0)	0 (0)	1 (2)
Not reported	6 (25)	7 (9)	15 (15)	6 (12)
Cut-points for SB (CPM and vector used) [original reference]				
25 CPM VA [116]	0 (0)	0 (0)	1 (1)	1 (2)
50 CPM VA [131]	0 (0)	0 (0)	1 (1)	0 (0)
60 CPM VA [125]	1 (4)	0 (0)	0 (0)	0 (0)
100 CPM VA [37,68,109,132–138]	4 (17)	31 (38)	40 (39)	31 (61)
100 CPM VM	0 (0)	1 (1)	1 (1)	0 (0)
120 CPM VM [108]	0 (0)	2 (2)	0 (0)	0 (0)
148 CPM VA [139]	3 (13)	1 (1)	0 (0)	0 (0)
150 CPM VA [140]	0 (0)	2 (2)	5 (5)	1 (2)
150 CPM VM [131]	0 (0)	1 (1)	2 (2)	1 (2)
184 CPM VA [69]	0 (0)	1 (1)	0 (0)	0 (0)
200 CPM VA [131,141]	0 (0)	0 (0)	1 (1)	1 (2)
200 CPM VM [116]	0 (0)	0 (0)	1 (1)	2 (4)
240 CPM VA [128]	1 (4)	0 (0)	0 (0)	0 (0)
250 CPM VA [131]	0 (0)	0 (0)	1 (1)	0 (0)
274 CPM VA [142]	2 (8)	0 (0)	0 (0)	0 (0)
384 CPM VM [125]	1 (4)	0 (0)	0 (0)	0 (0)
500 CPM VA [143]	0 (0)	1 (1)	0 (0)	0 (0)
720 CPM VM [69]	0 (0)	1 (1)	0 (0)	0 (0)
796 CPM VA [144]	2 (8)	1 (1)	0 (0)	0 (0)
820 CPM VM [128]	1 (4)	0 (0)	0 (0)	0 (0)
1068 CPM VA [124]	1 (4)	0 (0)	0 (0)	0 (0)
1204 CPM VA [145]	2 (8)	0 (0)	0 (0)	0 (0)
1260 CPM VM [123]	0 (0)	1 (1)	0 (0)	0 (0)
1452 CPM VA [145]	2 (8)	0 (0)	0 (0)	0 (0)
1488 CPM VA [146]	2 (8)	0 (0)	0 (0)	0 (0)
1592 CPM VA [145]	3 (13)	0 (0)	0 (0)	0 (0)
1932 CPM VM [70]	0 (0)	1 (1)	0 (0)	0 (0)
2652 CPM VM [124]	2 (8)	0 (0)	0 (0)	0 (0)
3300 CPM VM [123]	0 (0)	1 (1)	0 (0)	0 (0)
3660 CPM VM [70]	0 (0)	1 (1)	0 (0)	0 (0)
Cut-points for PA intensity classification [original reference]				
Aguilar-Farías et al. [116]	0 (0)	0 (0)	1 (1)	3 (6)
Aittasalo et al. [147]	0 (0)	1 (1)	0 (0)	0 (0)
Andersen et al. [143]	0 (0)	1 (1)	0 (0)	0 (0)
Butte et al. [128]	1 (4)	0 (0)	0 (0)	0 (0)
Chandler et al. [70]	0 (0)	1 (1)	0 (0)	0 (0)
Copeland et al. [109]	0 (0)	0 (0)	0 (0)	9 (18)
Costa et al. [125]	1 (4)	0 (0)	0 (0)	0 (0)
Crouter et al. [123]	0 (0)	1 (1)	0 (0)	0 (0)
Davis et al. [141]	0 (0)	0 (0)	0 (0)	1 (2)
Evenson et al. [68]	8 (34)	36 (45)	0 (0)	0 (0)
Freedson et al. [37]	0 (0)	1 (1)	30 (29)	14 (27)
Freedson et al. [140]	0 (0)	8 (10)	0 (0)	0 (0)
Grydeland et al. [148]	0 (0)	1 (1)	0 (0)	0 (0)
Hänggi et al. [108]	0 (0)	2 (2)	0 (0)	0 (0)
Hildebrand et al. [61]	0 (0)	1 (1)	0 (0)	0 (0)
Jimmy et al. [126]	1 (4)	1 (1)	0 (0)	0 (0)
Johansson et al. [124]	2 (8)	0 (0)	0 (0)	0 (0)
Matthews et al. [149]	0 (0)	1 (1)	6 (6)	8 (16)
Mattocks et al. [150]	0 (0)	1 (1)	0 (0)	0 (0)
Metzger et al. [138]	0 (0)	0 (0)	3 (3)	1 (2)
Pate et al. [139]	4 (17)	1 (1)	0 (0)	0 (0)
Pruitt et al. [151]	0 (0)	0 (0)	0 (0)	1 (2)
Pulakka et al. [152]	2 (8)	0 (0)	0 (0)	0 (0)
Puyau et al. [144]	2 (8)	2 (2)	0 (0)	0 (0)
Reilly et al. [142]	3 (13)	0 (0)	0 (0)	0 (0)
Romanzini et al. [69]	0 (0)	1 (1)	0 (0)	0 (0)
Santos-Lozano et al. [117]	0 (0)	1 (1)	1 (1)	1 (2)
Sasaki et al. [107]	0 (0)	0 (0)	6 (6)	2 (4)
Sirard et al. [145]	3 (13)	0 (0)	0 (0)	0 (0)
Treuth et al. [134]	0 (0)	4 (5)	0 (0)	0 (0)
Troiano et al. [137]	0 (0)	0 (0)	8 (8)	4 (8)
Trost et al. [113]	0 (0)	2 (2)	0 (0)	0 (0)
Vähä-Ypyä et al. [153]	0 (0)	0 (0)	1 (1)	0 (0)

Table 4 – continued
Summary of the criteria used for data collection protocols and data processing from articles reviewed by age group (see **ESM 1** for the criteria used by each of the studies listed in this table)

Reference	Pre-schoolers (n=24) n (%)	Children and adolescents (n=81) n (%)	Adults (n=103) n (%)	Older adults (n=51) n (%)
Van Cauwenberghe et al. [146]	3 (13)	0 (0)	0 (0)	0 (0)
Vanhelst et al. [154]	0 (0)	1 (1)	0 (0)	0 (0)
Zhu et al. [155]	0 (0)	2 (2)	0 (0)	0 (0)
Zisko et al. [156]	0 (0)	0 (0)	0 (0)	1 (2)
PAEE algorithms [original reference]				
Butte et al. [128]	1 (1)	0 (0)	0 (0)	0 (0)
Crouter et al. [157]	0 (0)	1 (1)	0 (0)	0 (0)
Crouter et al. [158]	0 (0)	0 (0)	2 (2)	0 (0)
Crouter et al. [112]	0 (0)	3 (4)	0 (0)	0 (0)
Ellis et al. [118]	0 (0)	0 (0)	1 (1)	0 (0)
Evenson et al. [68]	0 (0)	1 (1)	0 (0)	0 (0)
Freedson et al. [140]	0 (0)	3 (4)	0 (0)	0 (0)
Hildebrand et al. [61]	0 (0)	1 (1)	0 (0)	0 (0)
Liu et al. [159]	0 (0)	1 (1)	0 (0)	0 (0)
Mattocks et al. [150]	0 (0)	1 (1)	0 (0)	0 (0)
Pate et al. [139]	1 (4)	0 (0)	0 (0)	0 (0)
Puyau et al. [144]	1 (4)	3 (4)	0 (0)	0 (0)
Santos-Lozano et al. [117]	0 (0)	1 (1)	1 (1)	1 (2)
Schmitz et al. [160]	0 (0)	1 (1)	0 (0)	0 (0)
Stec et al. [161]	0 (0)	0 (0)	1 (1)	0 (0)
Treuth et al. [134]	0 (0)	3 (4)	0 (0)	0 (0)
Trost et al. [113]	0 (0)	3 (4)	0 (0)	0 (0)
WET	0 (0)	1 (1)	1 (1)	1 (2)
WET + Freedson et al. [37]	0 (0)	1 (1)	3 (3)	2 (4)
WET + Sasaki et al. [107]	0 (0)	1 (1)	1 (1)	1 (2)
Zakeri et al. [162]	1 (4)	0 (0)	0 (0)	0 (0)
Zhu et al. [163]	0 (0)	1 (1)	0 (0)	0 (0)
Sleep algorithm [original reference]				
Barreira et al. [111]	0 (0)	1 (1)	0 (0)	0 (0)
Cole-Kripke et al. [39]	0 (0)	1 (1)	1 (1)	3 (6)
Sadeh et al. [38]	1 (4)	2 (2)	4 (4)	0 (0)
Tudor-Locke et al. [110]	0 (0)	4 (5)	0 (0)	0 (0)

Note 1: Studies using several criteria have been considered in each criterion, thus, in these cases percentages do not have to sum to 100%.
Note 2: Criteria used for the data collection and processing validated with other devices but have been applied to GT3X/+ data have been also considered.
[a] Other placements used for different aims from physical PA classification, PAEE, or sleep estimation (e.g., PA type identification, light sensor validation).
[b] Non-wear-time definition expressed as: minimum minutes of 0 CPM – minimum minutes for before and after the allowance windows – maximum of minutes of allowance.

Abbreviations in this page:
CPM: counts per minute
ESM: electronic supplementary material
METs: metabolic equivalents
PAEE: physical activity-related energy expenditure
WET: work energy theory

Forty-four percent (N=103) of the included studies were conducted in adults (46% validation/calibration studies); 34% (N=81) in youth (30% validation/calibration studies); 22% (N=51) in older adults (11% validation/calibration studies); and 10% (N=24) in preschoolers (13% validation/calibration studies).

Studies including two or more age groups are summarized in both age group sections in this review. **Table 4** presents the criteria used for data collection and processing by age group. A list of references for each of the criteria is found in **ESM 3**. The information provided in **Table 4** and **ESM 3** allows researchers to make comparisons between studies that have used the same data collection and processing criteria.

Figure 9 shows the percentage of studies that did not report key methodological issues by age group. Fifteen to twenty percent of the studies reviewed did not report criteria such as sampling frequency, epoch length and a non-wear time definition, and 60-80% of studies did not report information on the filter used.

ESM 3 (scan or click here)

Table 5
Summary of studies comparing hip- and wrist-worn GT3X/+ accelerometers

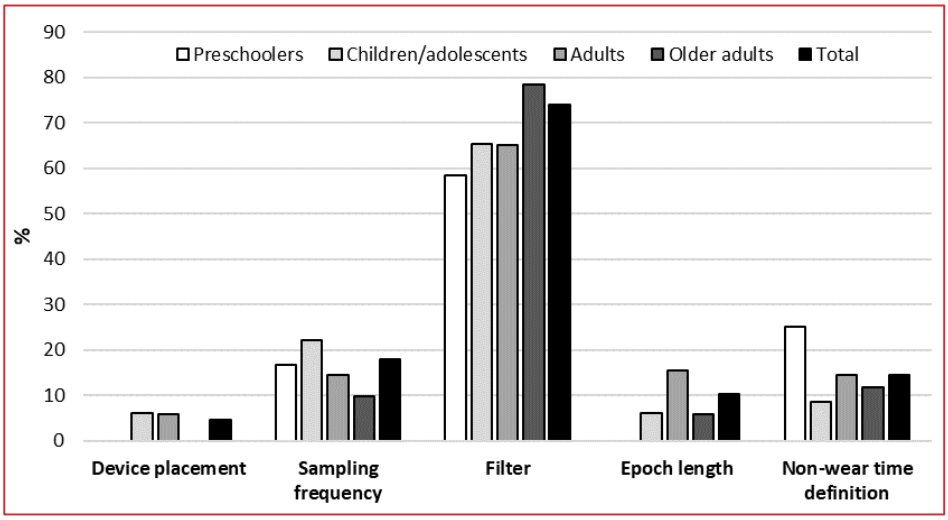
References	Age group	Aims (principal outcomes studied in italics)	Main findings/conclusions
Ellis et al. [118]	Adults	To compare GT3X/+ worn on right hip and non-dominant wrist, and the added value of heart rate data, for predicting *PA type* and *PAEE estimation*	In estimating PAEE, both device positions produced comparable results. The wrist GT3X/+ was superior predicting activities with significant arm movement, while the hip GT3X/+ was superior for predicting locomotion
Fairclough et al. [114]	Children	To compare right hip and non-dominant wrist *compliance*, and to compare *PA* derived from wrist and hip raw data	Wrist placement was associated with superior compliance compared with the hip. Raw accelerations were significantly higher for the wrist compared with the hip
Hildebrand et al. [61]	Children and adults	To compare raw GT3X/+ output from right hip and non-dominant wrist and to develop *PAEE* equations for each placement	The output from the wrist monitor was higher during more intense activities but similar or lower during SB activities. Hip PAEE equation showed a higher accuracy.
Hjorth et al. [164]	Children	To compare GT3X/+ *sleep scoring* from hip and non-dominant wrist with existing algorithms	Hip-worn and wrist-worn GT3X/+ cannot be used interchangeably for estimating sleep-related behaviours
Ozemek et al. [165]	Adults	To test the reliability of GT3X+ placed on the hip, dominant wrist and ankle in measuring *activity counts* recorded by axis 1, 2, 3 and VM during daily living	GT3X/+ worn on the hip, wrist and ankle showed a high test-retest agreement across all axes and VM. Lower variability in activity counts was observed in hip placement compared to wrist- or ankle-worn accelerometers
Slater et al. [166]	Adults	To examine the GT3X/+ validity for *sleep scoring* from right hip and left wrist compared to polysomnography using the same algorithm	The wrist-worn GT3X+ provided more valid measures of sleep but with only moderate capability to detect periods of wake during the sleep period. With Sadeh's algorithm[38] GTX3+ Actigraph worn on the hip does not provide valid or accurate measures of sleep
Staudenmayer et al. [115]	Adults	To develop algorithms for dominant wrist to estimate: *METs-hours, minutes in PA, minutes in SB vs not, and minutes in locomotion vs not*, validate them against indirect calorimetry and compare them against previous algorithms	The wrist models, applied to 15 s epoch, estimated METs better than a previously developed model that used CPM measured at the hip
Stec et al. [161]	Adults	To estimate the optimal placement (right wrist, right hip or right ankle) to attach the GT3X/+ for *PAEE estimation* during resistance exercise	The hip-worn GT3X obtained better results for estimating PAEE in resistance exercise
Tudor-Locke et al. [167]	Adults	To compare GT3X/+ *step outputs* obtained from right hip and non-dominant wrist	In lab conditions, hip detected more steps than wrist independently of the filter. In free-living, wrist produced higher step counts than hip. Hip step counts were more accurate than wrist in controlled conditions.

Abbreviations in this page:
CPM: counts per minute
ESM: electronic supplementary material
METs: metabolic equivalents
PA: physical activity
PAEE: physical activity-related energy expenditure
SB: sedentary behaviour
VM: vector magnitude

Table 5 presents the studies that have compared the differences in several outcomes when the GT3X/+ device was simultaneously worn on the wrist and hip. The optimal place to attach the GT3X/+ should be chosen based on reliability, validity and compliance. **Table 6** shows the references for the studies sorted by age group and placement site that have developed time in SB and PA cut-points, PAEE prediction equations, and sleep algorithms. **Table 7** shows the intensity cut-points used in the included studies together with the pre-processing criteria used in the study which developed each set of cut-points. Therefore, the practical considerations provided for each age group are based on the results from the validation/calibration studies (see **Table 3**).

In the following sub-sections, we will focus only on information from validation/calibration studies presented in **ESM 1**. Device placement, sampling frequency, and valid day and valid week correspond to data collection protocols (i.e., pre-processing stage) and the remaining criteria correspond to processing criteria (i.e., processing stage).

ESM 1 (scan or click here)



Abbreviations in this page:
PA: physical activity
ROC-AUC: receiver operating characteristic area under the curve
SB: sedentary behaviour

Figure 9
Percentage of the 235 included papers that did not report key methodological issues, separated by age group.

Device placement

Pre-schoolers

In young pre-schoolers Johansson et al. [124] reported receiver operating characteristic area under the curve (ROC-AUC) data for intensity thresholds between 0.88 to 0.98 using a left wrist mounted GT3X+. Similarly, a ROC-AUC of 0.90-0.94 was reported by Costa et al. [125] using a hip placement, suggesting high potential for both placements to correctly classify PA intensity in pre-schoolers.

Children and adolescents

A higher compliance for wrist-worn versus hip-worn devices has been reported in children/adolescents [114]. However, similar wear time was achieved in protocols using 24h waist-worn compared to 24h wrist-worn accelerometers [168].

In regards to cut-points to classify time in SB and PA intensity, non-dominant wrist placement achieved a lower ROC-AUC (0.64-0.89) [70] compared to the dominant wrist (0.83-0.94) [123] and hip (>0.90) for all cut-points [69,108]. Furthermore, Hildebrand et al. [61] found a greater percentage of the explained variance when using algorithms from the hip compared to the wrist (78% for hip; 71% for wrist).

Previously developed sleep algorithms for the wrist placement were tested on the hip and wrist by Hjorth et al. [164]. They obtained an classification accuracy between 86.6% and 89.9% for the algorithms tested (developed with GT1M) [38,39] in hip compared to wrist measurement. Hip placement overestimated total sleep time compared to the wrist (60.1 vs. 73.8 min per day for wrist and hip, respectively). Finally, Tudor-Locke et al. [110] developed an algorithm to identify bedtime for the hip-worn GT3X/+, and Barreira et al. [111] refined and validated these in a free-living environment against self-reported participant logs, where they obtained a non-significant absolute difference of 9±36 min.

Abbreviations in this page:
PA: physical activity
PAEE: physical activity-related energy expenditure
VMCounts: activity counts in the vector magnitude
TEE: total energy expenditure

Adults

Minimal differences between contralateral hips were found for VMCounts (effect size: 0.016, p=0.619) and wear time (effect size: 0.040, p=0.213) [136]. The reliability of the GT3X/+ attached to the hip, wrist and ankle was studied by Ozemek et al. [165], who found high correlations from 0.824 to 0.998 in VMCounts between pairs of devices under simulated activities of daily living.

Staudenmayer et al. [115] demonstrated greater accuracy for PA classification when the device was placed on the wrist compared to previously developed cut-points with the accelerometer placed on the hip. Additionally, they found that newly-developed algorithms could also categorize behaviours in a lab setting (e.g., sitting, standing, riding in a vehicle, walking, and running) better for the wrist compared to the hip placement. Ellis et al. [118] achieved better performance with a wrist model to predict household simulated activities, however, the hip model outperformed the wrist model on locomotion prediction (i.e., slow walk, brisk walk and jogging) as well as PAEE estimation. In contrast, Ellis et al. [127] and Hildebrand et al. [61] obtained a higher accuracy (5% more on average) and a larger explained variance (81% for hip vs. 75% for wrist), respectively for the hip compared to the wrist placement to classify PA type and intensity. Stec et al. [161] found a significant correlation between VMCounts and TEE from hip but not from wrist worn accelerometers during resistance exercise (Pearson correlations for hip=0.50, p=0.005; and wrist=-0.25, p=0.18).

In regards to step counting, Tudor-Locke et al. [167] found higher accuracy for step counting from hip compared to wrist devices in controlled conditions against direct observation. No data about placement comparisons were found in adults for sleep-related behaviour estimations.

Older adults

The hip has been the most commonly used placement for studies in older adults. Only one study by Choi et al. [122] placed the GT3X on the dominant wrist to validate their non-wear time algorithm. However, we found no data comparing different device placement in this age group for any of the accelerometer outcomes.

Sampling frequency

Due to an insufficient number of studies this section data from all age groups are combined. GT3X records accelerations at a sampling frequency of 30 Hz. However, with the release of GT3X+, the manufacturer allowed users to select the sampling frequency between 30 and 100 Hz. Brønd and Arvidsson [169] demonstrated that sampling frequency had an effect on activity counts (i.e., a difference of +90 CPM for a slow walk, +180 CPM for a fast walk, +103 CPM for

a slow run and +1601 CPM for a fast run at a sampling frequency of 40 Hz compared to 30 Hz). Since the filtering process was developed for 30 Hz, sampling frequencies in multiples of 30 produce the most accurate estimates. Particularly, these authors [169] observed that 30, 60 or 90 Hz produced similar VACounts whereas sampling frequencies at 40, 50, 70, 80 or 100 Hz offset the filter resulting in an increased number of VACounts.

Table 6
Studies developing cut-points for sedentary behaviour and physical activity intensity classification, PAEE, and sleep algorithms used in the articles reviewed and ordered by age group.

Age group	Hip		Dominant wrist		Non-dominant wrist	
Pre-schoolers	**VM**	**VA**	**VM**	**VA**	**VM**	**VA**
Cut-points	[125,126,128,152]	[125,126,128,139,142,145,146,152]	Not found	Not found	[124][a]	[124][a]
PAEE	[128,162]	[139]	Not found	Not found	Not found	Not found
Sleep	Not found	Not found	Not found	Not found	Not found	Not found
Children and adolescents	**VM**	**VA**	**VM**	**VA**	**VM**	**VA**
Cut-points	[61,69,108,117,126,147]	[68,69,148,150,154,155,126,132–134,137,140,143,144]	[123]	[123]	[61,70]	[70]
PAEE	[61,112,117,163]	[68,112,113,134,140,144,150,157]	Not found	Not found	[61]	Not found
Sleep	Not found	[110,111]	Not found	[38]	Not found	[38]
Adults	**VM**	**VA**	**VM**	**VA**	**VM**	**VA**
Cut-points	[107,117,153]	[37,131,133,137,138,149]	[115][b]	Not found	[61]	Not found
PAEE	[61,107,117,118,161]	[37,117,158]	Not found	Not found	[61,118]	Not found
Sleep	Not found	Not found	Not found	[38,39]	Not found	[38,39]
Older adults	**VM**	**VA**	**VM**	**VA**	**VM**	**VA**
Cut-points	[116,117,156]	[109,116,117,133,137,138,141,151,156]	Not found	Not found	Not found	Not found
PAEE	[117]	Not found	Not found	Not found	Not found	Not found
Sleep	Not found	Not found	Not found	[39]	Not found	[39]

[a] Cut-points developed on the left wrist
[b] Algorithm developed with machine learning, not usual cut-points

Valid day and valid week

We cannot present the information in this section separately for each age group due to the lack of studies. As Toftager et al. [170] reported, increasing the requirements for what is considered a valid day (i.e., the number of hours per day) and a valid week (i.e., the number of valid days with valid data) led to a decrease in sample size and therefore the study's power.

In the National Health And Nutrition Examination Survey (NHANES) 2003-2006, where participants wore accelerometers during waking hours, only 40%-70% of them achieved a minimum of 10 hours/day of wear time for ≥6 days, while in the NHANES 2011-2012, where participants were asked to wear the accelerometers 24 hours/day, the wear time achieved was 21-22 hours/day for ≥6 days [171].

Aadland et al. [135] examined how many days were needed to obtain an intraclass correlation coefficient (ICC) of 0.80 with different hours per day wear time criteria (≥8 hours/day, ≥10 hours/day and ≥12 hours/day). ICCs for a single day did not differ much for all variables when the wear time criteria increased (i.e., ICCs=0.20-0.53 for ≥8 hours/day, ICCs=0.21-0.53 for ≥10 hours/day, ICCs=0.23-0.52

Abbreviations in this page:
CPM: counts per minute
ICC: intraclass correlation coefficient
NHANES: National Health And Nutrition Examination Survey
PAEE: physical activity-related energy expenditure
VACounts: activity counts in the vertical axis
VA: vertical axis
VM: vector magnitude

Abbreviations in this page:
CPM: counts per minute
ICC: intraclass correlation coefficient
LFE: low-frequency extension
LPA: light physical activity
MVPA: moderate-to-vigorous physical activity
SB: sedentary behaviour

for ≥12 hours/day). The number of days needed for an ICC of 0.80 decreased with a more demanding wear time criterion (from 8.3 to 6.4 days for time in SB; from 4.4 to 3.7 days for LPA; and from 8.5 to 7.0 days for MVPA, all adjusted for wear time). Although the registration period is usually one week, two weeks were analysed in the aforementioned study. Also, Donaldson et al. [172] reported that 4 days of measurement would be comparable to one week for estimating time in SB (r^2=0.91).

Filter

Pre-schoolers

No data about the influence of the filter selected (i.e., normal vs. LFE) was found in pre-schoolers.

Children and adolescents

Hjorth et al. [164] used normal and LFE filters on GT3X+ data from hip mounted accelerometers during the night. Total activity counts were (on average) 2815 counts per night period higher with the LFE filter compared to the normal one. Assuming 8h of sleep, this means approximately 6 CPM more when the LFE filter is enabled. Therefore, total sleep time was 9 min per night higher with the normal filter compared to the LFE filter when using a hip mounted GT3X+.

Adults

Lyden et al. [173] found the normal filter more accurate compared to the LFE filter to identify time in SB and breaks in SB with hip-placed GT3X against direct observation. Ried-Larsen et al. [174] and Cain et al. [175] observed less SB and more PA with the LFE filter enabled. Non-wear time estimation was similar between filters in the study published by Cain et al. [175]. For sleep-related behaviours, Cellini et al. [176] found similar results for total sleep time and sleep efficiency with both filters in a short sleeping time of 2 hours. The use of LFE filter increased the step count by an average of approximately 6000 steps per day in free-living [167].

Older adults

Wanner et al. [177] observed a mean difference of +37.8±19.5 CPM when enabling the LFE filter compared to the normal filter. Therefore, less time in SB and more time in all PA intensities were observed with the LFE filter due to the influence on activity counts. The normal filter appears to be more accurate than the LFE filter when compared with the NL-100 pedometer [178].

Epoch length

Pre-schoolers

We did not find any information on the influence of epoch length on accelerometer output in pre-schoolers. However, several studies used a 5 s epoch based on the belief that the activity pattern of very young children is intermittent and shorter epoch lengths might be suitable to capture very short bouts of movement [124–126,179].

Abbreviations in next page:
CPM: counts per minute
LPA: light physical activity
LFE: low-frequency extension
METs: metabolic equivalents
MPA: moderate physical activity
ROC: receiver operating characteristic
SB: sedentary behaviour
VA: vertical axis
VM: vector magnitude
VPA: vigorous physical activity
VVPA: very vigorous physical activity

Table 7
Cut-points for SB, LPA, MPA, VPA and VVPA activity used in the articles reviewed and ordered by age-group in which they were validated

Age group / reference	Placement	Filter	Epoch	Axis	SB	LPA	MPA	VPA	VVPA
Pre-schoolers									
Butte et al. [128]	Right hip	Normal	60 s	VA	≤ 240	2120-4449	-	-	≥4450
				VM	≤ 820	3908-6111	-	-	≥6112
Costa et al. [125]	Right hip	LFE	5 s	VA	≤ 5	-	≥ 165	-	-
				VM	≤ 96.12	-	≥ 361.94	-	-
Jimmy et al. [126]	Right hip	Normal	5 s	VA	-	-	≥ 133	134-193	194-233
				VM	-	-	≥ 246	247-316	317-381
Johansson et al. [124]	Left wrist	Normal	5 s	VA	≤ 89	90-439	-	≥ 440	-
				VM	≤ 221	222-729	-	≥ 730	-
Pate et al. [139][a]	Right hip	Not reported	15 s	VA	≤ 37	38-419	420-841	≥ 842	-
Pulakka et al. [152]	Right hip	Normal	15 s	VA	-	-	≥ 35	-	-
				VM	-	-	≥ 208	-	-
Reilly et al. [142][a]	Right hip	Not reported	60 s	VA	≤ 274	-	-	-	-
Sirard et al. [145][a]									
3 years old	Right hip	Not reported	15 s	VA	≤ 301	302-614	615-1230	≥ 1231	-
4 years old					≤ 363	364-811	812-1234	≥ 1235	-
5 years old					≤ 398	399-890	891-1254	≥ 1255	-
Van Cauwenberghe et al. [146][b]	Right hip	Not reported	15 s	VA	≤ 372	373-584	585-880	≥ 881	-
Children and adolescents									
Aittasalo et al. [147][b]	Hip	Not reported	Raw: 100 Hz	VM	≤26.9 m*g*	27-332 m*g*	332-558 m*g*	≥558 m*g*	-
Andersen et al. [143][a]	Hip	Not reported	60 s	VA	≤499	500-1999	2000-2999	3000-4499	4500-32767
Chandler et al. [70]	Non-dominant wrist	Normal	5 s	VA	≤ 161	162-529	530-1461	≥ 1462	-
				Axis 2	≤ 132	133-445	446-998	≥ 999	-
				Axis 3	≤ 113	114-372	373-776	≥ 777	-
				VM	≤ 305	306-817	818-1968	≥ 1969	-
Crouter et al. [123] (ROC analysis)	Dominant wrist	LFE	5 s	VA	≤ 105	-	262-564	≥ 565	-
				VM	≤ 275	-	416-777	≥ 778	-
Crouter et al. [123] (regression analysis)	Dominant wrist	LFE	5 s	VA	≤ 35	36-360	361-1129	≥ 1130	-
				VM	≤ 100	101-609	610-1809	≥ 1810	-
Freedson et al. [140][a]	Right hip	Not reported	60 s	VA	≤ 149	150-499	500-3999	4000-7599	≥7600
Grydeland et al. [148]	Right hip	LFE	60 s	VA	≤ 100	101-2000	2001-6000	≥ 6001	-
Hänggi et al. [108]	Right hip	Normal	1 s	VM	≤ 2	3-56	≥56	-	-
Hildebrand et al. [61][b]	Right hip	Not reported	Raw: 60 Hz.	VM	-	≤ 142 m*g*	142-464 m*g*	≥ 464 m*g*	-
	Non-dominant wrist	Not reported			-	≤ 201 m*g*	201-707 m*g*	≥ 707 m*g*	-
Jimmy et al. [126]	Right hip	Normal	5 s	VA	-	-	-	5 METs: 193 6 METs: 233	-
				VM	-	-	-	5 METs: 316 6 METs: 381	-
Matthews et al. [133][a]	Right hip	Not reported	60 s	VA	≤ 100	-	-	-	-
Mattocks et al. [150][a]	Right hip	Not reported	60 s	VA	-	-	3581-6129	≥ 6130	-
Puyau et al. [144][a]	Right hip	Not reported	15 s	VA	≤ 199	200-799	800-2049	≥ 2050	-
Ridgers et al. [132][b]	Right hip	Not reported	60 s	VA	≤ 100	-	-	-	-
Romanzini et al. [69]	Hip	Not reported	15 s	VA	≤ 46	47-606	607-817	≥ 818	-
				VM	≤ 180	181-756	757-1111	1112	-
Santos-Lozano et al. [117]	Right hip	Normal	60 s	VM	-	-	≤ 2114	2115-6548	≥ 11490
Treuth et al. [134][a]	Right hip	Not reported	30 s	VA	≤ 50	51-1499	1500-2600	> 2600	-
	Right hip	Not reported	60 s	VA	≤ 100	101-2999	3000-5200	> 5200	-
Troiano et al. [137]	Right hip	Not reported	60 s	VA	≤ 100	101-2019	2020-5998	≥5999	-
Vanhelst et al. [154][b]	Hip	Not reported	60 s	VA	≤ 400	401-1900	1901-3918	≥ 3919	-
Zhu et al. et al. [155]	Right hip	Not reported	60 s	VA	-	-	2860-3839	≥ 3840	-

Table 7 – continued
Cut-points for SB, LPA, MPA, VPA and VVPA activity used in the articles reviewed and ordered by age-group in which they were validated

Age group / reference	Placement	Filter	Epoch	Axis	SB	LPA	MPA	VPA	VVPA
Adults									
Freedson et al. [37][a]	Right hip	Not reported	60 s	VA	≤ 99	100-759	760-5724	5725-9498	≥ 9499
Hildebrand et al. [61][b]	Right hip	Not reported	Raw: 60 Hz.	VM	-	≤ 69 m*g*	69-259 m*g*	≥ 260 m*g*	-
	Non-dominant wrist	Not reported			-	≤ 101 m*g*	101-429 m*g*	≥ 430 m*g*	-
Kozey-Keadle et al. [131]	Right hip	LFE	60 s	VA	≤ 150	-	-	-	-
Matthews et al. [180][a]	Right hip	Not reported	60 s	VA	-	-	760-5998	≥ 5999	-
Matthews et al. [133][a]	Right hip	Not reported	60 s	VA	≤ 100	-	-	-	-
Metzger et al. [138][a]	Right hip	Not reported	60 s	VA	< 100	100-2019	-	-	-
Santos-Lozano et al. [117]	Right hip	Normal	60 s	VM	-	-	≤ 3208	3209-8565	≥ 11593
Sasaki et al. [107]	Right hip	Normal	60 s	VM	-	≤ 2690	2691-6166	6167-9642	≥ 9643
Troiano et al. [137][a]	Right hip	NR	60 s	VA	≤ 100	101-2019	2020-5998	≥ 5998	-
Older adults									
Aguilar-Farias et al. [116]	Right hip	LFE	1 s	VA	< 1	-	-	-	-
			15 s		≤ 9	-	-	-	-
			60 s		≤ 24	-	-	-	-
			1 s	VM	< 1	-	-	-	-
			15 s		≤ 69	-	-	-	-
			60 s		≤ 199	-	-	-	-
Copeland et al. [109][a]	Right hip	Not reported	60 s	VA	≤ 99	100-1040	≥ 1040	-	-
Davis et al. [141][a]	Right hip	Not reported	60 s	VA	≤ 199	200-1999	2000-3999	≥ 4000	-
Kozey-Keadle et al. [131]	Right hip	LFE	60 s	VA	≤ 150	-	-	-	-
Matthews et al. [133][a]	Right hip	Not reported	60 s	VA	≤ 100	-	-	-	-
Metzger et al. [138]	Right hip	LFE	60 s	VA	≤ 149	150-2019	≥ 2020	-	-
Pruitt et al. [151]	Right hip	Not reported	60 s	VA	-	-	-	-	-
Santos-Lozano et al. [117]	Right hip	Normal	60 s	VM	-	-	≤ 2751	2752- 9359	≥9360
Troiano et al. [137][a]	Right hip	Not reported	60 s	VA	≤ 100	101-2019	2020-5998	≥ 5998	-
Zisko et al. [156]									
Men	Right hip	Not reported	60 s	VA	≤ 55	56-266	267-1971	1972-3878	≥ 3879
				VM	≤ 610	611-1652	1653-3016	3017-4581	≥ 4582
Women				VA	≤ 59	60-212	213-1217	1218-3157	≥ 3158
				VM	≤ 464	465-1076	1077-2424	2425-4078	≥ 4079

[a] Cut-points developed with the 7164 or the GT1M model (ActiGraph, Pensacola, FL, USA)
[b] Cut-points not expressed in counts, but in other units

Abbreviations in this page:
CPM: counts per minute
LPA: light physical activity
LFE: low-frequency extension
MPA: moderate physical activity
MVPA: moderate-to-vigorous physical activity
SB: sedentary behaviour
VA: vertical axis
VM: vector magnitude
VPA: vigorous physical activity
VVPA: very vigorous physical activity

Children and adolescents

Aibar et al. [181], compared the effect of different epoch lengths (3-60 s) on PA intensity, and found a progressive decrease in the time spent in MVPA intensity as the epoch length increased. Furthermore, they found that smaller epoch lengths increased the resolution of the measure, thus increasing the time spent in VPA intensity [181]. They suggested to use shorter epoch lengths (e.g., 3-15 s) in children.

Adults

No information on epoch length was found for adults.

Older adults

No information on epoch length was found for older adults.

Non-wear time definition

Pre-schoolers

No information on non-wear time was found for pre-schoolers.

Children and adolescents

Toftager et al. [170] showed that the longer the non-wear time duration the greater the number of participants that were included in the analyses. Furthermore, as the non-wear time duration increased the average CPM decreased (e.g., average PA level: 641 CPM with strings of 10 min of consecutive zeros compared to 570 CPM

with strings of 90 min of consecutive zeros) [170]. Since Toftager et al. [170] compared different non-wear time definitions without a criterion, it is not possible to conclude which algorithm was more or less valid.

Adults

Peeters et al. [182] compared six different definitions of non-wear time (i.e., 20, 60 and 90 min with and without allowance of 2 min of low accelerations). It was observed that 20 min of 0 CPM without allowing for interruptions resulted in a lower misclassification (5.9%) and a similar ROC-AUC (0.94) than 60 min (6.7%, ROC-AUC=0.94) and 90 min (7.4%, ROC-AUC=0.93) [182]. However, in these conditions, more participants did not meet the non-wear time criteria (32 out of 34 participants, i.e., 6% sample loss) compared to 60- or 90-min algorithms (33 and 34 out of 34 participants, i.e., 3% and 0% sample loss, respectively).

Older adults

Keadle et al. [120] compared the Troiano et al. algorithm [137], which uses a minimum of 60 min of 0 CPM with an allowance of 2 min of interruptions, with the Choi et al. algorithm [183], which uses a minimum of 90 min of 0 CPM with the same allowance as the Troiano algorithm plus two 30 min windows of 0 CPM before and after that allowance. They concluded that the algorithm by Choi et al. [183] was the best to identify wear time compared with diary records of the participants. The same conclusion was obtained in a later study by Choi et al. [122], especially when this algorithm was implemented for wrist-worn accelerometers, because the wrist placement is more sensitive to detect non-wear time than the hip [122].

Registration period: waking vs. 24-hours

Due to an insufficient number of studies this section combines all age groups. Recent large scale studies such as NHANES (2011-2012) and the International Study of Childhood Obesity, Lifestyle and Environment (ISCOLE) 2012-2013 [184] have used a 24-hour protocol (24h). Tudor-Locke et al. [168] found higher wear time compliance with 24h protocols compared to waking hour protocols, with this finding being consistent across different countries.

Sedentary behaviour and physical activity intensity classification

Pre-schoolers

Two studies comparing several cut-points developed from the VACounts from hip mounted devices were found [179,185]. Janssen et al. [185] supported the use of Evenson et al. [68] SB cut-point due to the higher classification accuracy compared to other cut-points [139,142,144,145,186], and recommended that the Pate et al. [139] cut-points are the best option for MVPA (all of them were developed with former models of ActiGraph). However, Kahan et al. [179]

Abbreviations in this page:
CPM: counts per minute
ISCOLE: International Study of Childhood Obesity, Lifestyle and Environment
NHANES: National Health And Nutrition Examination Survey
ROC-AUC: receiver operating characteristic area under the curve
SB: sedentary behaviour
VACounts: activity counts in the vertical axis

Abbreviations in this page:
LPA: light physical activity
MPA: moderate physical activity
MVPA: moderate-to-vigorous physical activity
PA: physical activity
ROC: receiver operating characteristic
ROC-AUC: receiver operating characteristic area under the curve
SB: sedentary behaviour
VMCounts: activity counts in the vector magnitude
VPA: vigorous physical activity

observed, in a small sample size (n=12), that Sirard et al. [145] cut-points showed the best agreement with direct observation for time in SB and MVPA compared to other cut-points [68,139,186] developed with former models of ActiGraph.

Four studies developed cut-points for time in SB and PA intensity for VMCounts from accelerometers worn on the hip [125,126,128,152]. Butte et al. [128] developed their cut-points using a 60 s epoch considering energy expenditure cut-points established through smoothing splines and ROC curves. The cut-points developed by Costa et al. [125] used 5 and 15 s epochs and were validated against direct observation. Jimmy et al. [126] developed and validated their cut-points utilizing a 5 s epoch based on indirect calorimetry. Pulakka et al. [152] developed one cut-point to differentiate SB/LPA from MVPA (they did not differentiate MPA from VPA, as done in the aforementioned cut-points) and validated it against direct observation. All of these studies obtained high ROC-AUC (0.89-0.98 for all cut-points).

Finally, Johansson et al. [124] was the only study that developed and validated cut-points for VMCounts from a wrist-worn accelerometer against direct observation in young pre-schoolers (15-36 months) obtaining a ROC-AUC of 0.89-0.98.

Children and adolescents

Zhu et al. [155] compared a set of cut-points for estimating time in SB and PA intensity developed using the VACounts, with the accelerometer worn on the hip in a sample of Chinese children. The authors observed a better accuracy with the cut-points proposed by Evenson et al. [68], Vanhelst et al. [154] and those internally developed [155] than with the rest of the cut-points tested [140,144,150] (all these cut-points were developed with former models of ActiGraph).

Five studies developed cut-points for VMCounts from the hip [69,108,126,187]. Peterson et al. [187] suggested that 150 CPM from hip mounted accelerations is the most accurate SB cut-point compared with direct observation. Hänggi et al. [108] developed their cut-points using a 1 s epoch in comparison with indirect calorimetry and obtained a ROC-AUC of 0.96 for SB, LPA and MPA. These cut-points [108] obtained better correlations with other brands of accelerometers than other VACounts-based cut-points [188]. Jimmy et al. [126] developed cut-points utilizing a 5 s epoch against indirect calorimetry and attained a ROC-AUC ranging from 0.89 to 0.94 for all intensities. Romanzini et al. [69] validated cut-points using a 15 s epoch against indirect calorimetry and obtained a ROC-AUC of 0.93-0.99. Finally, Santos-Lozano et al. [117] validated cut-points utilizing a 60 s epoch against indirect calorimetry and found the lowest ROC-AUC (0.6-0.8).

VM cut-points from the wrist placement were developed in three studies [61,70,123]. Chandler et al. [70] validated cut-points for the non-dominant wrist using a 5 s epoch against direct observation and attained a ROC-AUC ranging between 0.64-0.89. A higher ROC-AUC was obtained by Crouter et al. [123] using cut-points for the dominant wrist which were developed in a 5 s epoch against indirect calorimetry (ROC-AUC of 0.83-0.94). It is important to highlight that Crouter et al. [123] applied linear regression models to the dominant wrist and obtained non-significant differences between the accelerometer outputs and indirect calorimetry (mean biases ranged from 2.2% to 8.4% for all cut-points).

Finally, we found two studies using metrics extracted directly from raw data instead of activity counts by ActiGraph. Aittasalo et al. [147] developed a method based on amplitude of accelerations from the hip's raw accelerations. These cut-points were validated against heart rate monitoring using an ordinal logistic regression and showed a correlation coefficient of 0.97. However, these results must be interpreted carefully since only were walking and running at different intensities were used during the development of the intensity cut points. Hildebrand et al. [61] used a linear regression analysis to establish the relation between an acceleration metric based on raw data (i.e., ENMO) and energy expenditure measured through indirect calorimetry. Then, from the developed regression equations, they defined two sets of cut-points for the hip and the non-dominant wrist. They obtained correctly classified values between 96-97% for SB and LPA, 33-55% for MPA and 68-80% for VPA.

Adults

Kozey-Keadle et al. [131] tested some VACounts-based cut-points and determined that 150 CPM using the VACounts from hip accelerations was the most accurate SB cut-point compared with direct observation. Santos-Lozano et al. [117] validated cut-points for PA intensity against indirect calorimetry and obtained a ROC-AUC between 0.6-0.8. Sasaki et al. [107] used a linear regression model to establish the relation between ActiGraph VMCounts from the hip and energy expenditure measured by indirect calorimetry. The mean differences between the METs predicted by the cut-points derived from the regression model and the actual METs were -0.3, -0.4 and 0.7 at MPA, VPA and VVPA intensities respectively.

Three studies developed cut-points from raw data metrics. Vähä-Ypyä et al. [153] developed an amplitude-domain method for raw hip accelerations. The cut-points were validated against heart rate monitoring and showed an excellent agreement (ROC-AUC=0.99 for all cut-points), however, they were not used during free-living conditions. Hildebrand et al. [61] validated regression models for the hip and the non-dominant wrist against indirect calorimetry and

Abbreviations in this page:
CPM: counts per minute
ENMO: Euclidean Norm minus 1G
LPA: light physical activity
MPA: moderate physical activity
PA: physical activity
ROC-AUC: receiver operating characteristic area under the curve
SB: sedentary behaviour
VM: vector magnitude
VPA: vigorous physical activity

Abbreviations in this page:
LPA: light physical activity
MPA: moderate physical activity
PA: physical activity
PAEE: physical activity-related energy expenditure
ROC-AUC: receiver operating characteristic area under the curve
SB: sedentary behaviour
VACounts: activity counts in the vertical axis
VMCounts: activity counts in the vector magnitude
VPA: vigorous physical activity
VVPA: very vigorous physical activity

defined cut-points from the regression equations generated. They obtained correctly classified values between 93-96% for SB and LPA, 54-59% for MPA and 89-92% for VPA. Finally, Staudenmayer et al. [115] developed a classifier for PA intensity based on decision trees for the dominant wrist and they obtained 75% of values correctly classified using indirect calorimetry. Within this context, they reported preliminary results that their model performs well in a free-living environment [115].

Older adults

Keadle et al. [120] observed that cut-points using the VACounts or VMCounts are not comparable. Unfortunately, they could not report which cut-points were the most accurate since they did not determine a criterion to compare the outcomes [120]. Aguilar-Farias et al. [116] validated SB cut-points utilizing VMCounts acquired from the hip with 1, 15 and 60 s epochs against ActivPAL^{3TM} (Pal Technologies Ltd., Glasgow, UK) and found a high classification accuracy (ROC-AUC of 0.82, 0.85 and 0.86 for 1, 15 and 60 s epochs respectively). Santos-Lozano et al. [117] validated MPA, VPA and VVPA cut-points against indirect calorimetry and obtained a ROC-AUC of 0.7 for all intensities examined.

Physical activity-related energy expenditure algorithms

Ten studies that developed PAEE algorithms were found [61,107,112,117,118,128,161–163,185]. Due to an insufficient number of studies that used doubly labelled water or room calorimetry as criteria, this section combines age groups. Only two studies validated their algorithms against doubly labelled water or room calorimetry in pre-schoolers. Butte et al. [128] developed cross-sectional time series and multivariate adaptive regression splines to predict PAEE using both GT3X+ and heart rate monitoring. They validated the algorithms under controlled conditions in room calorimetry and in free-living conditions utilizing doubly labelled water [128]. The multivariate adaptive regression splines obtained a better prediction of PAEE against room calorimetry, i.e., inter-method mean difference equal to 0.006±0.085 kcal/min, however, the cross-sectional time series model achieved a better prediction in free-living conditions, using doubly labelled water (mean difference 41±97 kcal/day) [128]. Zakeri et al. [162] used the same two statistical methods described above with GT3X/+ and heart rate monitoring. They obtained better prediction with the cross-sectional time series model against room calorimetry (i.e., 0.001±0.070 kcal/min), but they did not validate the method in a free-living environment [162]. All these studies were carried out with hip-worn GT3X/+.

Sleep-related behaviours

Pre-schoolers

We did not find any study comparing different sleep algorithms in this age group. Only Meredith-Jones et al. [189] used the Sadeh et al. [38] sleep algorithm to identify sleep time in pre-schoolers. However, this algorithm was developed in an older sample (10-25 years), and the results should be interpreted cautiously.

Abbreviations in this page:
ICC: intraclass correlation coefficient

Children and adolescents

Hjorth et al. [164] compared the performance of existent sleep algorithms from the hip versus the non-dominant wrist placements. Despite the fact that these algorithms were developed for wrist accelerations, they obtained good accuracy (86.6-89.9%) [38,39] at both placements; however, the hip-worn device overestimated total sleep time compared to the wrist (60.1-73.8 min per day). These findings may be affected by the fact that bedtime was reported by participants using diaries and not through an algorithm.

Tudor-Locke et al. [110] developed an algorithm to detect bedtime for the hip-worn GT3X/+, and Barreira et al. [111] refined and validated it in a free-living environment against self-reported bedtime. They obtained a non-significant absolute difference of 9±36 min of bedtime per night [111].

Adults

Cellini et al. [176] found an accuracy of 82.8% for classifying epoch-by-epoch sleep or awake status against polysomnography (i.e., the gold standard to measure sleep patterns) using the Sadeh et al. [38] sleep algorithm. However, they found an overestimation of total sleep time (i.e., inter-method mean difference equal to 8.80 minutes) and sleep efficiency (i.e., inter-method mean difference equal to 14.53%), as well as an underestimation of sleep onset latency (ICC=0.56) and awakenings after sleep onset (ICC=0.54) during a two-hour sleep protocol. Rosenberger et al. [190] observed a mean difference of 4 min of sleep time for the Sadeh et al. [38] sleep algorithm compared to the Z-machine (portable monitor to measure brain activity which is relatively comparable to polysomnography [191]). Equally, Slater et al. [166] reported good accuracy of the Sadeh et al. [38] algorithm to detect total sleep time and moderate validity for awakenings after sleep onset against polysomnography from the wrist, but not from the hip. Finally, Zinkhan et al. [192] tested the performance of the Cole et al. [39] algorithm for the hip, even though it was developed for wrist accelerations. They observed a limited agreement with total sleep time measured by polysomnography (mean difference of 81.1 min per night).

Abbreviations in this page:
PA: physical activity
PAEE: physical activity-related energy expenditure
SB: sedentary behaviour

Older adults

No data about sleep algorithms were found in the papers reviewed for older adults.

Step counting

Only data on step counting estimated by ActiLife software are available in adults. Tudor-Locke et al. [167] found a higher accuracy for step counting from the hip mounted monitors over wrist mounted ones under controlled conditions against direct observation. Under free-living conditions, the wrist-worn accelerometer detected more steps than the hip-worn one independently of the filter used. See the Filter sections for more information about how different filters influence step counting [167].

Discussion

The use of objective methods when assessing time in SB, PA, PAEE and sleep in various research settings has increased enormously as indicated by the large number of articles included in this review. Accelerometry has several advantages over questionnaires and self-report methods [31,104], however data collection and processing criteria have a large impact on the interpretation of the data. Thus, predetermined decisions about data collection and processing in relation to the study participants and the objective of the study are important when planning research in this area.

A major finding of this review is that many of the studies did not report on key methodological issues. Therefore, data cannot be accurately compared between studies and readers may have difficulties interpreting how various methodological decisions may have influenced the main findings/conclusions. We recommend that future studies should report the complete set of criteria included in the present review in order to improve data comparability and reproducibility.

The next subsections provide practical considerations for every criterion based on the critical information extracted from the validation/calibration studies reviewed.

Device placement

Traditionally, cut-points to assess time in SB and PA variables, as well as PAEE estimates, were developed with the device placed on the hip, while algorithms to assess sleep-related behaviours were more commonly developed based on wrist accelerations. Ideally, researchers may want to collect accelerometer data using a 24-hour protocol with one accelerometer attached to either the hip or the wrist and be able to estimate time in SB, PA and sleep-related behaviours. However, lack of validated algorithms in some age groups [61,70,110,111,115,123] preclude this.

Abbreviations in this page:
PA: physical activity
PAEE: physical activity-related energy expenditure
SB: sedentary behaviour

We decided not to differentiate between the right or left hip because no significant differences were found by Aadland et al. [136]. We consider that both the hip and the wrist are feasible places to attach the GT3X/+. Better compliance for wrist compared to hip worn devices has been reported in children and adolescents [114], but similar wear time was found in large-scale studies in adults using hip and wrist placements [168]. We therefore cannot confirm the general belief supporting better compliance for wrist-worn devices. More studies are needed to investigate compliance differences between wrist- and hip-worn devices as well as the extent to which these differences influence the validity and reliability of accelerometer outcomes.

There are only a few studies directly comparing two placement sites using the GT3X/+ and they have consistently shown more accurate classification of time in SB and PA intensity as well as estimates of PAEE when the accelerometer was worn on the hip compared to the wrist [61,69,70,108,118,123,161]. However, one study found a better performance for the wrist-worn device for PA intensity classification [115], see **Table 5**. Step count also differs greatly depending on the device placement, i.e., more steps (>2500) were counted when wearing the accelerometer on the wrist compared to the hip in free-living conditions [167]. When studied under controlled conditions, hip placement has shown more accurate step counting than wrist placement from a speed of 54 m/min and upwards (at lower speeds, accuracy was better in the wrist) [167]. The lower accuracy for the wrist-worn devices could be due to the fact that accelerations such as brushing teeth might be interpreted as steps when the device is placed on the wrist, but not on the hip, nevertheless, this is just a hypothesis that needs to be confirmed by data under free-living conditions.

In regards to sleep algorithms, Hjorth et al. [164] compared the functioning of two algorithms applied to hip data against wrist data, finding an overestimation of the sleep time and a high accuracy (86.6 and 89.9 for each algorithm) from the hip compared to the wrist. However, it is important to note that these investigators imputed sleep and wake time manually from logs kept by the participants. The use of logs by the participants might explain the high accuracy achieved using a wrist-developed sleep algorithm on the hip.

Sampling frequency

Our recommendation is to use the highest sampling frequency possible, as we cannot anticipate future data processing needs. However, given the issues associated with other sampling frequencies other than 30 Hz or its multiples as described in the Sampling frequency section, sampling frequencies in multiples of 30 Hz seem to produce more accurate estimates when processing the signal using

Abbreviations in this page:
LFE: low-frequency extension
PA: physical activity

the methods proposed by ActiGraph. Therefore, the most reasonable conclusion for the time being would be to use 90 Hz when researchers are using the manufacturer methods, and 100 Hz when researchers are filtering and processing the signal on their own.

Valid day and valid week

To ensure that data are representative of an entire day, it is necessary to establish how many hours of wear time are required. It is also necessary to set how many valid days are needed to be representative of the total assessment period, which is usually one week (7 complete days). Wear time criteria for a valid day depends on the registration period, i.e., waking hours or 24-hours. In studies in which the accelerometer is worn for 24-h periods to assess both PA and sleep-related behaviours, the number of hours required for a day to be considered valid has to be larger than studies in which the accelerometer was taken off at night.

Similarly, increasing requirements for a valid day and a valid week provides more reliable data (more information can be found in **Table 7** from the study by Aadland et al. [135]), however, it results in greater sample loss. Our recommendation is to test different criteria to get the best compromise between sample size (and therefore optimal statistical power) and reliability of the measure. However, a minimum of 4 days of valid data is recommended as was suggested in a previous systematic review [193].

Filter

When movements (accelerations) occur at too low or high frequencies, ActiGraph interprets that this acceleration might not be compatible with human movement and should therefore be excluded from the analyses (e.g., if someone is using a drill). The GT3X/+ filtering process to exclude this kind of acceleration is implemented in the ActiLife software. This software allows users to choose between two different filters when processing the data: normal (default) and LFE filters.

The algorithms for these filters are proprietary information. It is known that a normal filter detects accelerations from a frequency range of 0.25-2.5 Hz, while the LFE filter establishes a lower threshold to capture slower movements; however, it is unknown exactly how much lower this threshold is. A weighting function is applied to the accelerations between the range of 0.25-2.5 Hz, so that the full weight (i.e., 1.0) is given to a frequency of acceleration of 0.75 Hz, and lower weighting is given to higher and lower movement frequencies progressively [194]. Accelerations at a frequency greater than 2.5 Hz are removed by the filter, although it is important to highlight that accelerations up to 3.4 Hz can be produced by the human body when performing PA at vigorous intensities when the device is attached to the hip (higher frequencies are achieved in the wrist) [195,196].

Therefore, ActiGraph's filtering process might remove accelerations associated with VPA, and consequently, minutes in VPA might be misclassified as MPA [169,197]. As the filter used has a large impact on the accelerometer outputs, it is alarming that 74% of the studies reviewed did not report this key information (**Figure 9**).

When selecting a set of cut-points or an algorithm to estimate a variable from activity counts, our recommendation is to select the same filter that was used in the validation study for the cut-points or algorithm employed (**Table 7** shows the filters used in all the cut-points identified). If cut-points or algorithms are not used, then researchers can decide which filter to use, we suggest using the LFE filter when low movements are of greater importance (e.g., when analysing time in SB, sleep or PA in older adults). From the studies discussed above, researchers and practitioners should be aware that enabling the LFE filter compared with the normal filter will result in decreased time in SB, greater time in PA at all intensities and an increase in the number of steps per day.

Epoch length

Activity counts produced by filtering raw accelerations need to be summed into specific time intervals or epoch in order to estimate PAEE, time spent in SB, in various levels of PA intensity, as well as estimating sleep/wake state, and this is usually done by applying specific intensity cut-points and algorithms.

Given that epoch length influences activity counts, it is important to use the same epoch length that was used in the validation study for the cut-points or algorithms (see **Table 7**). Epoch length should also be taken into account when comparing data from different studies. In young people (from pre-schoolers to adolescents), shorter epochs (1-15 s) are recommended to capture short bouts of activity occurring frequently in these age groups. In adults and older adults there are currently no data comparing the effect of epoch length on the outcomes studied. However, our own unpublished data suggest that selecting a 1 s *vs.* a 60 s epoch length has a marked impact on the accelerometer outcomes, i.e., +45-60min/day in MVPA using a 1 s compared to a 60 s epoch. This large impact on the accelerometer outputs warrants further research on this topic in order to obtain more comparable and accurate data. Considering sleep measurements there is an overall consensus for using a 60 s epoch (probably due to the stable movement pattern during sleep), as all sleep algorithms have been validated using that epoch length [38,39,110,111].

Non-wear time definition

In free-living studies, accelerometers are usually removed during water-based activities e.g., swimming or showering and when sleeping (in some studies). As a result, individuals might forget to

Abbreviations in this page:
LFE: low-frequency extension
MPA: moderate physical activity
MVPA: moderate-to-vigorous physical activity
PA: physical activity
PAEE: physical activity-related energy expenditure
SB: sedentary behaviour
VM: vector magnitude
VPA: vigorous physical activity

Abbreviations in this page:
CPM: counts per minute
LPA: light physical activity
MPA: moderate physical activity
PA: physical activity
ROC-AUC: receiver operating characteristic area under the curve
SB: sedentary behaviour
VPA: vigorous physical activity
VVPA: very vigorous physical activity

wear the accelerometer for a day(s) or part of day. Consequently, non-wear time must be identified (e.g., by a diary or algorithms) and excluded from data before analysis. Otherwise, this time is likely categorized as SB time. Generally, algorithms to detect non-wear time consist of intervals of time with consecutive 0 CPM with or without an allowance of several minutes in which small accelerations are allowed, with optional windows of 0 CPM before and after this allowance. Toftager et al. [170] studied the effect of different non-wear time definitions and concluded that the most accurate algorithm might differ among subgroups of children/adolescents. For example, studies focused on overweight adolescents might need to set a longer time of consecutive 0 CPM, since they have higher SB that can be misclassified as non-wear time.

More studies are needed to examine the accuracy of different non-wear time detection algorithms in all age groups. Based on the reviewed studies, we cannot recommend a non-wear time definition for pre-schoolers, children or adolescents. For adults, 20 min of consecutive 0 CPM without allowance showed the lowest misclassification error; however, it may result in slightly more loss of data (6% of the sample size [182]). As the accuracy between 20 min and 60 min of consecutive 0 CPM was similar (i.e., the ROC-AUC was virtually identical=0.94), we suggest to use 60 min of consecutive 0 CPM without allowing for interruptions in counts in this period for adults, to avoid the risk of misclassification of non-wear time as SB. In older adults, we recommend the Choi et al. algorithm [122], which consists of 90 min of 0 CPM with an allowance of 2 min of activity when it is placed between two 30 min windows of 0 CPM. This algorithm outperformed other algorithms on the detection of non-wear time [122] compared with the non-wear time reported by participants.

Registration period: waking vs. 24-hours

In line with recent and large-scale studies [184], we suggest registration periods of 24h instead of waking hours (more recording time, therefore more valid data). This is mainly due to an interest in assessing sleep-related behaviours and better compliance.

Sedentary behaviour and physical activity intensity classification

Traditionally, time in SB and PA intensity have been estimated based on the number of activity counts accumulated in a certain period (epoch length). Cut-points are the thresholds of activity counts used to categorize activity as SB, LPA, MPA, VPA or VVPA. **Table 7** presents the cut-point values (expressed as counts per time unit) that are currently available for SB, and for LPA, MPA, VPA and VVPA by age group. It is important to keep in mind that although the GT3X/+ is a triaxial accelerometer, the data are provided separately for the 3 axes plus the VM, so that it is still possible to use the data

registered only in the VA and apply it to the previously developed algorithms for such axis.

When applying cut-points to a specific data set, it is recommended to follow the same data collection and processing criteria which were used in the original validation/calibration study (see **Table 7**). All derived intensity thresholds are influenced by the activities chosen when performing the calibration studies. Thus, it is impossible to recommend the most appropriate set of intensity thresholds for free-living assessment. Also, different generations of ActiGraph devices have shown to be comparable under controlled conditions [107,198], but not in a free-living environment [107,174,175,199]. This suggests that if a certain cut-point was developed, for instance using the VACounts from the GT1M, that cut-points may not be used for data collected with the GT3X/+ VACounts, since they are not fully comparable. Therefore, our recommendations are based on cut-points developed only with GT3X/+ accelerometers. This review shows the need for future meta-analytic studies summarizing cut-points for each age group in order to obtain a set of cut-points with a wide range of activities influencing its development. Finally, across the studies reviewed, we have observed a widely accepted criterion to define PA intensity in the studies validating cut-points against indirect calorimetry, i.e., 1-1.5 MET for SB, 1.5-3 MET for LPA, 3-6 MET for MPA and > 6 MET for VPA.

The criteria considered for cut-point recommendations are: 1) the cut-points cover the whole activity spectrum (i.e., SB, LPA, MPA and VPA), 2) calorimetry as an objective criterion is better than direct observation; 3) for young populations, cut-points developed in short epochs; 4) the number and type of activities included in the study that derived the cut-points; and 5) results obtained in comparison with the criterion.

Pre-schoolers experience a rapid anatomical development and their patterns of PA change dramatically during the first years of life; therefore, the age of the sample is very important in pre-schoolers. Thus, for the hip placement we recommend Costa et al. [125] cut-points for early pre-schoolers (2-3 years old) and Jimmy et al. [126] cut-points for older pre-schoolers (4-6 years old). They have been developed in short epochs which enables the devices to capture small bouts of VPA, which is typical for this age group, while obtaining high accuracy in their validation. For young pre-schoolers (15-36 months) using the wrist placement, we recommend Johansson et al. [124] cut-points developed using a 5 s epoch because they obtained similar accuracy to hip-developed cut-points.

For children, we recommend to use the Hänggi et al. [108] cut-points developed in 1 s epoch for the hip. For adolescents, the Romanzini et al. [69] cut-points developed utilizing a 15 s epoch

Abbreviations in this page:
LPA: light physical activity
MPA: moderate physical activity
PA: physical activity
SB: sedentary behaviour
VA: vertical axis
VACounts: activity counts in the vertical axis
VM: vector magnitude
VPA: vigorous physical activity
VVPA: very vigorous physical activity

Abbreviations in this page:
CPM: counts per minute
MPA: moderate physical activity
PA: physical activity
PAEE: physical activity-related energy expenditure
ROC-AUC: receiver operating characteristic area under the curve
SB: sedentary behaviour
VPA: vigorous physical activity

appears appropriate. Both of these obtained excellent classification accuracy (ROC-AUC > 0.90 for all cut-points) and cover almost the whole spectrum of PA intensities. For the dominant wrist, and working with counts data, we recommend Crouter et al. [123] cut-points and for the non-dominant wrist Chandler et al. [70] cut-points. If a researcher is interested in working directly with raw data, Hildebrand et al. [61] cut-points seem to be the best options since they were validated against indirect calorimetry and they obtained relatively high accuracy, except for MPA and VPA (33-80%).

For adults, 150 CPM measured using the VACounts from hip accelerations are the best option to estimate time in SB [131]. For PA intensity classification, we recommend Sasaki et al. [107] cut-points developed utilizing a linear regression equation. Staudenmayer et al. [115] and Hildebrand et al. [61] cut-points are the only alternative at the moment to estimate PA from the dominant and the non-dominant wrists respectively, considering that raw data metrics have to be used to apply them, not activity counts.

For older adults, we only found the SB cut-points proposed by Aguilar-Farias et al. [116] and the PA cut-points by Santos-Lozano et al. [117]. By combining these cut-points we can assess the whole spectrum of PA levels, which is the only option at present.

Physical activity-related energy expenditure algorithms

PAEE can be estimated using algorithms applied to GT3X/+ data. Since the same movement can produce different energy expenditure depending on the characteristics of the individuals, caution is advised when interpreting PAEE estimated from accelerometers. It is worth highlighting that PAEE algorithms developed in a laboratory or a controlled setting are influenced by the activities selected in the study, while only studies under free-living conditions using doubly labelled water as a criterion to test validity can actually measure PAEE. Thus, in this review we have only focused on studies using doubly labelled water and room calorimetry as a criterion.

The criteria considered for PAEE algorithm recommendations were: 1) free-living studies are better than lab/controlled studies; 2) for young populations, algorithms developed in short epochs; 3) whether cross-validation was performed; and 4) results obtained in comparison with the criterion.

Different generations of ActiGraph devices are not fully comparable in free-living conditions [107,174,175,199], thus, our recommendations are based on PAEE algorithms developed only with GT3X/+ accelerometers. PAEE has been expressed differently across studies which needs to be considered when choosing a suitable algorithm. Overall, our conclusion is that more validation studies during free-living conditions utilizing doubly labelled water are needed in all age groups.

Abbreviations in this page:
LFE: low-frequency extension

For pre-schoolers, we recommend to use the algorithm proposed by Butte et al. [128] as it is the only one validated in free-living conditions against doubly labelled water and they obtained a high accuracy. There are no algorithms for wrist accelerations in this age group. Likewise, we do not recommend algorithms for the rest of age-groups since none of them were developed using doubly labelled water or room calorimetry as a criterion.

Sleep-related behaviours

The ActiGraph GT3X/+ can identify sleep-related behaviours from movement/non-movement patterns by applying sleep algorithms to activity counts. The overall conclusion is that more studies developing and validating sleep algorithms for the wrist and the hip mounted ActiGraphs are needed in all age groups. However, based on the aforementioned information and the ages of the samples in the validation studies for sleep algorithms, we recommend use of the Barreira et al. [111] algorithm in children and adolescents when the accelerometer is attached to the hip to document bedtime that is not reported by the participants. Also, we recommend the Sadeh et al. [38] algorithm if it is placed on the wrist to score the sleep time and the rest of sleep-related behaviours. Moreover, we recommend the Sadeh et al. [38] algorithm for young adults (i.e., up to 30 years of age) and Cole-Kripke et al. [39] algorithm for older adults (i.e., >30 years of age) (in both cases with the accelerometer placed on the wrist).

Step counting

We have reviewed studies using the default step counting function by ActiLife. In this regard, we recommend to use the normal filter when the step count is a variable of interest, as it has demonstrated to be more comparable to other criterion devices than the LFE filter.

Limitations and Strengths

Several limitations need to be acknowledged. Studies with earlier models than GT3X/+ (e.g., GT1M) have not been included in our review, so our recommendations are limited to the triaxial ActiGraph models (GT3X/+). Another limitation is that for certain age groups and for some accelerometer criteria analysed, the number of studies were small; therefore, the recommendation should be revisited when more studies on those topics are available. In addition, the field of accelerometry is rapidly developing and continuously changing. Therefore, reviews are needed every few years in order to update the recommendations provided in this review. Another major issue is that proprietary algorithms used by the manufacturer when processing the data to obtain activity counts are unavailable to the public and these affect outputs. Future work using the raw acceleration

Abbreviations in this page:
ESM: electronic supplementary material
PAEE: physical activity-related energy expenditure

signal (i.e., before any filtering is applied) should overcome this problem. Finally, another limitation is that some of our recommendations are based on few studies and should be interpreted cautiously. Thus, further studies such as a formal meta-analysis may provide the most optimal intensity thresholds for the different intensity thresholds.

The strengths of this review are: 1) the inclusion of a large number of studies, summarizing the methodologies used in each of them, which will allow for more accurate comparability of the data; 2) the separate sections for the validation/calibration studies in order to provide guidance and recommendations to researchers and practitioners; 3) the inclusion of all age groups in one single review, which will allow researchers to find/read the information about the age group they are working with/interested in; and 4) the set of tables included in this review were developed to assist researchers in their decision making process (see the examples included in the Practical Implications section).

Practical Implications

This review will help researchers and practitioners to make better decisions when designing their study and processing the data from the GT3X/+ accelerometer in order to obtain the most accurate and comparable information. Here, we provide some hypothetical examples illustrating how the information presented in the tables in this review can be used.

- A researcher intends to evaluate accelerometry in a new study and needs to know where to place the accelerometer. **Table 5** summarizes the most important results obtained when comparing the outputs from the GT3X/+ attached to the hip versus the wrist and provides recommendations depending on the variables to be analysed by age group.
- A researcher has collected accelerometry data and wishes to compare the data with those from other studies to generate an accurate and meaningful discussion. **Table 4** lists the criteria used for data collection protocols and data processing in studies. **ESM 3** lists all articles that have been used for each of these criteria.
- A researcher has collected accelerometry data with the device placed on the wrist (for example) and wishes to know which cut-points, PAEE or sleep algorithms can be applied to those data. **Table 6** and **Table 7** will help the researcher answer these questions.
- A researcher has decided to apply a specific set of cut-points based on the characteristics of his/her sample but is uncertain which exact setting was used in the original study (and is aware that it is recommended that the same settings be used

ESM 3 (scan or click here)

to ensure the cut-points are applied correctly to the new data). **Table 7** lists all criteria needed to correctly apply these cut-points (i.e., placement, filter, vector and epoch).

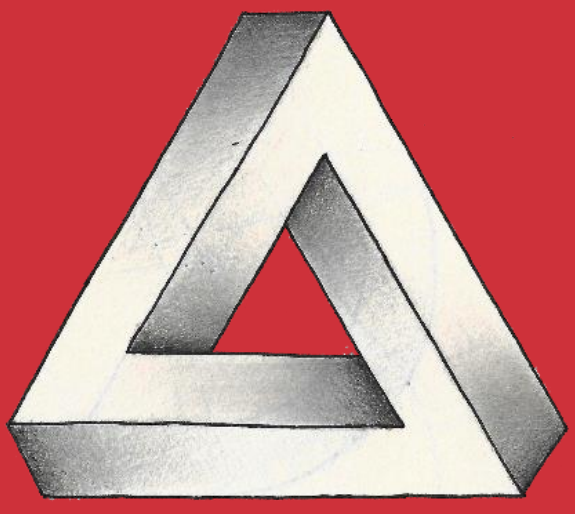
Conclusion

We suggest that researchers who assess time in SB, PA, PAEE, sleep-related behaviours and/or steps using GT3X/+ select the specific placement, sampling frequency, filter, epoch length, non-wear time definition, valid days and valid week criteria, SB and PA intensity classification, PAEE and sleep algorithms depending on the population's age (i.e., pre-schoolers, children and adolescents, adults or older adults). Likewise, when selecting a specific cut-point or algorithm, it is important to apply the same criteria as in the original validation/calibration study. Moreover, this review has identified some issues in the studies using the GT3X/+ during the last 5 years, such as that many studies do not report all of the criteria used in their analyses (see **Figure 9**). Future studies are recommended to report the criteria as summarized in the present review.

Although ideally researchers should select all the data collection and processing criteria before the assessment period, it is important to note that only the placement and sampling frequency criteria have to be decided *a priori* (i.e., before the measurement period), while the rest of processing decisions can be made *a posteriori*. This is important since new and better analytical methods might emerge after a study was planned, and they should be considered and tested, at least as sensitivity analyses. The preliminary evidence comparing wrist and hip placements seems to support the idea that a similar compliance can be achieved wearing the accelerometer on the wrist or on the hip, while wearing it on the hip might produce more accurate estimates of PAEE and better time in SB and PA intensity classifications; however, these notions need to be confirmed or refuted in future studies. We recommend to record raw data for complete days (i.e., 24h periods), so that collected data will have the maximum potential for future analyses. The summary tables presented in this systematic review will help researchers to make better decisions on how to design and process the GT3X/+ data.

Abbreviations in this page:
CPM: counts per minute
ENMO: Euclidean Norm minus 1*G*
LPA: light physical activity
MPA: moderate physical activity
PA: physical activity
ROC-AUC: receiver operating characteristic area under the curve
SB: sedentary behaviour
VM: vector magnitude
VPA: vigorous physical activity

Study II



GGIR: a research community-driven open-source R package for generating physical activity and sleep outcomes from multi-day raw accelerometer data

Migueles JH, Rowlands AV, Huber F, Sabia S, van Hees VT.
J Meas Phys Beh. 2019 Jul;2:188-196
DOI: 10.1123/jmpb.2018-0063

Contents

Abstract

Background | Recent technological advances have transformed the research on PA initially based on questionnaire data to the most recent objective data from accelerometers. The shift to availability of raw accelerations has increased measurement accuracy, transparency, and the potential for data harmonization. However, it has also shifted the need for considerable processing expertise to the researcher. Many users do not have this expertise. The R package GGIR has been made available to all as a tool to convert multi-day high resolution raw accelerometer data from wearable movement sensors into meaningful evidence-based outcomes and insightful reports for the study of human daily PA and sleep.

Aim | This paper aims to provide a one-stop overview of GGIR package, the papers underpinning the theory of GGIR, and how research contributes to the continued growth of the GGIR package.

Results | The package includes a range of literature-supported methods to clean the data and provide day-by-day, as well as full recording, weekly, weekend, and weekday estimates of PA and sleep parameters. In addition, the package also comes with a shell function that enables the user to process a set of input files and produce csv summary reports with a single function call, ideal for users less proficient in R. GGIR has been used in over 90 peer-reviewed scientific publications to date.

Conclusion | The evolution of GGIR over time and widespread use across a range of research areas highlights the importance of open source software development for the research community and advancing methods in physical behaviour research.

Abbreviations in this page:
PA: physical activity

Key Points

Question
Is there any alternative to closed commercial software for accelerometer data analysis in the PA field?

Findings
The GGIR open-source software is able to convert accelerometer raw data into meaningful PA, SB, and sleep indicators from several accelerometer brands (ActiGraph, GENEActiv, Axivity, among others)

Meaning
The GGIR software presented in this paper facilitates the processing and extraction of insightful PA and sleep variables of the data collected with these so-called raw data accelerometers.

Introduction and motivation

Abbreviations in this page:
PA: physical activity

Human PA and sleep are popular areas of research because of their important role in health outcomes [10,200]. PA and sleep have traditionally been quantified with diaries and questionnaires, but wearable sensors have gained momentum since the 1990's. In the beginning, wearable movement sensors (i.e., accelerometers) typically performed onboard signal processing and only stored derived output to reduce battery consumption and memory requirements. However, following a general movement towards more transparent and open science and thanks to technological evolution towards smaller, cheaper and power efficient sensors, accelerometers now tend to store 'raw' data for offline processing and analysis. The data recorded are typically expressed in gravitational acceleration (g) because this is the reference point for acceleration value calibration, reflecting both the movement and gravitational component [67]. However, this technological advance is counterbalanced by the large amount of data collected per measurement (typically $2 \cdot 10^8$ data points per week of measurement) and the necessity to process the data to obtain meaningful variables that can be used in standard statistical analysis and software. Many PA and sleep researchers do not have the expertise necessary to process and analyse raw accelerometer data. The **GGIR software** presented in this paper facilitates the processing and extraction of insightful PA and sleep variables of the data collected with these so called raw data accelerometers from three widely used sensor brands [41]. The individual algorithms that are embedded in GGIR have been described across a number of published papers. The fast advances in wearable sensor technology over the last decades comes with the price of mandatory development of scientific software to ensure a good valorisation of the newly available sensors [201]. However, scientific software instruments need to be subjected to a peer review process as it is the case for other methodological components (e.g., algorithms and study protocols). Previous publications related to GGIR focused on specific algorithms, such as sleep detection, and their validity. However, those publications did not focus on GGIR as a generic piece of scientific software that connects all these algorithms and adds a range of essential extra functionalities, e.g., time zone and daylight-saving time handling, to provide value to an entire research community far beyond the specificities of those studies. Therefore, this paper aims to provide a one-stop overview of GGIR package, the papers underpinning the theory of GGIR, and how research contributes to the continued growth of the GGIR package.

GGIR vignette
(scan or click here)

How open-source software, e.g., GGIR, can contribute to advances in the field of PA measurement

Abbreviations in this page:
PA: physical activity

GGIR contributes to scientific discovery by enabling researchers to study (human) PA and sleep using accelerometers without prerequired programming expertise. GGIR is appropriate for use across a wide variety of study designs, e.g., variations in measurement duration, in sample frequency, instructions given to the participant, and study populations.

GGIR publication list
(scan or click here)

Applications of GGIR have been reported in over 90 peer reviewed journal publications since its first release in 2013, with 24 in 2017 and 48 in 2018 based on looking up the citations to the key journal publications underlying GGIR. Additionally, nine methodological papers were written to motivate and evaluate parts of GGIR. Previous accelerometer software has been commercial (e.g., Actilife) and/or restricted to one brand of accelerometer (e.g., Actilife, GENEActiv PC combined with Excel macros). GGIR facilitates the processing and analysis of data from three of the most widely used brands of research-grade movement sensors (GENEActiv by ActivInsights Ltd, ActiGraph by Actigraph LLC, and Axivity by Axivity Ltd) using open-source generic brand agnostic methods, potentially providing a means for harmonization of data from large surveys globally [10]. Further, GGIR is continually updated to include innovations developed by the software team and/or users expediting the application of novel analytics in research [66,202,203]. GGIR is very much a community driven development as testified by: the multiple contributors from both health research and technological backgrounds; the fact that many publications that use GGIR are not co-authored by the development team; the existence of a support and maintenance service by V. van Hees where GGIR-users can hire Vincent's time as freelancer to help address specific needs from the user-community (www.movementdata.nl). This service has fuelled a range of package upgrades, and is one of the possible ways to sustain open source software like GGIR, and; the availability of a user-forum to make it possible for users to reach out to each other (see **Table 8**).

Previously, accelerometers were typically either used to assess waking PA or sleep and circadian rhythms, but rarely tailored for both research areas. The new generation of accelerometers, worn night and day, allows the measurement of both PA and sleep using the same tool. Moreover, in several studies such as UK Biobank [204], a sleep diary was not used to define a sleep window as commonly done in sleep research. The GGIR package allows differentiation of waking from sleep periods and provides sleep quality parameters. Using the GGIR package, the Colaus study reported better sleep efficiency among those more physically active [205]. Several papers using UK Biobank data have now been published and show the

advantage of using accelerometer instead of questionnaire data to identify for example the genetics of sleep [206–208].

Table 8
Key metadata on the current version of GGIR

Current code version	1.6-7
Permanent link to code/repository used of this code version	https://github.com/wadpac/GGIR
Legal Code License	LGPL (≥ 2.0, < 3)
Code versioning system used	Git
Software code languages, tools, and services used	r, C++, Travis-CI
Compilation requirements, operating environments & dependencies	64-bit operating system & R environment version 3.2.3 and up (64-bit) & R packages: MASS, signal, zoo, mmap, bitops, matlab, GENEAread, tuneR, testthat, covr, knitr, rmarkdown, data.table, Rcpp
If available Link to developer documentation/manual	https://cran.r-project.org/web/packages/GGIR/GGIR.pdf
User forum	https://groups.google.com/forum/#!forum/rpackageggir
Works on Operating Systems	Windows, Linux, and OSx

Here are some further examples of research studies that were facilitated by the use of GGIR. Results from the Whitehall II study showed: 1) the association between PA and BMI was much stronger when using accelerometer data rather than questionnaire data [203], 2) obese people with an unfavourable metabolic profile had a lower level of PA than those with a healthy metabolic profile, which was not evident when using questionnaire data [209], 3) the association between MPA and VPA and healthy ageing was evident whether the activity was performed in short (<10 minutes) or long (≥10 minutes) bouts [210]. Results from the UK Biobank also showed: 1) people with cardiometabolic disease are less physically active and tend to engage more in SB that last more than 30 minutes [211], 2) short bursts of very high intensity PA are associated with better bone health in pre- and post-menopausal women [212].

Abbreviations in this page:
BMI: body mass index
LGPL: Lesser General Public License
MPA: moderate physical activity
VPA: vigorous physical activity
PA: physical activity
SB: sedentary behaviour

A typical experimental setting

Studies vary in size from a few dozen participants in clinical or methodological studies [213], a few thousand in epidemiological cohorts [214,215], and a hundred thousand in biobanks like UK Biobank [204]. The researcher configures an accelerometer with the desktop software supplied by the accelerometer manufacturer. Next, the accelerometer is given or sent to the participant who wears it on their wrist or other body location (depending on instructions) day and night for usually seven consecutive days, although different measurement periods have also been reported. After the period of wear, the participant returns the accelerometer either in person or by post. The data file is then downloaded with the same desktop software from the accelerometer manufacturer. The file size typically ranges between 0.5 and 1.5 GB depending on specific format, sensor brand, and recording duration. GGIR can either run on a local computer or integrated in parallel processing tasks on a computing cluster when working with large numbers of datafiles. GGIR generates

Abbreviations in this page:
BMI: body mass index
GPL: General Public License
LGPL: Lesser General Public License
MPA: moderate physical activity
VPA: vigorous physical activity
PA: physical activity
SB: sedentary behaviour

reports in csv-format that can then be loaded in the researcher's preferred statistical software (e.g., R, SAS, SPSS, or Stata).

Related work

Software

A few other software tools exist to work with multi-day raw accelerometer data (**Table 9**). Actilife (ActiGraph, Pensacola, FL, USA) is a closed source commercial software designed for the accelerometer hardware developed by the same company. The GENEActiv PC software by the developer of the GENEActiv (ActivInsights Ltd, Kimbolton, UK) in combination with freely available Excel macros (available from https://open.geneactiv.org) is designed for the accelerometer hardware developed by the same company. GENEAclassify is an open-source R package primarily aimed at facilitating the segmentation and classification of accelerometer data produced by the GENEAactiv accelerometer [216]. OMGui by Dan Jackson and colleagues (Open Movement, Newcastle University, UK) is an open source GUI implemented in C# and developed for the Open Source hardware AX3 [217]. In addition to the monitor's own binary .CWA format, OMGUI can also create and load brand agnostic csv and audio WAVE file formats [218]. Pampro by Tom White (Cambridge, UK) is implemented in Python, available on GitHub and designed to process data from multiple wearable sensor brands and data formats using methodologies replicated from publications, including the publications describing parts of GGIR [219]. Pampro includes an Open Source license BiobankAccelerometerAnalysis [220] by Aiden Doherty (Oxford, UK) and colleagues is an open source tool designed to provide a minim set of basic outcomes for the UK Biobank accelerometer data, utilizing some of the key algorithms from GGIR [204]. The BiobankAccelerometerAnalysis software has recently been enhanced with activity type classification for Axivity AX3 .cwa accelerometer files [221].

Table 9
Related software

	GGIR	Actilife	GENEActiv PC + macros	GENEAclassify	OMGUI	Pampro	Biobank-Accelerometer-Analysis
Executable version online	YES	NO	YES	YES	YES	YES	YES
Source code available online	YES	NO	NO, but YES for macros	YES	YES	YES	YES
Open source software (OSS) license	LGPL (copyleft) OSS license	No OSS license	NO OSS license	GPL (copyleft) OSS license	Custom (permissive) OSS license	OSS license	BSD 2-Clause License (permissive)
Primary programming language	R	Unknown	Unknown	R	C#	Python	Java and Python

Literature

The main publications underlying GGIR described: 1) the potential of using wrist worn raw data accelerometers for estimating human daily energy expenditure [222], 2) how raw acceleration data can be meaningfully aggregated [67], 3) a method to calibrate the acceleration signals based on the recording itself with no need for additional data collection [66], 4) methods for sleep detection when the accelerometer is worn on the wrist with [40] and without [202] use of a sleep diary, and 5) comparisons with other research methods [212,214,223–225].

Abbreviations in this page:
BMI: body mass index
GPL: General Public License
LGPL: Lesser General Public License
MPA: moderate physical activity
PA: physical activity
SB: sedentary behaviour
VPA: vigorous physical activity

Software description

Software architecture

R package GGIR has been released with an Open Source LGPL 2 license on CRAN the central repository for R packages since 2013 and on GitHub [www.github/wadpac/GGIR] since December 2016. GGIR can run on Windows, Linux, and OSx (this and additional key facts about GGIR are summarised in **Table 8**).

The package comes with the following core functionalities: load data; extract signal metrics (also called signal features); detect when the sensor was not worn and impute these periods if requested by the researcher; detect the sleep period time window and sleep episodes within it; segment the data according to conventional heuristic threshold techniques; specify which parts of the recording should be considered based on the researchers knowledge about the study design (e.g., participant started wearing the accelerometer for a certain number of hours after starting the record); and finally to store convenient summary reports in csv and pdf format (**Figure 10**).

The package is split in five parts that group functionalities in logical processing order and in line with how the package historically evolved. The parts are numbered from 1 to 5 and the corresponding function names refer to the part number: *g.part1*, *g.part2*, *g.part3*, *g.part4*, and *g.part5*. The parts should be executed sequentially with milestone data automatically being stored in between parts to facilitate re-processing parts without having to go back to the original raw input data. One shell function allows for interacting with the five parts and all underlying functionalities from one central point: *g.shell.GGIR*. The shell function takes all arguments from the five parts mentioned above. In this way, the users can interact with GGIR from one single function call to function *g.shell.GGIR* and easily share their call to allow for replicating the analysis on a different dataset. The duration of GGIR depends on computer specifications, input arguments, and data characteristics. Part 1 (*g.part1*) is the most time consuming taking up at least 80% of the processing time and lasting

around 10 minutes when applied to a 7-day data file using GGIR's default argument settings.

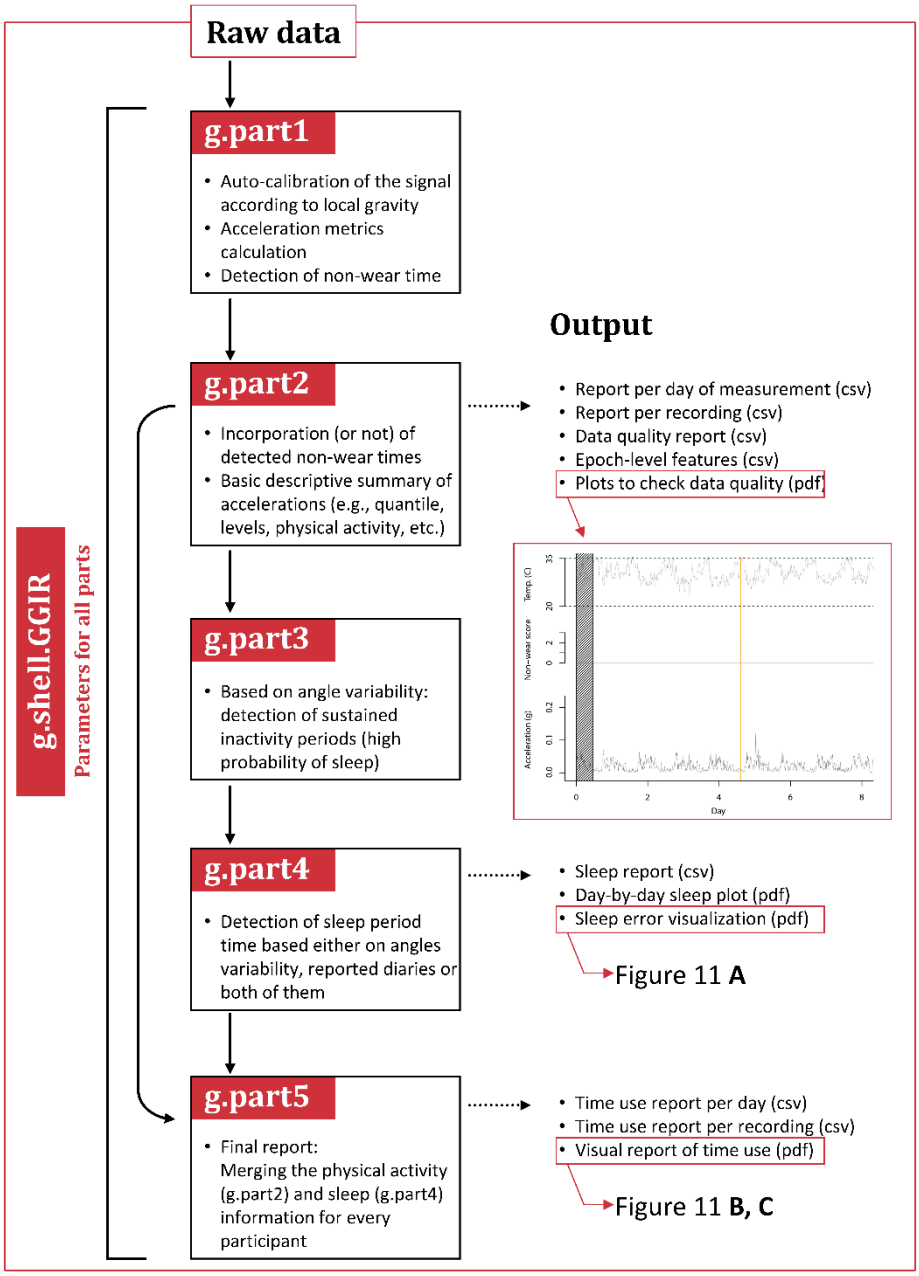


Figure 10
Overview of main steps and output in GGIR workflow

The code builds a folder structure with a depth of two to store the milestone data per participant in .RData files (including collections of data.frame and vector objects) and the analysis reports per dataset (see **Figure 10**). The analysis reports are in .csv format to give the user the flexibility to process their quantitative results in other statistical or data processing environments. There are many variables stored in the reports, an introduction to these variables can be found in the **package vignette**, while a more elaborate discussion is found in the **package manual**.

GGIR manual
(scan or click here)

In GitHub the code is supported by continuous integration with Travis-CI (https://travis-ci.org/wadpac/GGIR). Starting in 2017 we have adopted the habit of writing unit tests, but as it was not done before 2017 not all parts of the code are covered by tests yet and current test coverage is 63%. In addition to unit tests, the development

has typically gone in close collaboration with GGIR end-users who checked code validity by close examination of the package output. We have created one video to introduce GGIR and a second video to provide a visual summary of how GGIR is typically used.

Abbreviations in this page:
ENMO: Euclidean norm minus 1*G*
PA: physical activity

Software dependencies

Most of the code is written in R, with a small part in C++ needed for reading the binary data from the AX3 accelerometer brand using R package *Rcpp* as a dependency [226]. R packages *bitops* and *matlab* are used to enable reading the binary data from the Genea accelerometer (the non-commercial precursor to the GENEActiv, Unilever Discover, UK). Dependencies are R packages GENEAread and mmap, which are used for reading GENEActiv accelerometer binary data. Finally, R package *tuneR* [227] is used to read wav format data, which is an optional export format for AX3 accelerometer data. The R package *signal* is used for frequency filtering [228], R package *zoo* is used for calculating a rolling median, sum and mean [229], and R package *data.table* is used at some steps to efficiently handle large amounts of tabular data.

GGIR introduction video
(scan or click here)

Software functionalities

Part 1 (g.part1)

The *g.part1* function searches the specified data directory recursively for files that could possibly represent acceleration data. Next, it automatically detects which accelerometer brand the file belongs to, the data format in which it is stored (.csv, .bin, .cwa, .wav), and extracts the file header using appropriate functions. Next, function *g.calibrate* is used to investigate calibration error which results in proposed correction coefficients as motivated and described in van Hees et al. 2014 [66]. Then, metrics essential for sensor wear detection, PA and sleep analysis are extracted from the raw data. Here, the user can choose one or multiple aggregation metrics out of a collection of most common metrics, e.g., ENMO, and control the window size over which the metrics are calculated. Additionally, a standard set of metrics is extracted per long time window (default 15 minutes) which are needed for the detection of accelerometer non-wear. The data loading and metric extraction takes place in approximately 24-hour blocks since putting the full file content in computer memory may not always be possible. The data window size is modifiable in case the 24-hour blocks are still too large. Additionally, the code evaluates available memory throughout the processing and shortens the window by 20% when available memory is getting too low. At the end of part 1 the signal metrics are stored as milestone data in an RData file (using a filename corresponding to the input accelerometer data file). For most of the analysis we use POSIX format for timestamps, which is the default in R, but for the exported time series

GGIR visual summary
(scan or click here)

Abbreviations in this page:
PA: physical activity

we transform these to ISO8601 format to facilitate usage in other software environments. Further, GGIR takes into account day saving time and time zone. The user is expected to explicitly provide time zone of the recording, this to avoid confusion about where the experiments took place, which may not always be the time zone of the machine on which the analyses are run, or the default time zone assumed by GGIR (UTC: Europe/London).

Part 2 (g.part2)

As part of *g.part2* function *g.impute* takes the milestone output from *g.part1*, optionally identifies unreliable signal sections (e.g., monitor not worn or signal clips near its extreme) and replaces these sections by imputed values (average of same time point on all other days of the measurement). Then, function *g.analyse* performs a descriptive analysis of the output and summarizes it per measurement, per day of measurement and conveniently per day type (i.e., weekdays and weekend days separately) as well as per segment(s) of each day (specified by user with argument *qwindow*). For example, if *qwindow* has value 'c(0,9,12,17,24)' the summaries will be derived for the time windows 0:00-24:00, 0:00-9:00, 9:00-12:00, 12:00-17:00, and 17:00-24:00. This could for instance be useful in research on the level of PA during specific segments of the day in children and workplace interventions in adults. Examples of summaries generated per time segment are time spent in specific acceleration magnitude ranges (bins), average acceleration metric value, and the timing of the least and most active five-hour window of the day (number of hours can be set by user). The user can also provide important knowledge about the experimental design, which is accounted for when generating the descriptive summaries, e.g., instruct GGIR to ignore the first hour of the measurement or all data before the first and after the last midnight of the recording. In addition, the descriptive summaries take into account the quality of the measurement. For example, the user can specify the necessary minimum number of valid hours per day. Days with fewer valid hours of data will be ignored in the person level descriptive summaries.

The reason why *g.part1* and *g.part2* are not merged is because *g.part1* takes much longer to run and involves only minor decisions of interest to the PA or sleep researcher. Function *g.part2* on the other hand is relatively fast and comes with all the decisions that directly impact on the variables that are of interest to the researcher. Therefore, the user may want to run *g.part1* overnight or on a computing cluster, while *g.part2* can then be the main playground for research.

Part 3 and 4 (g.part3 and g.part4)

The functions *g.part3* and *g.part4* provide functionality for estimating sleep when the accelerometer was worn on the wrist as

described in van Hees et al. [40,202]. Although these functions can be applied to accelerometer data from any wear location, currently no scientific literature exists to support the interpretation for attachment locations other than the wrist. In short, *g.part3* detects time episodes with a sustained lack of change in arm angle, which are referred to as sustained inactivity bouts. The user can configure the parameters used for this and can specify multiple parameter values to facilitate comparisons. The *g.part3* function only generates milestone data to facilitate the work done in *g.part4*. Output from *g.part3* for example includes the start and end times of the sustained inactivity bouts, and estimated start and end of the sleep period time window. Then, the *g.part4* function gives the user the option to either use a sleep diary or to rely on the estimated sleep period time window from *g.part3* [202]. Another option is to specify a general time window when individuals within the study population are assumed to be in bed, which is probably more naïve and primarily implemented for reference purposes. Sustained inactivity bouts that occur within the sleep period time window are considered sleep episodes, and sustained inactivity bouts outside the sleep period time window are considered rest, potential nap period, or undetected short episodes of monitor non-wear time. Further, *g.part4* offers the user the option to exclude the first and last night. Both *g.part3* and *g.part4* store a record of the amount of available valid and invalid data per night, and an indicator of whether sleep diary data was available for each specific night used. To facilitate a relatively quick inspection, sleep detection plots of the classification are exported as pdf files. To facilitate identifying obvious mistakes in sleep diary data *g.part4* provides the option to visualise the differences between accelerometer-based estimates and sleep diary; this was for example used in van Hees et al [40] to identify a dozen of problematic nights out of 27000 nights (**Figure 11**, panel A).

Abbreviations in this page:
LPA: light physical activity
MPA: moderate physical activity
VPA: vigorous physical activity

Part 5 (g.part5)

The *g.part5* function takes the output from *g.part2* and *g.part4* to describe time spent in 16 time use classes, composed of: nighttime or daytime behaviour; sustained inactive or sleep, other inactive, LPA, MPA, or VPA behaviour, and; un-bouted, short, medium or long bouts of behaviour. The number of 16 classes is the default, this can be adjusted. Next, the time series of epoch level classes are segmented by day based on one or both available definitions of a day: 1) From waking up to waking up the next day, which means that a full night of sleep is included per day and that the duration of days vary, or 2) From midnight to midnight which equals a calendar day, meaning that a night of sleep is likely to be split between days. In the latter definition days are by default assumed to end at midnight, but the user also has the option to change the hour on which the day ends.

Abbreviations in this page:
LPA: light physical activity
MPA: moderate physical activity
PA: physical activity
VPA: vigorous physical activity

From these daily segments *g.part5* calculates the time spent in these classes, the number of occurrences of these classes, and the average acceleration within the classes.

The thresholds to describe level of PA level intensity (LPA, MPA, and VPA) have their origin in energy expenditure research, but in GGIR these levels should be interpreted as thresholds to mark ranges in acceleration values. However, if the user specifies the thresholds wisely he/she may be able to interpret the levels as indicators of energy metabolism. The concept of behavioural bouts may also need clarification: *g.part5* calculates the total time spent in behavioural bouts, which is a time segment that meets user-specified criteria on the range of acceleration, the percentage of time during which those criteria on the range need to be met, and the maximum duration of breaks allowed to define a bout. The user selects one of four metrics for bout calculation using argument *bout.metric*. This functionality overlaps somewhat with *g.part2* which also calculates time spent in bouts, but *g.part5* does it in a much more advanced way with also attention to inactivity bouts, LPA bouts, as well as unbouted behaviour and sleep. All this information is stored in csv files and visual summaries are stored in pdf files (**Figure 11**, panels B and C). The *g.part5* function offers the user a large freedom to explore multiple parameters simultaneously (thresholds, minimal and maximal bout durations and percentage criteria to define a bout) resulting in potentially hundreds or even thousands of output variables to enable the study of physical behaviour.

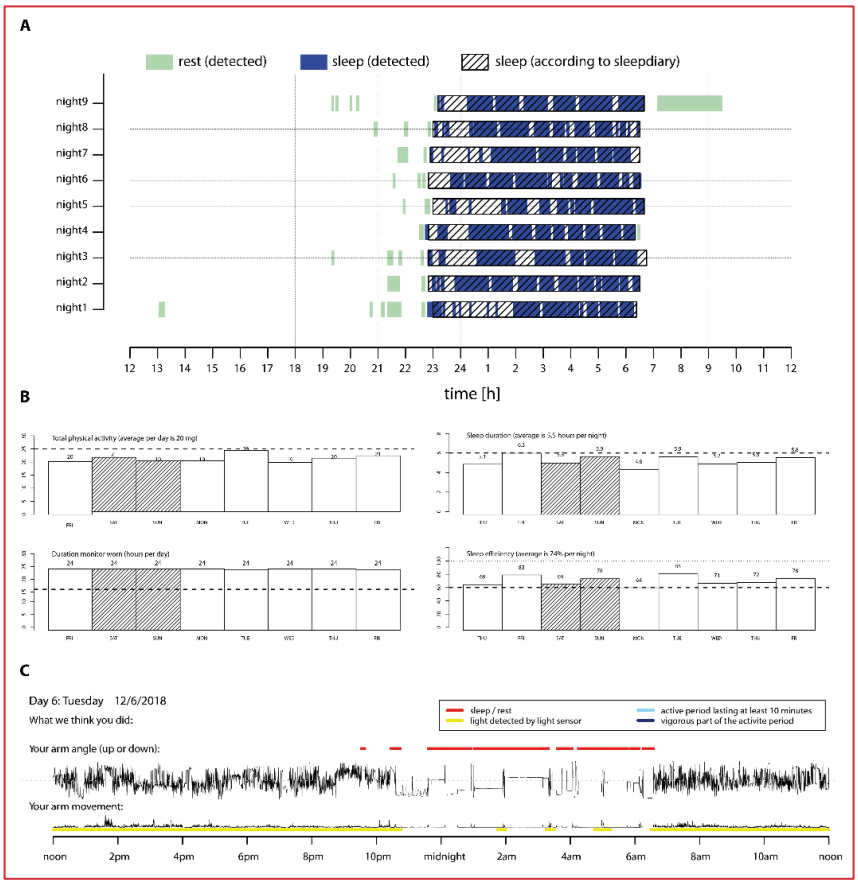


Figure 11
Visual output reported by GGIR.
A) Rest and sleep detected from acceleration features and sleep diary (g.part4).
B) Bar plots with information on key physical activity and sleep variables (g.part5).
C) Visual summary of the physical activity and sleep patterns (g.part5).

Impact

The main value of GGIR is that it offers a broad set of functionalities ranging from data quality handling to 24hours/7days time use characterization of PA and sleep utilizing literature supported methods and is study population agnostic making it suitable for a wide range of research areas. GGIR can be operated without significant prior programming expertise.

At the same time, the user experiences freedom in the specification of input arguments and the selection of output variables. Being fully open source, GGIR can be adapted and extended according to the needs of the respective research project. GGIR is currently being used by the PA and sleep research communities, with over 90 peer-reviewed journal papers published using the software since 2013, with the number published per year increasing rapidly.

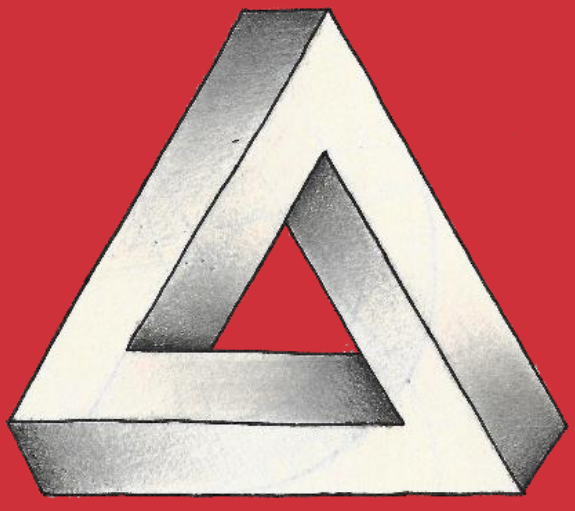
By facilitating analysis of raw accelerometer data, GGIR enables the analysis of objective measures of PA and sleep and encourages raw data accelerometers to be used more readily in research studies that aim to understand the importance of PA and sleep for human health. Moving away from closed commercial accelerometer software and self-report questionnaires to assess PA and sleep will improve knowledge, allowing better translation in public health recommendations.

Abbreviations in this page:
PA: physical activity

Conclusion

Most of the knowledge on PA and sleep comes from studies using questionnaire data or, more recently, proprietary PA metrics generated using commercial software. With the rapid advancement in technology over recent decades, raw acceleration movement sensors can now be used in large scale studies. However, the data they generate are not straightforward to analyse. GGIR provides a tool for researchers to derive variables that characterize PA and sleep assessed in an objective manner. In addition, as the code is open source, it can be used in part or as a whole making it flexible to research needs. It further facilitates a reproducible analysis of the raw data which is key to generating conclusions in clinical and observational research settings. Previously, widely used software packages for analysing accelerometer data have been proprietary and there has been little opportunity for most researchers to feed into how they can be developed. This demonstrates the importance of open source software development for the research community and for the advancement of methods in physical behaviour research.

Study III



Accelerometer data processing and energy expenditure estimation in pre-schoolers

Migueles JH, Nyström CD, Henriksson P, Cadenas-Sanchez C, Ortega FB, Löf M.
Med Sci Sports Exerc. 2019 Mar;51(3):590-598
DOI: 10.1249/MSS.0000000000001797

Contents

Abstract

Background | Given that PAEE is the most malleable component of TEE, any researcher and health care professional dealing with energy balance and PA is interested in accurate estimations of PAEE. Accelerometer could provide such estimates. To the best of our knowledge, no study has compared the performance of ActiGraph's activity counts and alternate summary metrics in the assessment of free-living TEE and PAEE.

Aims | To assess the capacity of different acceleration metrics from wrist accelerations to estimate TEE and PAEE using doubly labelled water in preschool children.

Methods | Thirty-nine pre-schoolers (5.5 T 0.1 yr) were included. TEE was measured using doubly labelled water during 14 d, and PAEE was then calculated using a predicted basal metabolic rate. Participants wore a wGT3X-BT accelerometer on their nondominant wrist for ≥5 d. We derived the following metrics from raw accelerations: VMCounts and LFECounts; and alternate summary metrics, such as ENMO, Euclidian norm of the high-pass-filtered accelerations (HFEN), the bandpass-filtered accelerations, the HFEN plus Euclidean norm of low-pass filtered accelerations minus 1g (HFEN+) and the mean amplitude deviation (MAD).

Results | Alternate summary metrics explained a larger proportion of the variance in TEE and PAEE than ActiGraph's activity counts (counts, 7–8 and 25% of TEE and PAEE; alternate summary metrics, 13%–16% and 35%–39% of TEE and PAEE). Adjustments for body weight and height resulted in an explanation of51% of PAEE by ENMO. All of the metrics adjusted for fat mass and fat-free mass explained up to 84% and 67% of TEE and PAEE, respectively.

Conclusion | ENMO and the other alternate summary metrics explained more of the variance in TEE and PAEE than the ActiGraph's activity counts in 5-yr-old children, suggesting further exploration of these variables in studies on physical activity and energy expenditure in pre-schoolers. Our results need confirmation in other populations with wider age groups and varying body compositions.

Abbreviations in this page:
ENMO: Euclidean norm minus $1G$
HFEN: high-pass filtered Euclidean norm
$HFEN_+$: HFEN plus the Euclidean norm minus $1G$
LFECounts: activity counts with the low-frequency extension filter in the vector magnitude
MAD: mean amplitude deviation
PA: physical activity
PAEE: physical activity-related energy expenditure
TEE: total energy expenditure

Key Points

Question
What is the capacity of wrist-worn accelerometers to estimate TEE and PAEE in pre-schoolers?

Findings
Open-source metrics explained more variance of TEE and PAEE than activity counts. The prediction capacity was moderate for the metrics, and high when body composition variables were added to the models

Meaning
Any researcher or practitioner interested in the energy balance in pre-schooler may consider wrist-worn accelerometers as a proxy to TEE and PAEE

Introduction

Energy imbalance generated by high-calorie intakes and low PAEE has led the current population to an obesity epidemic, which has caused an array of health consequences throughout the lifespan [230]. Furthermore, given that PAEE is the most malleable component of TEE, any researcher and health care professional dealing with energy balance and PA is interested in accurate estimations of PAEE. Doubly labelled water is the gold-standard method to assess TEE and PAEE, but its high cost makes it unfeasible in large populations [231]. Instead, accelerometers, which are widely available and less expensive, allow for the measurement of PA with the potential to assess PAEE [41,61,94,232–235].

Abbreviations in this page:
NHANES: National Health And Nutrition Examination Survey
PA: physical activity
PAEE: physical activity-related energy expenditure
TEE: total energy expenditure

Previous studies have shown poor estimations of TEE and PAEE using accelerometers, only explaining between 14-31% of the variation in PAEE in preschool aged children [95,234,235]. Methods proposed to improve these estimations include: (i) the combination of heart rate monitoring with accelerometers, which does not seem to improve estimations when using comparable acceleration metrics [236,237] or (ii) the inclusion of body composition measures in the models, which has substantially improved accelerometer estimations of free-living TEE and PAEE [95,234].

The increasing use of accelerometers in epidemiological studies has resulted in a myriad of ways to collect and process accelerometer data [233]. In regards to data collection, the majority of studies have attached accelerometers to the hip or waist looking for a proximal placement to the body's centre of gravity. However, some large-scale studies such as the NHANES and the UK biobank have placed accelerometers on the wrist due to the belief it promotes a higher wear compliance [171,233], which makes the study of wrist placement to estimate TEE and PAEE a matter of interest. In regards to data processing, previous studies usually processed raw accelerations using manufacturer proprietary filters to obtain activity counts or steps [95,235,238]. These manufacturer-processed counts and steps are not comparable between devices [175,199], halting the progression of the accelerometer estimates of PA and PAEE. New devices store raw accelerations, which has made it possible to obtain alternative acceleration metrics to activity counts and steps using open-source filtering methods [67,147]. Although these new metrics (hereinafter referred to as alternate summary metrics) seem to be comparable across certain devices, they are not necessarily comparable [225,239]. For example, Rowlands et al. suggested that the ActiGraph GT9X raw accelerations are slightly lower than other brands, although more research is needed to confirm this [225].

To the best of our knowledge, no study has compared the performance of ActiGraph's activity counts and alternate summary

Abbreviations in this page:
BFEN: band-pass filtered Euclidean norm
BMR: basal metabolic rate
ENMO: Euclidean norm minus 1*G*
HFEN: high-pass filtered Euclidean norm
$HFEN_+$: HFEN plus the Euclidean norm minus 1*G*
LFECounts: activity counts with the low-frequency extension filter in the vector magnitude
MAD: mean amplitude deviation
MINISTOP: Mobile-based INtervention Intended to STop Obesity in Pre-schoolers
PAEE: physical activity-related energy expenditure
SD: standard deviation
TEE: total energy expenditure
VMCounts: Activity counts in the vector magnitude

metrics in the assessment of free-living TEE and PAEE. Thus, the purpose of the present study was to assess the capacity of the ActiGraph's activity counts and alternate summary metrics from wrist accelerations to estimate TEE and PAEE assessed with doubly labelled water. For this purpose, we utilized data collected in the MINISTOP trial [95,240] with data on energy metabolism and 24-hour wrist accelerations in five-year-olds.

Methods

Participants and study design

This validation study was conducted in forty parent and child dyads participating in the MINISTOP trial [240]. These 40 parents and children were recruited at the follow-up measurement at 5.5 years of age [94,95]. More information about the project is published elsewhere [94,95]. One child did not fulfill the wearing time criteria for the ActiGraph (i.e., ≥ 16 hours/day for ≥ 4 days) and thus this analysis includes 39 preschool children. The Research and Ethics Committee, Stockholm, Sweden approved the study and informed consent was collected from both parents (2013/1607-31/5; 2013/2250-32). The MINISTOP study is registered as a clinical trial (https://clinicaltrials.gov/ct2/show/NCT02021786).

Table 10
Descriptive characteristics of participants

	All (n=39)	Boys (n=21)	Girls (n=18)
Physical characteristics			
Age (years)	5.5 ± 0.1	5.5 ± 0.2	5.5 ± 0.1
Height (kg)	114.2 ± 4.5	115.0 ± 5.4	113.3 ± 3.1
Weight (cm)	20.5 ± 4.3	20.6 ± 4.3	20.3 ± 4.3
BMI (kg/m^2)	15.6 ± 2.3	15.5 ± 2.0	15.7 ± 2.6
BMI categories[a] (n, %)			
Normal weight	36 (92%)	19 (90%)	17 (94%)
Overweight	1 (3%)	1 (5%)	0 (0%)
Obese	2 (5%)	1 (5%)	1 (6%)
Fat mass (kg)	5.4 ± 2.7	5.1 ± 2.5	5.9 ± 3.0
Fat-free mas (kg)	15.1 ± 2.0	15.6 ± 2.1	14.5 ± 1.7
Energy expenditure			
TEE (kJ/24h)	6030 ± 691	6272 ± 720	5749 ± 548
PAEE[b] (kJ/24h)	1460 ± 436	1560 ± 447	1343 ± 405
BMR (kJ/24h)	3967 ± 402	4084 ± 406	3831 ± 361
Physical activity level	1.52 ± 0.12	1.54 ± 0.12	1.50 ± 0.12
Acceleration metrics			
Valid days	7.5 ± 1.1	7.3 ± 1.2	7.8 ± 1.1
Non-wear time (min/day)	7.7 ± 14.2	8.6 ± 17.4	6.6 ± 9.5
VMCounts (counts/5s)	218.3 ± 38.0	218.3 ± 44.2	218.3 ± 30.4
LFECounts (counts/5s)	225.8 ± 38.4	225.6 ± 44.7	226.0 ± 30.7
ENMO (m*g*)	47.3 ± 11.9	50.5 ± 13.8	43.5 ± 8.1
BFEN (m*g*)	161.3 ± 25.2	163.6 ± 29.3	158.5 ± 19.9
HFEN (m*g*)	162.5 ± 25.3	165.0 ± 29.4	159.6 ± 19.9
$HFEN_+$ (m*g*)	150.9 ± 25.7	155.0 ± 30.2	146.1 ± 18.7
MAD (m*g*)	61.0 ± 12.9	63.9 ± 15.0	57.7 ± 9.3

Data are presented as mean ± SD
[a] Cole et al. [64] reference standards were used to classify the participants into weight status groups.
[b] PAEE was calculated as TEE multiplied by 0.9 minus BMR to correct for dietary induced thermogenesis (commonly assumed to correspond to 10 % of TEE)

Body composition and energy expenditure estimation

The children's TEE and total body water were measured with doubly labelled water during 14 days. Details on the dosing of doubly labelled water, collection of urine samples before and after the dose, analyses of isotopic enrichments using isotope ratio mass spectrometry and calculation of carbon dioxide and total body water have been provided previously [94,95]. The Weir equation was applied to derive TEE from carbon dioxide production [99] assuming a food quotient of 0.85 [100]. For this study sample (n=39), the quotient between the N_D and the N_O was 1.039 ± 0.008. Fat-free mass was calculated from total body water assuming that fat-free mass contains 76.4% water [102]. Thereafter, fat mass was calculated as the difference between body weight and fat-free mass. Over the 14-day measurement period, no major change in body weight was observed (n=39; 0.07 ± 0.32 kg). We applied prediction equations based on weight [101] to estimate BMR. Thereafter, PAEE was calculated as TEE multiplied by 0.9 minus BMR. This includes a reduction in TEE by 10% to adjust for energy expended due to dietary induced thermogenesis.

Accelerometer data processing

PA was monitored with the wGT3X-BT accelerometer (ActiGraph, Pensacola, FL, USA) placed on the non-dominant wrist. Participants were instructed to wear the ActiGraph for the first 7 days of the 14-day doubly labelled water period (24-hours per day); however, some of them wore the accelerometer for more than 7 days, and thus, all available data were used. Devices were initialized to collect data at 50 Hz, as this sampling frequency is sufficient to capture wrist daily motion [96,97]. In order to remove the influence of gravitational acceleration and noise from the raw accelerations, we extracted the acceleration metrics as described below. We derived VMCounts and LFECounts from the ActiLife 6.2.0 software (ActiGraph, Pensacola, FL, USA) over 5 second epochs. Simultaneously, we used the GGIR package v. 1.6-12 (https://cran.r-project.org/web/packages/GGIR/) implemented in the R software (http://cran.r-project.org) to automatically export the following alternate summary metrics: *ENMO*, Euclidean Norm Minus One G (i.e., $\sqrt{x^2 + y^2 + z^2}$ – 1 g) with negative values rounded to zero after auto-calibration of the raw accelerations [67]; *BFEN*, Euclidean norm of the band-pass frequency filtered accelerations on the auto-calibrated raw accelerations with cut-off frequencies of 0.2 and 15 Hz [67]; *HFEN*, Euclidean norm of the high-pass frequency filtered accelerations on the auto-calibrated raw accelerations with a cut-off frequency of 0.2 Hz [67]; *HFEN+*, HFEN plus ENMO from the low-pass filtered accelerations [67]; *MAD*, mean amplitude deviation, i.e.,

Abbreviations in this page:
BFEN: band-pass filtered Euclidean norm
BMR: basal metabolic rate
ENMO: Euclidean norm minus 1G
HFEN: high-pass filtered Euclidean norm
HFEN+: HFEN plus the Euclidean norm minus 1G
LFECounts: activity counts with the low-frequency extension filter in the vector magnitude
MAD: mean amplitude deviation
PAEE: physical activity-related energy expenditure
TEE: total energy expenditure
VMCounts: activity counts in the vector magnitude

Abbreviations in this page:
ESM: electronic supplementary material
PAEE: physical activity-related energy expenditure
rMSE: root mean square error
SD: standard deviation
TEE: total energy expenditure

mean distance of the data points for the mean of the epoch using the Euclidean norm of the auto-calibrated raw accelerations [147].

The code of the GGIR package was modified in order to apply the same methods to all of the derived metrics (including ActiGraph's activity counts and alternate summary metrics), e.g., same treatment of non-wear time, detection of waking and sleeping hours, etc. For more information about the processing methods, see **ESM 4**. The code is available at https://github.com/jhmigueles/MINISTOP_data_processing.

Statistical analyses

Descriptive characteristics of the different metrics were calculated by descriptive or frequency analyses. Linear regression analyses were used to determine the amount of variation in TEE and PAEE explained by the different acceleration mean metrics (i.e., daily average 5-second epochs values) and in combination with body composition measures to estimate TEE and PAEE (i.e., identical models for each of the metrics). Body composition outcomes used in previous literature were included in the present study [234,241,242].

The out-of-sample performance of the developed equations was studied using a 13-fold cross-validation approach, which has demonstrated to out-perform the leave-one-out cross-validation when it is applied to linear regression models [243]. Briefly, we randomly split the sample into 13 groups of 3 participants each. Then, to calculate the estimated TEE and PAEE for each participant, we used the linear regression equation developed with those participants not belonging to this group (i.e., the 36 remaining participants after excluding one group of 3 participants). Mean differences between the measured TEE and PAEE with doubly labelled water and the estimated TEE and PAEE from the cross-validated equations were compared using independent t-tests. Bias and root mean square error (rMSE) of the estimations from the cross-validated equations were also calculated. Finally, we performed Bland-Altman plots to examine the agreement between the measured TEE and PAEE with doubly labelled water and the estimated TEE and PAEE from the cross-validated equations [244]. According to this method, predicted PAEE minus measured PAEE (y-axis) was plotted against the average of the predicted and measured PAEE (x-axis) for all 39 children. The mean difference as well as the limits of agreement (±2SD) were then calculated. To test for a trend within methods (i.e., whether the magnitude of the error changes across the x-axis) a linear regression model was fitted between the x and y axis. The same procedure was repeated for TEE.

All analyses were performed in R software v. 3.4.1. Statistical significance was defined as $p<0.05$.

ESM 4 (scan or click here)

Results

Age, anthropometric characteristics, energy expenditure and acceleration metrics are presented in **Table 10** as means and SDs. For the 39 participants, 3 recorded valid data for 10 days, 3 for 9 days, 12 for 8 days, 15 for 7 days, 5 for 6 days and 1 for 5 days. Therefore, 97% of the sample had ≥6 days.

Table 11 and **Table 12** show the performance and the regression equations of the acceleration metrics adjusted for body composition measures to estimate TEE and PAEE, respectively. For TEE, alternate summary metrics as unique predictors explained 13-16% and ActiGraph's activity counts explained 7-8% of the variance. For PAEE, alternate summary metrics explained 35-39% and ActiGraph's activity counts 25% of the variance. Among alternate summary metrics, ENMO and MAD provided slightly higher proportions of the explained variance in TEE and PAEE in most of the models. Adjustments for body composition measures considerably improved the estimations of TEE compared to the acceleration metrics alone, i.e., explained variance in TEE increased up to 67-76% for body weight and height and 81-84% for fat mass and fat-free mass. Likewise, explained variance in PAEE increased up to 31-51% adjusting for body weight and height and 64-67% for fat mass and fat-free mass. Models using fat mass and fat-free mass as covariates explained similar proportions of the variance of TEE and PAEE for all of the acceleration metrics (81-84% in TEE and 64-67% in PAEE).

Abbreviations in this page:
BFEN: band-pass filtered Euclidean norm
ENMO: Euclidean norm minus 1*G*
HFEN: high-pass filtered Euclidean norm
$HFEN_+$: HFEN plus the Euclidean norm minus 1*G*
LFECounts: activity counts with the low-frequency extension filter in the vector magnitude
MAD: mean amplitude deviation
PAEE: physical activity-related energy expenditure
SD: standard deviation
TEE: total energy expenditure
VMCounts: activity counts in the vector magnitude

Table 11
Linear regression models to assess the TEE estimation validity of each metric adjusted for body composition estimates

Covariates Metric	Equations to estimate TEE (kJ/day)	Adj. r^2	Standard error of the estimation	*P* model
Unadjusted				
VMCounts	kJ/day = 4773.93 + 5.75·VMCounts*	0.08	663.8	0.050
LFECounts	kJ/day = 4749.83 + 5.67·LFECounts	0.07	664.2	0.051
ENMO	kJ/day = 4844.85 + 25.06·ENMO**	0.16	631.2	0.006
BFEN	kJ/day = 4321.63 + 10.59·BFEN*	0.13	645.4	0.015
HFEN	kJ/day = 4310.10 + 10.58·HFEN*	0.13	645.0	0.015
$HFEN_+$	kJ/day = 4372.53 + 10.98·($HFEN_+$)**	0.14	638.9	0.010
MAD	kJ/day = 4663.06 + 22.41·MAD**	0.15	635.5	0.011
Height (H) + Weight (W)				
VMCounts	kJ/day = -6689.20 + 5.27·VMCounts** + 30.02·W + 95.9·H***	0.67	397.2	<0.001
LFECounts	kJ/day = -6729.48 + 5.2·LFECounts** + 29.87·W + 96.07·H***	0.67	397.4	<0.001
ENMO	kJ/day = -5690.9 + 24.27·ENMO*** + 42.54·W* + 84.93·H***	0.76	339.5	<0.001
BFEN	kJ/day = -6564.66 + 9.92·BFEN*** + 36.77·W + 89.66·H***	0.74	368.2	<0.001
HFEN	kJ/day = -6573.15 + 9.94·HFEN*** + 36.9·W + 89.59·H***	0.72	367.0	<0.001
$HFEN_+$	kJ/day = -6491.85 + 10.56·($HFEN_+$)*** + 38.3·W + 88.81·H***	0.74	352.4	<0.001
MAD	kJ/day = -5719.85 + 21.07·MAD*** + 42.60·W + 83.98·H***	0.74	354.6	<0.001
Fat mass (FM) + fat-free mass (FFM)				
VMCounts	kJ/day = 271.88 + 4.48·VMCounts** - 47.87·FM* + 334.76·FFM***	0.81	297.6	<0.001
LFECounts	kJ/day = 248.18 + 4.42·LFECounts** - 48.05·FM* + 334.99·FFM***	0.81	297.7	<0.001
ENMO	kJ/day = 678.26 + 18.04·ENMO*** – 25.8·FM + 307.97·FFM***	0.84	277.1	<0.001
BFEN	kJ/day = 118.18 + 7.87·BFEN*** – 37.01·FM + 321.62·FFM***	0.83	282.7	<0.001
HFEN	kJ/day = 110.06 + 7.88·HFEN*** – 36.73·FM + 321.32·FFM***	0.83	281.9	<0.001
$HFEN_+$	kJ/day = 216.74 + 8.12·($HFEN_+$)*** – 32.63·FM + 316.41·FFM***	0.84	277.3	<0.001
MAD	kJ/day = 524.22 + 15.93·MAD*** – 28.79·FM + 311.33·FFM***	0.83	282.3	<0.001

*$P < 0.05$
**$P < 0.01$
***$P < 0.001$

Table 12
Linear regression models to assess the PAEE estimation validity of each metric adjusted for body composition estimates

Covariates Metric	Equations to estimate TEE (kJ/day)	Adj. r^2	Standard error of the estimation	*P* model
Unadjusted				
VMCounts	kJ/day = 4773.93 + 5.75·VMCounts*	0.08	663.8	0.050
LFECounts	kJ/day = 4749.83 + 5.67·LFECounts	0.07	664.2	0.051
ENMO	kJ/day = 4844.85 + 25.06·ENMO**	0.16	631.2	0.006
BFEN	kJ/day = 4321.63 + 10.59·BFEN*	0.13	645.4	0.015
HFEN	kJ/day = 4310.10 + 10.58·HFEN*	0.13	645.0	0.015
$HFEN_+$	kJ/day = 4372.53 + 10.98·($HFEN_+$)**	0.14	638.9	0.010
MAD	kJ/day = 4663.06 + 22.41·MAD**	0.15	635.5	0.011
Height (H) + Weight (W)				
VMCounts	kJ/day = -6689.20 + 5.27·VMCounts** + 30.02·W + 95.9·H***	0.67	397.2	<0.001
LFECounts	kJ/day = -6729.48 + 5.2·LFECounts** + 29.87·W + 96.07·H***	0.67	397.4	<0.001
ENMO	kJ/day = -5690.9 + 24.27·ENMO*** + 42.54·W* + 84.93·H***	0.76	339.5	<0.001
BFEN	kJ/day = -6564.66 + 9.92·BFEN*** + 36.77·W + 89.66·H***	0.74	368.2	<0.001
HFEN	kJ/day = -6573.15 + 9.94·HFEN*** + 36.9·W + 89.59·H***	0.72	367.0	<0.001
$HFEN_+$	kJ/day = -6491.85 + 10.56·($HFEN_+$)*** + 38.3·W + 88.81·H***	0.74	352.4	<0.001
MAD	kJ/day = -5719.85 + 21.07·MAD*** + 42.60·W + 83.98·H***	0.74	354.6	<0.001
Fat mass (FM) + fat-free mass (FFM)				
VMCounts	kJ/day = 271.88 + 4.48·VMCounts** - 47.87·FM* + 334.76·FFM***	0.81	297.6	<0.001
LFECounts	kJ/day = 248.18 + 4.42·LFECounts** - 48.05·FM* + 334.99·FFM***	0.81	297.7	<0.001
ENMO	kJ/day = 678.26 + 18.04·ENMO*** – 25.8·FM + 307.97·FFM***	0.84	277.1	<0.001
BFEN	kJ/day = 118.18 + 7.87·BFEN*** – 37.01·FM + 321.62·FFM***	0.83	282.7	<0.001
HFEN	kJ/day = 110.06 + 7.88·HFEN*** – 36.73·FM + 321.32·FFM***	0.83	281.9	<0.001
$HFEN_+$	kJ/day = 216.74 + 8.12·($HFEN_+$)*** – 32.63·FM + 316.41·FFM***	0.84	277.3	<0.001
MAD	kJ/day = 524.22 + 15.93·MAD*** – 28.79·FM + 311.33·FFM***	0.83	282.3	<0.001

*$P < 0.05$
**$P < 0.01$
***$P < 0.001$

Abbreviations in this page:
BFEN: band-pass filtered Euclidean norm
ENMO: Euclidean norm minus 1*G*
HFEN: high-pass filtered Euclidean norm
$HFEN_+$: HFEN plus the Euclidean norm minus 1*G*
LFECounts: activity counts with the low-frequency extension filter in the vector magnitude
MAD: mean amplitude deviation
PAEE: physical activity-related energy expenditure
SD: standard deviation
TEE: total energy expenditure
VMCounts: activity counts in the vector magnitude

In exploratory analyses, we additionally tested a model using fat mass, fat-free mass and height as covariates that resulted in poorer estimations than that of the fat mass and fat-free mass models presented. Finally, since adjustments for age and sex did not improve the estimation capacity of the models, they were not used in any of our regression models (data not shown).

On average, estimations of TEE and PAEE from the 13-fold cross-validated equations were not significantly different to the TEE and PAEE assessed using doubly labelled water (**Table 13**). **Figure 12** shows the Bland-Altman plots using VMCounts and ENMO models as an example. Inverse associations were observed in the Bland-Altman plots for all of the metrics (r = -0.36 to -0.53, p = 0.001-0.02, see **Figure 12**).

Discussion

The main findings of this study were: 1) alternate summary metrics (i.e., ENMO, BFEN, HFEN, $HFEN_+$ and MAD) explained a larger proportion of the variance in TEE (+2-12%) and PAEE (+1-20%) than ActiGraph's activity counts (i.e., VMCounts and LFECounts) from the wrist-worn wGT3X-BT accelerometer in pre-schoolers; 2) Equations combining acceleration metrics and body composition measures explained up to 84% and 67% of the variation in TEE and PAEE, respectively; and 3) Our cross-validation developed equations produced a non-biased assessment of TEE and PAEE compared to the

doubly labelled water method, although we suggest to cross-validate the developed equations in other samples to confirm these findings. Among the alternate summary metrics, ENMO and MAD provided slightly higher proportions of the variance explained in TEE and PAEE in most of the models. This together with the fact that ENMO and MAD are easier to calculate and interpret since they do not use frequency filters and have a unit of measurement (i.e., *G*'s) suggest a need of further exploration of these variables in studies on PA and PAEE in pre-schoolers.

Table 13
Bias, 95% CIs and rMSE from comparisons between the estimated TEE and PAEE from the cross-validated equations and the measured TEE and PAEE with doubly labelled water

Covariates / Metric	TEE estimation (kJ/day)			PAEE estimation (kJ/day)		
	Bias	*P* values[a]	rMSE	Bias	*P* values[a]	rMSE
Unadjusted						
VMCounts	-11.22	0.920	682.83	-4.10	0.948	382.66
LFECounts	-10.96	0.922	683.01	-4.07	0.948	383.00
ENMO	-16.00	0.881	657.48	0.22	0.997	340.52
BFEN	-15.06	0.891	670.23	-3.48	0.952	353.79
HFEN	-14.95	0.891	669.79	-3.42	0.953	353.04
$HFEN_+$	-14.50	0.894	664.76	-2.45	0.965	346.13
MAD	-15.39	0.887	661.37	-2.61	0.963	346.16
Height + Weight						
VMCounts	28.35	0.700	451.37	25.72	0.663	361.95
LFECounts	28.62	0.698	452.32	25.94	0.661	362.59
ENMO	26.06	0.661	364.13	21.54	0.666	306.36
BFEN	23.38	0.725	406.86	21.03	0.693	326.47
HFEN	23.61	0.722	406.22	21.21	0.690	326.08
$HFEN_+$	25.29	0.689	387.31	22.12	0.667	315.71
MAD	25.46	0.689	390.39	22.18	0.675	324.91
Fat mass + fat-free mass						
VMCounts	-17.74	0.731	315.89	-4.08	0.925	266.25
LFECounts	-17.40	0.736	315.80	-3.77	0.931	266.28
ENMO	-9.07	0.849	292.24	-0.48	0.991	261.42
BFEN	-15.65	0.748	299.71	-3.04	0.942	255.62
HFEN	-15.15	0.756	298.40	-2.62	0.950	255.12
$HFEN_+$	-9.96	0.832	288.02	0.93	0.982	254.60
MAD	-11.36	0.812	293.04	-0.78	0.985	263.66

[a] p-values from paired t-tests

Similar to previous studies [95,222,234,235], we found an improvement in the estimation of TEE and PAEE when body size and composition measures are used as covariates (i.e., body weight, height, fat mass and fat-free mass). As expected, since the muscle mass in fat-free mass is a major determinant of energy expenditure, the highest proportion of the variance in TEE and PAEE was explained when adding body composition measures. For some of our models, fat mass was independently associated with PAEE and TEE. This observation may seem intriguing, however, it agrees with previous findings in children aged 1.5 years [234], and supports further investigations of associations between body composition and energy metabolism in pre-schoolers. It is also important to note that the use of ENMO together with body weight and height explained as much as half of the variance in PAEE. Although this finding needs confirmation in populations with wider ranges of age and BMI, it suggests that there may be a potential in the future to use acceleration metrics like

Abbreviations in this page:
BFEN: band-pass filtered Euclidean norm
BMI: body mass index
CI: confidence interval
ENMO: Euclidean norm minus 1*G*
HFEN: high-pass filtered Euclidean norm
$HFEN_+$: HFEN plus the Euclidean norm minus 1*G*
LFECounts: activity counts with the low-frequency extension filter in the vector magnitude
MAD: mean amplitude deviation
PAEE: physical activity-related energy expenditure
rMSE: root mean square error
SD: standard deviation
TEE: total energy expenditure
VMCounts: activity counts in the vector magnitude

Abbreviations in this page:
ENMO: Euclidean norm minus 1*G*
LFECounts: activity counts with the low-frequency extension filter in the vector magnitude
PAEE: physical activity-related energy expenditure
rMSE: root mean square error
TEE: total energy expenditure
VMCounts: activity counts in the vector magnitude

ENMO in combination with simple anthropometrics for larger studies intended to estimate free-living PAEE as an alternative method to doubly labelled water.

Previous studies in pre-schoolers estimating energy expenditure from accelerometers have used different devices and wear placements, i.e., hip and chest [128,234,235]. Butte et al. used cross-sectional time series and multivariate adaptive regression splines to estimate TEE from the Actiheart (Camntech Limited, http://www.camntech.com) and the ActiGraph GT3X+ and body composition measures [128]. They found a rMSE of 105 and 139 kcal/day using hip VMCounts, heart rate, steps, sex, age, height and weight for the estimation of TEE, for the cross-sectional time series and the multivariate adaptive regression splines models, respectively. Accordingly, our models adjusted for weight and height produced a rMSE ranging from 356 to 439 kJ/day for ENMO and LFECounts respectively (~ 85 to 105 kcal/day) for the estimation of TEE. Sijtsma et al. [235] found 31% of the PAEE variance explained by CPM as a unique predictor from the Tracmor accelerometer (Philips DirectLife, Amsterdam, the Netherlands). Our results showed slightly higher proportions of the PAEE explained variance from alternate summary metrics (i.e., 36-39%) and lower from the ActiGraph's activity counts (i.e., 28%). It is important to note that activity count metrics are brand-dependent, so the counts from the Tracmor are not the same metric as the VMCounts and LFECounts used in this study. In regards to TEE, Sijtsma et al. [235] found that activity counts from the Tracmor adjusted for body weight and height (among other confounders) could predict 29% of the variation. The models in the present study adjusted for body weight and height are able to predict 67-74% of the variation in TEE for all of the acceleration metrics. Henriksson et al. [234] used the chest-worn Actiheart and they found that 76% of the variance in TEE could be explained by activity counts, fat mass and fat-free mass in children aged 1.5-3 years old. Our wrist-based models explained between 81-83% of the variance in TEE when adjusted for the same body composition measures. These results show that previous findings regarding the hip-worn accelerometers out-performing the wrist-worn accelerometers in laboratory conditions [245] might not follow the same trend in a free-living setting. Although we cannot directly compare the performance of hip- versus wrist-worn accelerometers since we only used the latter placement. Wrist accelerations reflect arm movements, so they could improve the detection of upper-body movements compared with hip accelerations, explaining the observed difference between the laboratory and free-living settings. Further, the comparisons with previous literature should be interpreted with caution since differences in age and body composition variability.

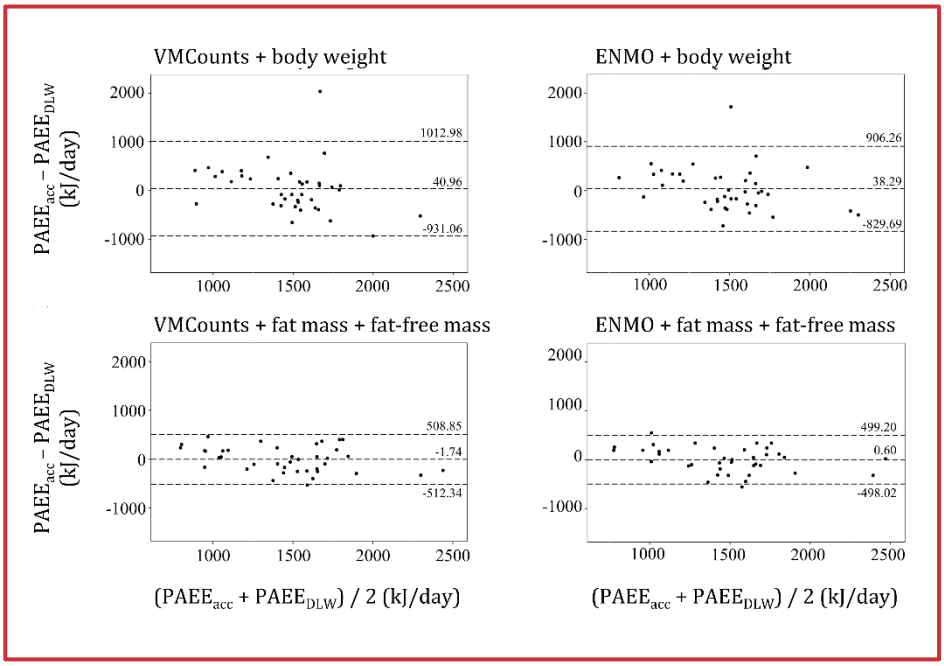


Figure 12
Example of Bland-Altman plots for the agreement between the estimated PAEE and the measured PAEE with doubly labelled water using the cross-validated equations from VMCounts and ENMO metrics (waking hours models).
Regression equations: A) y = -0.47x + 1456; r = -0.63; p <0.001; B) y = -0.49x + 1460; r = -0.53; p < 0.001; C) y = -0.37x + 1482; r = -0.41; p = 0.009; D) y = - 0.41x + 1480; r = -0.36; p = 0.02.

To the best of our knowledge, only three studies have assessed the performance of wrist-worn accelerometers to estimate free-living energy expenditure in any age group [67,95,222]. Firstly, Van Hees et al. [67,222] found that between 26 and 36% of the variation in PAEE could be explained by different alternate summary metrics (i.e., ENMO, BFEN, HFEN and HFEN$_+$) calculated from a wrist-worn accelerometer and adjusted for body weight in non-pregnant women. This is in concordance with our results in pre-schoolers, i.e., we found that between 33-37% of the variance in PAEE was explained from the same alternate summary metrics and body weight (data not shown). Secondly, we extend our previous findings in five-year-old children [95] to also include alternate summary metrics in addition to LFEcounts. In this context it is relevant to note that our previous results regarding LFECounts in these same participants [95] have been slightly improved from 24 to 28% of the variance explained in PAEE due to new decisions regarding the data processing of the raw accelerations. These decisions included: (i) the clipping score (i.e., abnormally sustained high accelerations due to malfunctioning of the devices -see **ESM 4**) was calculated to detect malfunctioning devices, (ii) non-wear time and the clipping score have been imputed by means of the rest of the recording days during the same time window since this procedure has shown to be better for the prediction of PAEE in adult women [222], (iii) we analysed all available days with enough wearing data in the present study while the number of days was restricted to 7 in the previous study [95], and (iv) one participant has been excluded from the present analyses for not meeting the inclusion criteria (i.e., 16 hours of valid data per day for at least 4 measurement days). We decided to change these processing criteria on the basis of the findings from recent studies

Abbreviations in this page:
BFEN: band-pass filtered Euclidean norm
ENMO: Euclidean norm minus 1G
ESM: electronic supplementary material
HFEN: high-pass filtered Euclidean norm
HFEN$_+$: HFEN plus the Euclidean norm minus 1G
LFECounts: activity counts with the low-frequency extension filter in the vector magnitude
MAD: mean amplitude deviation
PAEE: physical activity-related energy expenditure

ESM 4 (scan or click here)

Abbreviations in this page:
BFEN: band-pass filtered Euclidean norm
ENMO: Euclidean norm minus 1G
LFECounts: activity counts with the low-frequency extension filter in the vector magnitude
MAD: mean amplitude deviation
PAEE: physical activity-related energy expenditure
rMSE: root mean square error
TEE: total energy expenditure
VMCounts: activity counts in the vector magnitude

ESM 5 (scan or click here)

[67,222] and to apply the same procedures to all of the acceleration metrics included.

The 13-fold cross-validation performed in this study to test the out-of-sample performance of the developed equations showed no difference between TEE and PAEE assessed using doubly labelled water, as well as a relatively low rMSE compared to a previous study [235]. However, rMSE is not small enough to ensure a good prediction at an individual level and researchers should be cautious about that. Bland-Altman plots indicated a tendency that the accelerometer overestimated low energy expenditure (PAEE and TEE) values while corresponding high values were underestimated. Cross-validation of these equations in different samples should be performed in order to confirm the out-of-sample performance of these equations. Furthermore, other models for instance utilizing machine learning, are worthy to be tested in future studies in order to evaluate if they may improve the estimation capacity of these equations.

In regards to the epoch-by-epoch analyses, estimation of PAEE with VMCounts is lower than the ENMO estimations at high intensities (see **Figure S1**, **ESM 5**), which can explain the higher performance of the alternate summary metrics over ActiGraph's activity counts in this study. The unique difference between these models is the processing technique used to derive VMCounts and ENMO, i.e., raw accelerations are processed in different ways to remove gravity acceleration and noise from the signal without affecting the accelerations produced by body movement (i.e., PA). ActiGraph applies a band-pass frequency filter from 0.05 to 2.5 Hz, whilst the alternate summary metrics either do not use a frequency filter (i.e., ENMO and MAD) [67,147] or place the high bound at 15 Hz (i.e., BFEN) [67]. Furthermore, although it has been found that accelerometers worn on the wrist during walking and running on a treadmill do not plateau, the authors hypothesized that the lack of a predominant plane of movement in this location during ambulatory behaviours could explain this [246]. However, it could be possible that the higher complexity in the free-living movement of the human wrist produces movements up to 23 Hz in a same plane [96]. In this scenario, the band-pass filtering used to obtain activity counts would remove high wrist accelerations. Therefore, raw accelerations processing should be specific for the body placement where the device has been attached to. It is important to highlight that data collected with sampling frequencies different than 30, 60 or 90 Hz can cause issues when converting the raw data to activity counts by ActiGraph. Therefore, our models could be affected by this, so caution is needed when interpreting the information from our VMCounts and LFECounts models. It is also important to note that this epoch-by-epoch estimation has been performed with the unique objective of finding an

explanation for the differences found across acceleration metrics, therefore, these equations should never be used to estimate energy expenditure at an epoch level for describing PA patterns.

The practical implications of this study should be highlighted. We provide information showing that, if possible, alternate summary metrics should be used since they explain a larger amount of the variation in PAEE and TEE. Furthermore, we provide different equations for the estimation of TEE and PAEE in five-year-old children (**Table 11** and **Table 12**). All this information may be useful to researchers when investigating associations between PA, energy expenditure and energy balance, body composition and various health outcomes in preschool children. The use of different acceleration metrics and body composition measures ease the implementation of these equations in different scenarios, i.e., depending on the availability of collected information. Additionally, it is important to note that the GGIR package used to process the raw accelerations in this study is an open source software to automatically process the raw signal from the most used research-grade devices, i.e., ActiGraph, Geneactiv (GEN; GeneActive, ActiveInsights, Kimbolton, Cambridgeshire, United Kingdom) and Axivity (Axivity Ltd, UK). Caution is needed when comparing these metrics across devices since we cannot confirm that they are fully comparable [225]. Thus, the acceleration metrics used in this study can be easily calculated without high technical expertise.

The generalizability of our findings also deserves some comments. Children in this nested validation study were not different in terms of weight, height, fat mass and fat-free mass (mean and SD) from the complete cohort studied in the MINISTOP study. Likewise, our participants can be considered fairly representative of Swedish populations in comparison with standard representative weight and height values [247], as well as in terms of fat mass and fat-free mass when compared with British reference data in children [248]. Thus, our children covered a wide range in weight, height and body composition, which agrees with Swedish children in general. However, our sample only included five-year-olds and primarily normal-weight children (only two were classified as obese). This limits generalizability beyond the age of five years and to obese children. Furthermore, our findings need confirmation including a formal cross-validation of our equations in another population which should include a wider age range and more obese children.

There are some limitations that need to be acknowledged. First of all, BMR was predicted to calculate PAEE and it could be argued that the use of body weight to predict BMR could increase the correlation in the regression models for the estimation of PAEE. We find this unlikely as body weight explained only a small fraction of PAEE ($r^2 < 0.001$, $p = 0.904$). Additionally, the protocol for the

Abbreviations in this page:
PA: physical activity
PAEE: physical activity-related energy expenditure
SD: standard deviation
TEE: total energy expenditure

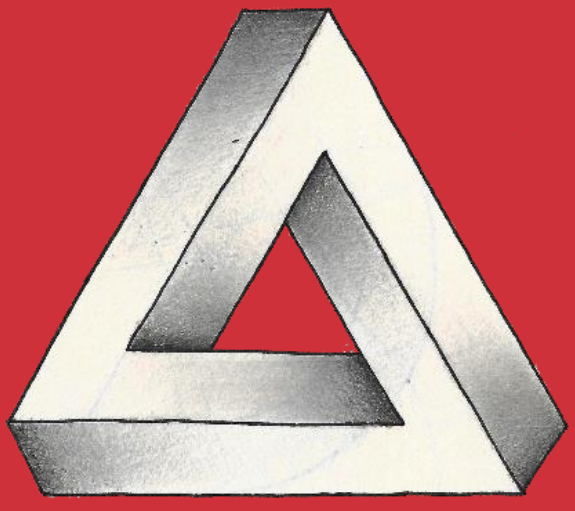
Abbreviations in this page:
BFEN: band-pass filtered Euclidean norm
ENMO: Euclidean norm minus 1*G*
HFEN: high-pass filtered Euclidean norm
$HFEN_+$: HFEN plus the Euclidean norm minus 1*G*
MAD: mean amplitude deviation
PAEE: physical activity-related energy expenditure
TEE: total energy expenditure

accelerometer data collection and the doubly labelled water protocol was not time consistent. Accelerometers were worn during 5-10 days whilst the doubly labelled water protocol required 14 days to ensure the maximum accuracy [249]. We replicated the analysis using only the doubly labelled water of the first 7 days to test whether the time difference could affect our findings. The results were similar for both TEE and PAEE (results not shown), which is probably due to the fact that day-to-day variation in TEE is small [250]. Although average TEE and PAEE were not different from reference TEE and PAEE, caution is advised when interpreting individual estimations from these equations. Finally, a formal cross-validation in a different sample is needed to confirm the current findings. Otherwise, our study has several strengths. First and foremost is the use of the gold-standard measure of TEE through doubly labelled water. Furthermore, we used and compared several methods to process the accelerometer raw signals in order to study how different acceleration metrics are related to TEE and PAEE. Finally, the high compliance for the accelerometer protocol allowed for a relevant representation of the PA patterns (i.e., 97% of the sample had ≥6 valid days).

Conclusion

In conclusion, a higher performance to predict PAEE and TEE was observed for the alternate summary metrics (i.e., ENMO, BFEN, HFEN, $HFEN_+$ and MAD) versus ActiGraph's counts, especially for the model including ENMO, body weight and height, which explained a relatively high proportion of the variance in PAEE (i.e., 50%). Further exploration of these variables in studies on PA and energy expenditure are warranted. Altogether, acceleration metrics from the wrist in combination with body composition measures explained up to 84% of the variance in TEE and 67% of the variation in PAEE after adjustments for body weight and height. Although confirmation of these findings in other populations is still needed, the results suggest that wrist-worn accelerometers have the potential to assess free-living TEE and PAEE in combination with body composition measures in pre-schoolers.

Study IV



Comparability of accelerometer signal aggregation metrics across placements and dominant wrist cut points for the assessment of physical activity in adults

Migueles JH, Cadenas-Sanchez C, Rowlands AV, Henriksson P, Shiroma EJ, Acosta FM, Rodriguez-Ayllon M, Esteban-Cornejo I, Plaza-Florido A, Gil-Cosano JJ, Ekelund U, van Hees VT, Ortega FB.
Sci Rep. 2019 Dec;9(1):18235
DOI: 10.1038/s41598-019-54267-y

Contents

Abstract

Background | Large epidemiological studies that use accelerometers for physical behaviour and sleep assessment differ in the location of the accelerometer attachment and the signal aggregation metric chosen.

Aims | This study aimed to assess the comparability of acceleration metrics between commonly-used body-attachment locations for 24 hours, waking and sleeping hours, and to test comparability of PA cut-points between dominant and non-dominant wrist.

Methods | Forty-five young adults (23 women, 18–41 years) were included and GT3X+ accelerometers were placed on their right hip, dominant, and non-dominant wrist for 7 days. We derived ENMO, Low-pass filtered ENMO (LFENMO), MAD and VMCounts over 5-second epochs from the raw accelerations. Metric values were compared using a correlation analysis, and by plotting the differences by time of the day.

Results | Cut points for the dominant wrist were derived using Lin's concordance correlation coefficient optimization in a grid of possible thresholds, using the non-dominant wrist estimates as reference. They were cross-validated in a separate sample (N = 36, 10 women, 22–30 years). Shared variances between pairs of acceleration metrics varied across sites and metric pairs (range in r^2: 0.19–0.97, all $p < 0.01$), suggesting that some sites and metrics are associated, and others are not. We observed higher metric values in dominant vs. non-dominant wrist, thus, we developed cut points for dominant wrist based on ENMO to classify SB (<50 mg), LPA (50–110 mg), MPA (110–440 mg) and VPA (≥440 mg).

Conclusion | Our findings suggest differences between dominant and non-dominant wrist, and we proposed new cut points to attenuate these differences. ENMO and LFENMO were the most similar metrics, and they showed good comparability with MAD. However, counts were not comparable with ENMO, LFENMO and MAD.

Abbreviations in this page:
ENMO: Euclidean norm minus 1*G*
LFENMO: ENMO of the low-pass filtered raw accelerations
LPA: light physical activity
MAD: mean amplitude deviation
MPA: moderate physical activity
PA: physical activity
SB: sedentary behaviour
VMCounts: activity counts in the vector magnitude
VPA: vigorous physical activity

Key Points

Question
How comparable are different acceleration metrics for the assessment of PA, SB, and sleep across common body attachment sites?

Findings
ENMO and LFENMO were the most similar metrics, and they showed good comparability with MAD. However, counts were not comparable with ENMO, LFENMO and MAD

Meaning
Our findings suggest differences in the PA measured from dominant and non-dominant wrist, and we proposed new cut points for the dominant wrist to attenuate these differences. Also, similar patterns can be identified by different acceleration metrics, which should be considered for studies comparability, reproducibility, and harmonization.

Introduction

Physical behaviours occurring during the 24 hours of the day consist of PA, SB and sleep. These three behaviours are of major public health interest due to their well-documented influence on health [8,10,251,252]. Objective methods to assess free-living PA range from cost-efficient tools such as pedometers to relatively expensive multi-sensor devices, e.g., Actiheart [253]. Accelerometers provide a balance of cost and feasibility and have been increasingly used in large epidemiological cohorts [254,255], for example in the Women's Health Study (WHS) and the NHANES, the UK Biobank, and the ISCOLE which collected data worldwide. Furthermore, accelerometers have been validated for the estimation of PA [61,69,70], SB [62,116,256] and sleep [38,40,257]. However, accelerometer utilization requires data collection and processing decisions which could affect the final estimations [233].

Data collection decisions start with the selection of the most appropriate anatomical location to attach the accelerometer [233]. Hip and wrist are the most frequently selected locations [233], and both have been demonstrated to be valid locations for classifying PA intensities and SB [61,69,70,233], as well as to assess sleep [38,40,257]. However, it is unclear how much the accelerometer outcome measures vary between body sites. Previous studies have found high correlations (i.e., ICCs > 0.9) between acceleration values from both wrists [258] with slightly lower values in the non-dominant wrist, although non-significantly different from the dominant wrist [259,260]. Likewise, moderate-to-high correlations between acceleration values from either wrist and the hip have been reported (i.e., r coefficients between 0.7 and 0.9) [260,261]. These studies had a focus on PA and SB (i.e., waking hours) and did not report associations during sleeping hours. Furthermore, cut points to estimate PA have been proposed for the non-dominant wrist [61,62] and hip [61,62,107] in adults, but not for the dominant wrist. Therefore, studies where accelerometers were placed on the dominant wrist do not have specific cut points proposed for their data, for instance the UK Biobank [204].

Data processing aims to remove the gravitational component and noise from the raw signal, in order to obtain an acceleration signal aggregation metric (from herein acceleration metric) intended to reflect body movement [67]. For example, acceleration metrics include ENMO [67], ENMO of the low-pass filtered raw accelerations (LFENMO) [204], MAD [147,262] as well as manufacturer-specific metrics such as VMCounts, among others (see definitions of these acceleration metrics in **Table 14**). To our knowledge, these acceleration metrics have not previously been compared to each other in the same study. Comparing these metrics using data from hip, dominant

Abbreviations in this page:
ENMO: Euclidean norm minus 1G
ICC: intraclass correlation coefficient
ISCOLE: International Study of Childhood Obesity, Lifestyle and Environment
LFENMO: ENMO of the low-pass filtered raw accelerations
MAD: mean amplitude deviation
NHANES: National Health And Nutrition Examination Survey
PA: physical activity
SB: sedentary behaviour
VMCounts: activity counts in the vector magnitude
WHS: Women's Health Study

Abbreviations in this page:
ENMO: Euclidean norm minus 1*G*
LFENMO: ENMO of the low-pass filtered raw accelerations
MAD: mean amplitude deviation
PA: physical activity
SB: sedentary behaviour
VMCounts: activity counts in the vector magnitude

wrist and non-dominant wrist and focusing on different periods of the day (i.e., 24 hours, waking and sleeping hours) could benefit researchers interested in either PA, SB and/or sleep.

Moreover, the movement pattern identified throughout the day by each of these metrics (i.e., acceleration metric values throughout the day) and the data from different body attachment sites may be useful to describe PA [263]. Only one study has investigated differences in movement patterns identified from hip, dominant wrist and non-dominant wrist [259]. This study only analysed VMCounts during waking hours in a sample of older adults [259]. The drawbacks of using brand-dependent VMCounts have been described (e.g., precludes comparison across studies, complicates the interpretation of results, summarizes raw data which may minimize their potential) [264], as well as the importance of moving forward with open-source derived metrics from raw accelerometer data [265,266]. Therefore, the present study aimed to: 1) study the comparability between different acceleration metrics across right hip, dominant wrist and non-dominant wrist attachment sites during different periods of the day (i.e., 24 hours, waking and sleeping hours); and 2) use previously established cut points for accelerations measured at the non-dominant wrist [61,62] to develop and cross-validate cut points in a separate sample for accelerations measured at the dominant wrist in a sample of young adults.

Table 14
Brief description of the acceleration metrics included

Acceleration metric	Frequency filter	Definition
ENMO	None	Euclidean norm minus one *g* of the raw accelerations, with resulting negative values rounded to zero and then averaged over 5 s epochs.
LFENMO	Low-pass	Euclidean norm minus one *g* of the low-pass filtered raw accelerations with resulting negative values rounded to zero (Butterworth 4th order filter; ω = 20 Hz).
MAD	None	Euclidean norm of each raw acceleration data point minus the mean of its correspondent 5 s epoch.
VMCounts	Band-pass	Counts are obtained by using a band-pass frequency filter to the raw signal (by default: ω_0=0.025 Hz, ω_1=2.5 Hz). The rest of information is mostly unknown.

Methods

Study design and participants

The present cross-sectional study analysed free-living data from a convenience sample composed of students and research personnel from the University of Granada, Spain. The study was carried out in two waves of 45 (23 women, 18-41 years old) and 36 (10 women, 22-30 years old) young adults, respectively. Wave 1 (cut-point calibration sample) was used to compare different acceleration metrics across body attachment sites and to develop cut points for the dominant wrist that are consistent with the only set of cut points proposed to estimate PA intensity from the non-dominant wrist in adults to date [61,62]. Wave 2 (cut-point cross-validation sample) data were used to cross-validate the new set of cut points in a

different sample of participants with similar characteristics. All participants were informed of the purpose of the study and written informed consent was obtained. This study was conducted according to the Declaration of Helsinki and approved by the Ethics Committee on Human Research of the University of Granada and the study was approved by the institutional review board of the University of Pittsburgh and National Institute on Aging.

Abbreviations in this page:
BMI: body mass index
ENMO: Euclidean norm minus 1 *G*
LFENMO: ENMO of the low-pass filtered raw accelerations
MAD: mean amplitude deviation
SD: standard deviation

Table 15
Descriptive characteristics of participants

	Cut-point calibration sample			Cut-point cross-validation sample[a]		
	All (N = 42)	Men (N = 19)	Women (N = 23)	All (N = 36)	Men (N = 26)	Women (N = 10)
Age (years)	27.3 (5.3)	27.2 (5.9)	27.4 (4.9)	24.3 (1.9)	24.4 (2.3)	24.1 (1.1)
Weight (kg)	67.8 (12.1)	78.3 (8.5)	59.1 (6.3)	70.3 (14.2)	77.0 (13.1)	58.3 (5.1)
Height (cm)	171.0 (8.3)	178.0 (1.7)	165.1 (5.2)	172.3 (9.7)	177.0 (7.8)	163.8 (6.6)
BMI (kg/m^2)	23.0 (2.6)	24.6 (1.7)	21.7 (2.5)	23.5 (3.6)	24.5 (3.8)	21.8 (2.5)
Acceleration metrics	**24 hours**	**Waking hours**	**Sleeping hours**	**24 hours**	**Waking hours**	**Sleeping hours**
Wear time (h/day)						
Right hip	23.9 (0.3)	17.2 (0.7)	6.7 (0.7)	-	-	-
Dom. wrist	24.0 (0.2)	17.0 (0.7)	6.9 (0.6)	23.7 (1.3)	16.2 (1.7)	7.5 (1.8)
Non-dom. wrist	24.0 (0.2)	17.0 (0.7)	7.0 (0.6)	23.6 (1.4)	16.2 (1.4)	7.4 (1.5)
ENMO (m*g*)						
Right hip	16.0 (5.6)	21.4 (7.7)	2.4 (1.3)	-	-	-
Dom. wrist	33.9 (7.6)	46.5 (10.6)	3.0 (1.5)	31.7 (14.0)	43.1 (13.8)	3.5 (2.3)
Non-dom. wrist	31.3 (6.8)	43.1 (9.9)	3.2 (1.4)	29.9 (12.9)	40.7 (12.8)	4.4 (4.8)
LFENMO (m*g*)						
Right hip	12.1 (4.7)	16.0 (6.4)	2.1 (1.1)	-	-	-
Dom. wrist	26.4 (6.4)	36.1 (8.9)	2.5 (1.2)	-	-	-
Non-dom. wrist	24.9 (5.9)	34.3 (8.6)	2.6 (1.2)	-	-	-
MAD (m*g*)						
Right hip	24.4 (6.9)	33.4 (9.4)	1.5 (1.7)	-	-	-
Dom. wrist	48.4 (8.9)	67.0 (12.2)	2.8 (2.2)	-	-	-
Non-dom. wrist	44.2 (8.5)	61.4 (12.0)	2.9 (2.1)	-	-	-
VMCounts/5s						
Right hip	41.4 (9.9)	56.3 (13.6)	3.6 (3.0)	-	-	-
Dom. wrist	176.9 (31.8)	242.6 (43.2)	14.6 (8.9)	-	-	-
Non-dom. wrist	164.9 (29.3)	226.7 (41.6)	15.0 (7.8)	-	-	-

Data are presented as mean (SD)
[a]Cut-point cross-validation sample data was only used to cross-validated cut points for dominant wrist based on ENMO, so they did not wear hip-worn accelerometers and only ENMO was derived.

Procedures

Participants' body weight and height were measured to the nearest 0.1 kg and 0.1 cm using an electronic scale (SECA 861, Hamburg, Germany) and a precision stadiometer (SECA 225, Hamburg, Germany). We calculated BMI as mass (kg) / height^2 (m^2). Participants were instructed to wear accelerometers (ActiGraph GT3X+, Pensacola, FL, USA) for seven complete days (24 hours per day). Cut-point calibration sample participants wore three accelerometers placed on the right hip, dominant wrist and non-dominant wrist. Cut-point cross-validation sample participants wore accelerometers on both wrists. All participants were instructed to remove the accelerometers during bathing and showering, to always wear and remove all of the accelerometers at the same time, and to keep a diary of the times they went to bed and got off the bed every day.

Accelerometers

ActiGraph GT3X+ is a triaxial accelerometer with a dynamic range of +/- 6 G. Accelerometers were initialized to capture and store accelerations at 100 Hz. Raw accelerations were then downloaded

Abbreviations in this page:
CI: confidence interval
ENMO: Euclidean norm minus 1*G*
LFENMO: ENMO of the low-pass filtered raw accelerations
LPA: light physical activity
MAD: mean amplitude deviation
MPA: moderate physical activity
PA: physical activity
SB: sedentary behaviour
SD: standard deviation
VMCounts: activity counts in the vector magnitude
VPA: vigorous physical activity

(".gt3x" files) and converted to ".csv" format using ActiLife v.6.13.3 (ActiGraph, Pensacola, FL, USA).

Data processing

Raw ".gt3x" files were loaded in the ActiLife software to export raw data in .csv format and to obtain VMCounts over 5 s epochs using the software's default filter (ω_0=0.025 Hz, ω_1=2.5 Hz). Next, raw ".csv" files were processed using the GGIR software (version 1.6-0, https://cran.r-project.org/web/packages/GGIR/) [67,200]. The processing methods of GGIR involved: 1) Auto-calibration of the data according to the local gravity [66]; 2) calculation of ENMO, LFENMO and MAD and inclusion of the previously obtained VMCounts over 5 seconds epochs (**Table 14**) to participant datasets to follow the same non-wear time calculation and processing decisions than the rest of the acceleration metrics; 3) detection of the non-wear time based on the raw acceleration from the three axes using a validated algorithm [67], briefly, each 15-min block was classified as non-wear time if the SD of 2 out of the 3 axes was lower than 13 m*g* during the surrounding 60-min moving window or if the value range for 2 out of the 3 axes was lower than 50 m*g*; 4) detection of sustained abnormal high accelerations, i.e., higher than 5.5 g; 5) imputation of detected non-wear time and abnormal high accelerations by means of the acceleration for the rest of the recording period during the same time interval than the affected periods; and 6) separation of waking and sleeping hours using a validated algorithm on the non-dominant wrist data and guided by logged timestamps by participants [40]. Logged times were 9 min (95% confidence intervals [$CI_{95\%}$]: -6 to 25 min) earlier and 17 min ($CI_{95\%}$: 2 to 32 min) later than accelerometer detected times for sleep onset and wake-up times, respectively. Finally, waking and sleeping hours detected from the non-dominant wrist were applied to the hip and dominant wrist measurements of each participant. All participants with at least 4 days with ≥16 hours wearing accelerometers were included in the analyses.

Data analysis

SB and time spent in each PA intensity (i.e., LPA, MPA and VPA) were estimated from the ENMO metric from the dominant wrist and non-dominant wrist-worn accelerometer data. Hildebrand et al.'s [61,62] ENMO cut points developed for the non-dominant wrist were applied to the dominant wrist and non-dominant wrist data. Additionally, we calculated the same variables using the validated cut points incremented by 5 and 10 m*g* [61,62] for the dominant wrist data only. Daily means of the acceleration metrics for 24 hours, waking hours and sleeping hours, as well as estimations of time spent in SB and PA intensity levels from the wrist-worn accelerometers were included in the analyses.

Statistics

Descriptive statistics were calculated as means and SDs. We used linear regressions to study the associations between the different acceleration metrics (i.e., ENMO, LFENMO, MAD and VMCounts) calculated from the same and different body attachment sites (i.e., right hip, dominant wrist and non-dominant wrist) (i.e., aim 1).

In order to study whether acceleration metrics identify a different movement pattern over the day (i.e., aim 2), we plotted 30-min averages of acceleration metrics across body attachment sites. As each acceleration metric has a different unit of measurement, we used z-scores when different acceleration metrics appeared in the same plot. Furthermore, we performed a curve analysis using statistical parametric mapping (SPM) [267] to compare accelerations from dominant and non-dominant wrists throughout the day. Acceleration data over the day were depicted as curves. These acceleration curves produced throughout the day are highly variable between individuals due to several factors (e.g., lifestyle, working schedule). To minimize this high variability and allow for a comparison of the curves, we sorted accelerations produced every day per participant in an increasing order. Therefore, all of the curves start with the periods of the day when activity was low (e.g., sleep, SB activities…) and finish with the periods of the day with the highest intensity activities, independently of the moment of the day when they occurred. T-tests were used to determine significant differences between the curves for dominant and non-dominant wrists. SPM involved 4 steps to compute the t-test analysis: 1) computing the value of a test statistic at each point in the normalized time series; 2) estimating temporal smoothness on the basis of the average temporal gradient; 3) an equally smooth random process is performed to compute the value of the test statistics above which only $\alpha = 5\%$ of the data would be expected to reach; 4) computing the probability that specific suprathreshold regions could have resulted from an equivalently smooth random process.

Finally, the Lin's concordance correlation coefficient (LCCC), two sample t-tests and mean absolute percent error (MAPE) were used to study the agreement between SB time and time spent in each PA intensity derived from non-dominant wrist with validated cut points (reference) and all cut points used for dominant wrist (see *Data processing* section) (i.e., aim 3).

Cut-point selection was made following these criteria: 1) closest VPA estimation if any; 2) closest MPA estimation if any (the upper threshold is already defined in step 1); 3) closest LPA estimation if any (the upper threshold is already defined in step 2). When two or more thresholds provided similar results, we tried to respect the original distance between the previously-established thresholds for

Abbreviations in this page:
ENMO: Euclidean norm minus 1*G*
LFENMO: ENMO of the low-pass filtered raw accelerations
LCCC: Lin's concordance correlation coefficient
LPA: light physical activity
MAD: mean amplitude deviation
MAPE: mean absolute percent error
MPA: moderate physical activity
PA: physical activity
SB: sedentary behaviour
SD: standard deviation
SPM: statistical parametric mapping
VMCounts: activity counts in the vector magnitude
VPA: vigorous physical activity

Abbreviations in this page:
CI: confidence interval
ENMO: Euclidean norm minus 1*G*
LFENMO: ENMO of the low-pass filtered raw accelerations
LCCC: Lin's concordance correlation coefficient
MAD: mean amplitude deviation
MAPE: mean absolute percent error
PA: physical activity
SPM: statistical parametric mapping
VMCounts: activity counts in the vector magnitude

non-dominant wrist [61,62]. The closest thresholds were selected based on mean differences, LCCC and MAPE. Then, the selected cut points were cross-validated using data from the cut-point cross-validation sample. All analyses were performed in R v.3.4.1 (https://cran.r-project.org/). The significance level was $P < 0.05$.

Results

Descriptive characteristics of participants

Out of the 45 participants from the cut-point calibration sample, three were excluded from the analyses for either not having accumulated enough wear time (n = 2) or having mis-placed the dominant wrist and non-dominant accelerometers for ≥ 1 day (n = 1). All of the cut-point cross-validation sample participants were included in the cross-validation analyses (i.e., aim 3). Descriptive characteristics of the included participants, as well as acceleration metric values for each body attachment site (i.e., right hip, dominant wrist and non-dominant wrists) across periods of the day (i.e., 24 hours, waking hours and sleeping hours) are presented in **Table 15**. Since wear time was high and practically the same for the right hip, dominant wrist and non-dominant wrist accelerometers, we did not delete unmatched non-wear time across accelerometers.

Comparison of each acceleration metric across body attachment sites

Table 3 presents the shared variances for every acceleration metric across different body attachment sites, i.e., hip, dominant wrist and non-dominant wrist. Overall, shared variance between wrists (r^2 between 0.56 [$CI_{95\%}$: 0.33 – 0.74] and 0.94 [$CI_{95\%}$: 0.89 – 0.97]) was higher than shared variance between any of the wrists and the hip (r^2 between 0.21 [$CI_{95\%}$: 0.03 – 0.45] and 0.88 [$CI_{95\%}$: 0.78 – 0.93]) for all the metrics analysed.

Table 16
Shared variance (r^2) for each acceleration metric across different body attachment sites (i.e., hip, dominant and non-dominant wrists).

	ENMO	LFENMO	MAD	VMCounts
24 hours				
Right hip vs. Dominant wrist	**0.37****	**0.28***	0.68**	0.40**
Right hip vs. Non-dominant wrist	**0.34****	**0.25***	0.77**	0.43**
Dominant vs. Non-dominant wrist	0.79**	0.78**	0.86**	0.71**
Waking hours				
Right hip vs. Dominant wrist	**0.37****	**0.31****	0.66**	0.38**
Right hip vs. Non-dominant wrist	**0.35****	**0.28***	0.75**	0.42**
Dominant vs. Non-dominant wrist	0.79**	0.79**	0.86**	0.70**
Sleeping hours				
Right hip vs. Dominant wrist	0.37**	0.21*	0.88**	0.75**
Right hip vs. Non-dominant wrist	0.39**	0.27*	0.88**	0.69**
Dominant vs. Non-dominant wrist	0.67**	0.56**	0.94**	0.92**

* $P < 0.01$
** $P < 0.001$

Figure 13 shows that acceleration values (for all of the metrics) for the wrists are higher than for the hip, with the highest values reached in the dominant wrist. Although the PA pattern seemed to be concordant for dominant wrist and non-dominant wrists, SPM analysis showed significant differences ($p < 0.001$) between the 50th and

the 90th percentile of the accelerations produced (when accelerations start to increase, indicating periods of PA, see **Figure 14**). In regards to the PA pattern identified from hip, besides recording lower values, peaks of activity were not totally concordant with those identified by the wrists.

Abbreviations in this page:
ENMO: Euclidean norm minus 1*G*
LFENMO: ENMO of the low-pass filtered raw accelerations
MAD: mean amplitude deviation
PA: physical activity
SPM: statistical parametric mapping
VMCounts: activity counts in the vector magnitude

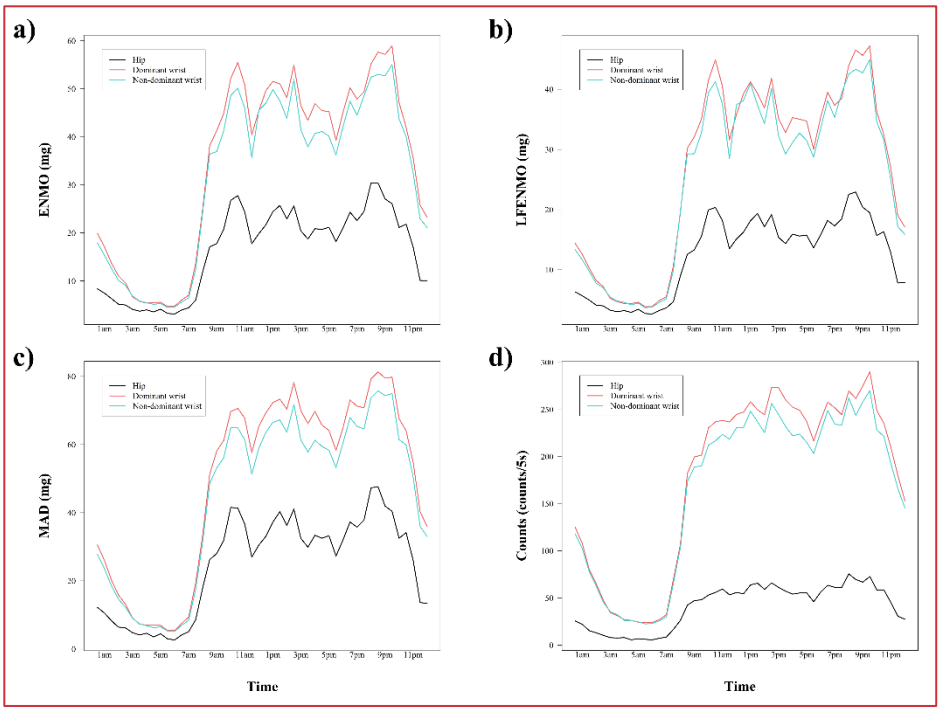


Figure 13
Means of ENMO (Panel A), LFENMO (Panel B), MAD (Panel C) and Counts (Panel D) over 30-min periods for the hip, dominant and non-dominant wrist.
Each data point is the average for this time interval for all participants from the cut-point calibration sample (n = 42).

Comparison of acceleration metrics derived from the same body attachment site

Table 4 shows the shared variances (r^2) between pair of acceleration metrics derived from the same body attachment site and period of the day (i.e., 24 hours, waking and sleeping hours). ENMO and LFENMO were the metrics which presented the highest shared variances among all of the metrics included (r^2 ranged from 0.95 [$CI_{95\%}$: 0.90 – 0.97] to 0.97 [$CI_{95\%}$: 0.93 – 0.98] for all locations and moments of the day). The lowest shared variances were found between LFENMO and VMCounts (r^2 = 0.19 [$CI_{95\%}$: 0.02 – 0.42]), and for LFENMO and MAD during sleeping hours for the hip (r^2 = 0.21 [$CI_{95\%}$: 0.03 – 0.45]). For the rest of the metrics, in general, they presented higher r^2 values during waking hours (r^2 between 0.38 [$CI_{95\%}$: 0.11 – 0.57] and 0.92 [$CI_{95\%}$: 0.85 – 0.96]) than during sleeping hours (r^2 between 0.32 [$CI_{95\%}$: 0.10 – 0.55] and 0.79 [$CI_{95\%}$: 0.65 – 0.88]). A graphical comparison of all of the acceleration metrics for each body attachment site can be found in **Figure 15**. While ENMO, LFENMO and MAD were describing almost the same movement pattern when derived from the same attachment site, VMCounts were more discordant in some periods of the day (e.g., between 9 and 11 am, VMCounts did not detect a peak of activity identified by the rest of the metrics in all of the placements, **Figure 15**).

Table 17
Shared variance (r^2) between different acceleration metrics derived from the same body attachment site (N = 42).

	ENMO vs. LFENMO	ENMO vs. MAD	ENMO vs. VMCounts	LFENMO vs. MAD	LFENMO vs. VMCounts	MAD vs. VMCounts
24 hours						
Right hip	0.97**	0.72**	0.46**	0.59**	0.34**	0.81**
Dominant wrist	0.96**	0.91**	0.55**	0.86**	0.51**	0.66**
Non-dominant wrist	0.97**	0.87**	0.49**	0.82**	0.47**	0.56**
Waking hours						
Right hip	0.97**	0.74**	0.48**	0.62**	0.38**	0.81**
Dominant wrist	0.97**	0.92**	0.54**	0.87**	0.51**	0.64**
Non-dominant wrist	0.97**	0.89**	0.52**	0.85**	0.51**	0.57**
Sleeping hours						
Right hip	0.95**	0.42**	0.32**	0.21**	0.19*	0.54**
Dominant wrist	0.97**	0.75**	0.44**	0.59**	0.37**	0.47**
Non-dominant wrist	0.97**	0.79**	0.42**	0.66**	0.37**	0.44**

* $P < 0.01$
** $P < 0.001$

Abbreviations in this page:
ENMO: Euclidean norm minus 1*G*
LFENMO: ENMO of the low-pass filtered raw accelerations
LCCC: Lin's concordance correlation coefficient
LPA: light physical activity
MAD: mean amplitude deviation
MAPE: mean absolute percent error
MPA: moderate physical activity
PA: physical activity
SB: sedentary behaviour
VMCounts: activity counts in the vector magnitude
VPA: vigorous physical activity

Cut-point replication

As the shared variance between ENMO derived from the dominant and non-dominant wrist was fairly high during waking hours (r^2 = 0.79 [$CI_{95\%}$: 0.65 – 0.88], Table 4), and the movement pattern over the day identified was very similar for both wrists (**Figure 13**), we replicated the cut points validated by Hildebrand et al. in the non-dominant wrist [61,62] using dominant wrist ENMO values (**Table 5**). After considering the criteria defined to select the new cut points, we found that the closest estimation between wrists was achieved with 50, 110 and 440 m*g* to classify SB and LPA, MPA and VPA, respectively. Criteria were: 1) closest VPA estimation if any, no threshold was selected since all of them were equally comparable; 2) closest MPA estimation if any, 110-430 m*g*, 110-435 m*g* and 110-440 m*g* provided similar results, so 110-440 m*g* was selected to respect the distance between previously-established cut points (i.e., 100-430 m*g*); 3) closest light PA estimation considering the upper threshold defined in step 2 (i.e., 110 mg). This selection of cut points was cross-validated using data from the wave 2; estimations were not significantly different between wrists for any of the intensities and showed high LCCCs (0.85-0.92) and low MAPEs (0.67-14.29%).

Discussion

Main findings

The main findings of this study were: (i) the dominant wrist showed systematically higher acceleration metric values than the non-dominant wrist, which were translated into different SB time and PA estimations when using the same cut points; ii) dominant and non-dominant wrist based estimations of PA became comparable by modifying the validated cut points, which was confirmed in a cross-validation sample; (iii) non-dominant and dominant wrist acceleration metrics shared a higher proportion of variance than between the hip and either wrist, while MAD was the metric with the highest shared variances across body attachment sites; (iv) overall, the metrics ENMO, LFENMO and MAD shared higher proportions of variance

than any of these metrics with VMCounts, especially when the metrics were derived from the wrist-worn accelerometers; (v) the movement pattern identified throughout the day was visually equivalent for any given acceleration metric (e.g., ENMO) from the dominant wrist and non-dominant wrist, and was also similar when comparing either wrist with the hip and across acceleration metrics derived from the same body site (with the exception of VMCounts). Altogether, these findings demonstrate the extent to which different factors related to data collection (e.g., anatomical wear location) and processing procedures (e.g., different acceleration metrics) could modify the final PA, SB time, and sleep estimations.

Abbreviations in this page:
ENMO: Euclidean norm minus 1G
LFENMO: ENMO of the low-pass filtered raw accelerations
MAD: mean amplitude deviation
PA: physical activity
SB: sedentary behaviour
VMCounts: activity counts in the vector magnitude

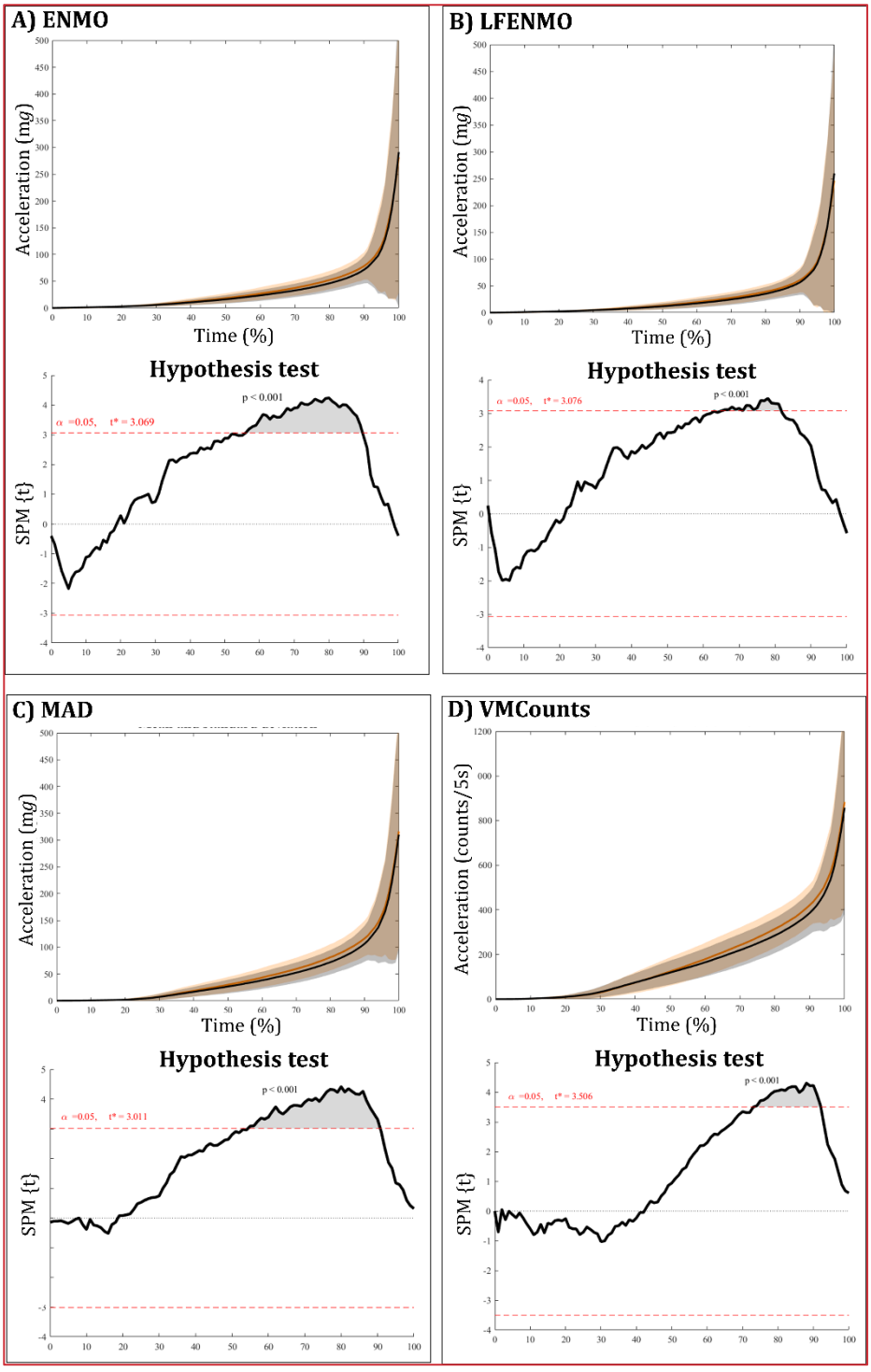


Figure 14
Comparison of means of ENMO (Panel A), LFENMO (Panel B), MAD (Panel C) and Counts (Panel D) sorted in an increasing order between dominant and non-dominant wrist.
Each data point is the average for all participants from the cut-point calibration sample (N = 42).
ENMO: Euclidean norm minus 1 g with negative values rounded to zero;
LFENMO: Euclidean norm minus 1 g of the low-pass filtered raw accelerations with negative values rounded to zero;
MAD: Mean amplitude deviation

Abbreviations in this page:
ENMO: Euclidean norm minus 1*G*
ISCOLE: International Study of Childhood Obesity, Lifestyle and Environment
LFENMO: ENMO of the low-pass filtered raw accelerations
MAD: mean amplitude deviation
NHANES: National Health And Nutrition Examination Survey
PA: physical activity
SB: sedentary behaviour
VMCounts: activity counts in the vector magnitude
WHS: Women's Health Study

Comparison of each acceleration metric across body sites

Comparisons between each metric derived from different body attachment sites add important knowledge to the field. Studies attaching accelerometers to the hip (e.g., WHS [268], NHANES 2003-2004 or ISCOLE), to the dominant wrist (e.g., UK Biobank), and to the non-dominant wrist (e.g., NHANES 2012-2013, and Whitehall II Study [214]) could reach conflicting conclusions in regards to PA, SB time, and/or sleep outcomes. This study quantifies these potential differences. Accordingly with previous evidence [258], our findings show higher associations between the dominant and non-dominant wrists than between either wrist and the hip for all the metrics included. Comparison of acceleration metrics across the body attachment sites revealed higher shared variances for MAD compared with the rest of the metrics. Furthermore, comparisons between the dominant wrist vs. hip, and the non-dominant wrist vs. hip were similar for all of the metrics included, as occurred in a previous study in adults [260]. In regards to the movement pattern throughout the day, we observed almost identical patterns between the dominant and non-dominant wrists (with slightly lower values for the non-dominant wrist). To a lesser extent, hip movement pattern was similar to those from wrists. This suggests that the relationship between the dominant and non-dominant wrist accelerations is linear, which is also supported by a previous study in adults [258]. This implies that activities across the day should be captured similarly by all sites, especially between the dominant and non-dominant wrist.

Comparison of acceleration metrics derived from the same body attachment site

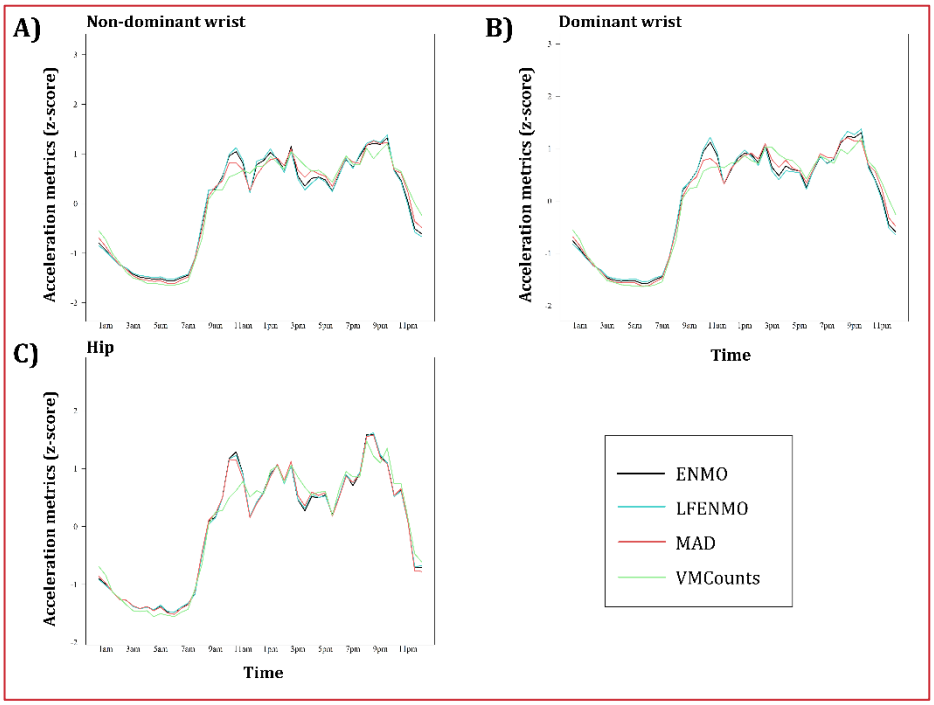


Figure 15
Means of ENMO, LFENMO, MAD and VMCounts, over 30-min periods for non-dominant wrist (Panel A), dominant wrist (Panel B) and hip (Panel C).
Each data point is the average for this time interval for all participants from the cut-point calibration sample (N = 42).
ENMO: Euclidean norm minus 1 g with negative values rounded to zero;
LFENMO: Euclidean norm minus 1 g of the low-pass filtered raw accelerations with negative values rounded to zero;
MAD: Mean amplitude deviation.

Table 18
Cut-point translation from previously proposed non-dominant wrist cut points for dominant wrist ENMO

	Mean (SD) non-dominant wrist (min/day)	Mean (SD) dominant wrist (min/day)	LCCC	Mean diff. [95% CI] (min)[b]	MAPE
Cut-points translation					
SB time thresholds (mg)					
45[a]	769 (69)	744 (71)	0.86	-25 [-55 to 6]	3.25%
50	**-**	**766 (69)**	**0.92**	**-3 [-33 to 27]**	**0.39%**
55	-	786 (67)	0.90	17 [-13 to 46]	2.21%
LPA thresholds (mg)					
45-100[a]	147 (29)	161 (34)	0.84	14 [0 to 28]	9.52%
45-105	-	170 (37)	0.73	23 [9 to 37]	15.65%
45-110	-	179 (38)	0.63	31 [17 to 46]	21.09%
50-100	-	139 (30)	0.91	-8 [-21 to 5]	5.44%
50-105	-	148 (33)	0.93	1 [-12 to 15]	0.68%
50-110	**-**	**157 (34)**	**0.89**	**9 [-4 to 23]**	**6.12%**
55-100	-	120 (27)	0.62	-27 [-40 to -15]	18.37%
55-105	-	129 (29)	0.78	-19 [-31 to -6]	12.93%
55-110	-	137 (31)	0.88	-10 [-23 to 3]	6.80%
MPA thresholds (mg)					
100-430[a]	104 (28)	120 (30)	0.72	16 [3 to 28]	15.38%
100-435	-	120 (30)	0.72	14 [0 to 28]	13.46%
100-440	-	120 (30)	0.72	16 [4 to 29]	15.38%
105-430	-	111 (28)	0.82	7 [-5 to 19]	6.73%
105-435	-	111 (28)	0.81	7 [-5 to 19]	6.73%
105-440	-	111 (28)	0.81	7 [-5 to 19]	6.73%
110-430	-	103 (26)	0.84	-1 [-13 to 10]	0.96%
110-435	-	103 (26)	0.84	-1 [-13 to 11]	0.96%
110-440	**-**	**103 (26)**	**0.84**	**-1 [-13 to 11]**	0.96%
VPA thresholds (mg)					
430[a]	9 (7)	9 (7)	0.95	0 [-3, 3]	0%
435	-	9 (7)	0.95	0 [-3, 3]	0%
440	**-**	**9 (7)**	**0.95**	**0 [-3, 3]**	0%
Cross-validation					
SB	750 (78)[c]	755 (94)[d]	0.85	5 [-36, 45]	0.67%
LPA	134 (34)[c]	133 (36)[d]	0.89	-2 [-18, 15]	1.49%
MPA	103 (44)[c]	97 (44)[d]	0.92	-6 [-26, 15]	5.82%
VPA	7 (7)[c]	7 (7)[d]	0.90	1 [-2, 4]	14.29%

Cut-point selection (values presented in bold) was made following these criteria: 1) closest VPA estimation if any, no threshold is selected since all of them were equally comparable; 2) closest MPA estimation if any, 110-430 m*g*, 110-435 m*g* and 110-440 m*g* provided similar results, so 110-440 m*g* was selected to respect the distance between previously-established cut points (i.e., 100-430 m*g*); 3) closest light PA estimation considering the upper threshold defined in step 2 (i.e., 110 m*g*).
Bold text indicates the cut-point selection based on the criteria defined above.
[a] Indicates original cut points validated in non-dominant wrist [61,62].
[b] Dominant wrist *minus* non-dominant wrist.
[c] Derived with the original cut points validated in non-dominant wrist, i.e., 45/100/430 [61,62].
[d] Derived with the cut points proposed in the present study for dominant wrist, i.e., 50/110/440.

Our findings show moderate to high associations between pairs of acceleration metrics derived from the same body attachment site. Likewise, previous findings by van Hees et al. [67] reported moderate to high shared variances between ENMO and other acceleration metrics (r^2 from 0.46 to 0.95) not included in the present study. Also in concordance with van Hees et al.'s study [67] in adults, our findings show stronger associations between pairs of metrics with none or minimal filtering (ENMO, LFENMO and MAD) than the comparisons of any of these metrics with VMCounts, which may be explained by the application of a frequency filter to the raw signal. We also found that the movement pattern identified throughout the day was visually similar for ENMO, LFENMO and MAD, while VMCounts did not identify some peaks of movement detected by the rest of the metrics. The current study complements the study by van Hees et al. [67] by using 24 hours of accelerometer recording for both the hip (only waking hours in the previous study) and both wrists, as well as including different acceleration metrics for comparison, i.e., LFENMO,

Abbreviations in this page:
CI: confidence interval
ENMO: Euclidean norm minus 1*G*
LFENMO: ENMO of the low-pass filtered raw accelerations
LCCC: Lin's concordance correlation coefficient
LPA: light physical activity
MAD: mean amplitude deviation
MAPE: mean absolute percent error
MPA: moderate physical activity
SB: sedentary behaviour
SD: standard deviation
VMCounts: activity counts in the vector magnitude
VPA: vigorous physical activity

Abbreviations in this page:
LPA: light physical activity
MAD: mean amplitude deviation
MPA: moderate physical activity
MVPA: moderate-to-vigorous physical activity
PA: physical activity
SB: sedentary behaviour
VMCounts: activity counts in the vector magnitude
VPA: vigorous physical activity

MAD and ActiGraph's VMCounts [67,147,153,204]. Furthermore, pairwise comparisons of acceleration metrics showed better agreement during waking hours than during sleeping hours in all the body attachment sites. However, it is important to note that absolute values of all acceleration metrics are lower during sleeping hours, which could produce these lower shared variances even when the absolute differences are rather small (see descriptive values in Table 2). To the best of our knowledge, this is the first study providing this comparison stratified by waking and sleeping hours.

Cut-points replication

The fact that there are only cut points available to assess PA and SB time from the non-dominant wrist makes their application to data from the dominant wrist controversial. Indeed, differences found between acceleration curves from the dominant and non-dominant wrists indicate the need to propose new cut points for the dominant wrist. Accordingly, we detected 25 min/day less of SB time, 14 min/day more of LPA, 16 min/day more of MPA and similar estimations of VPA using the original cut points (i.e., for non-dominant wrist) in data collected from the dominant wrist. The linearity in the association between wrists (i.e., consistent movement pattern detected from both the dominant and non-dominant wrists) make it possible to adapt cut points developed for one of the wrists to the other by applying slightly different new cut points. This strategy has been used previously by Rowlands et al. to replicate hip-based MVPA cut points using non-dominant wrist data [223]. In the present study, we develop cut points for the dominant wrist using previously validated cut points for the non-dominant wrist as reference [61,62]; and then, we cross-validated these newly developed cut points in a different sample. Estimations from previously validated cut points on the non-dominant wrist and their translation to the dominant wrist were almost equal and highly correlated. The cut points we propose in this study for the dominant wrist could help to obtain equivalent and comparable estimations of PA between studies using the dominant wrist with studies using the non-dominant wrist.

This study complements existing information by using a long-term measurement (7 days) for these comparisons since previous studies used 1-day measurements [258,260]. Furthermore, our stratified analyses for waking and sleeping hours allow an understanding of how acceleration metrics agree or disagree for PA, SB time and sleep-related estimations. Notably, our 24-hour-based comparisons across body attachment sites and acceleration metrics are similar to waking-hour-based comparisons.

This study has practical implications for studies using the same acceleration metric, but attaching accelerometers to different body sites or vice versa. Furthermore, accelerometers are not only used to

estimate PA and SB time, but also to assess sleep. Researchers focused on any of these behaviours can benefit from the comparisons presented in this study across acceleration metrics and body attachment sites during waking and sleeping hours, since these associations were different depending on the period of the day analysed. Finally, this study provides information to quantify methodological discrepancies across studies, as it provides cut points to get similar PA estimations from dominant wrist and non-dominant wrist. We suggest these cut points are used to obtain comparable estimations across studies. Differences found between sites and acceleration metrics do not constitute different associations between SB, PA, and/or sleep outcomes with health parameters. Whether associations with health parameters differ depending on data collection and processing decisions should be studied by future research.

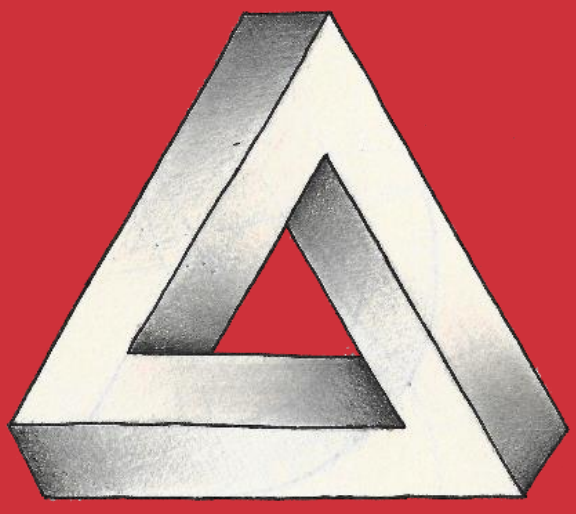
The main limitation of this study is the lack of a criterion that would allow us to assess the accuracy of each acceleration metric in the measurement of PA and/or sleep. Likewise, the lack of an energy expenditure measure precludes us from deriving cut points for dominant wrist against a criterion. Thus, although our derived cut points may be of great value to identify PA from the dominant wrist, these cut points should be tested against an energy expenditure measure in future studies. Another limitation is the use of a convenience sample and all analyses were only carried out with data from one accelerometer brand (ActiGraph GT3X+), which could limit the generalization of our findings to other devices [269,270], or even to different generations of the same brand [174]. Strengths of the present study were: 1) the fact that we used consistent data processing techniques with all the metrics (e.g., same calculation of non-wear time or waking and sleeping hours detection) which allow for a direct comparison between metrics and body attachment sites; and 2) the fact that our participants reached high wear times, allowing for a complete range of daily living accelerations.

Abbreviations in this page:
ENMO: Euclidean norm minus 1*G*
LFENMO: ENMO of the low-pass filtered raw accelerations
MAD: mean amplitude deviation
PA: physical activity
SB: sedentary behaviour
VMCounts: activity counts in the vector magnitude

Conclusion

In conclusion, our findings suggest higher acceleration metric values in the dominant wrist compared with the non-dominant wrist. These differences could be attenuated by applying the new set of cut points provided in this manuscript. Furthermore, ENMO and LFENMO were the metrics that compared the best, and to some extent, they also showed good comparability with MAD for daily average values and for the movement pattern identified throughout the day. However, VMCounts were demonstrated to be less comparable to the previously-mentioned metrics. Future studies should test which of these metrics and body locations are the best to accurately capture PA against a criterion (e.g., calorimetry).

Study V



Comparability of published cut-points for the assessment of physical activity: Implications for data harmonization

Migueles JH, Cadenas-Sanchez C, Tudor-Locke C, Löf M, Esteban-Cornejo I, Molina-Garcia P, Mora-Gonzalez J, Rodriguez-Ayllon M, Garcia-Marmol E, Ekelund U, Ortega FB.
Scand J Med Sci Sports. 2019 Apr;29(4):566-574
DOI: 10.1111/sms.13356

Contents

Abstract

Background | Accelerometers are the method of preference to assess PA in research at the moment. The rapid development of calibration studies has resulted in a number of cut-points to quantify time spent in SB and PA intensities in children. Information on the comparability between cut-points is limited.

Aims | This study aimed to compare estimations of SB and time spent in PA intensities in children with overweight or obesity across different age-appropriate cut-points based on different body-worn attachment sites and acceleration metrics.

Methods | A total of 104 overweight/obese children (10.1 ± 1.1 years old, 43 girls) concurrently wore ActiGraph GT3X+ accelerometers on their right hip and non-dominant wrist for 7 days (24 hours). ENMO, VACounts, and VMCounts were derived. We calculated estimates of SB and LPA, MPA, VPA, and MVPA using different published cut-points for children. The prevalence of children meeting the recommended 60 min/d of MVPA was calculated.

Results | The time spent in SB and the different PA intensities largely differed across cut-points based on different attachment sites and acceleration metrics (i.e., SB = 11-252 min/d; LPA = 10-217 min/d; MPA = 1-48 min/d; VPA = 1-35 min/d; MVPA = 4-66 min/d). Consequently, the prevalence of children meeting the recommended 60 min/d of MVPA varied from 8% to 96% of the study sample.

Conclusion | The present study provides a comprehensive comparison between available cut-points for different attachment and acceleration metrics in children. Furthermore, our data clearly show that it is not possible (and probably will never be) to know the prevalence of meeting the PA guidelines based on accelerometer data since apparent differences range from almost zero to nearly everyone meeting the guidelines.

Abbreviations in this page:
ENMO: Euclidean norm minus 1 *G*
LPA: light physical activity
MPA: moderate physical activity
MVPA: moderate-to-vigorous physical activity
PA: physical activity
SB: sedentary behaviour
VACounts: activity counts in the vertical axis
VMCounts: activity counts in the vector magnitude
VPA: vigorous physical activity

Key Points

Question
Do the use of different cut-points substantially affect PA and SB estimations?

Findings
Large differences were found, independently of the body attachment site and the acceleration metric used

Meaning
The prevalence of physical inactivity drastically differed across cut-points from nearly none to nearly everyone meeting the public health recommendations on PA

Introduction

Abbreviations in this page:
NHANES: National Health And Nutrition Examination Survey
PA: physical activity
SB: sedentary behavior

Accurate and objective estimations of daily time in SB and PA are important to estimate the prevalence of populations meeting the current PA guidelines, to assess the success of interventions aiming to increase PA in specific populations, to explore population activity trends, and to quantify the dose-response impact of SB and PA on health [254]. Accelerometers are feasible tools to objectively assess time in SB and PA in large-scale studies, but their utilization requires standardized data collection (e.g., attachment site) and processing criteria (e.g., how to filter the raw accelerations), both demonstrating a high potential to affect the estimation of PA [233]. Additionally, protocols and methods vary largely across studies which aims to develop cut-points (e.g., differences in the exercise protocols or the measurement of energy expenditure), resulting in differences in the identification and application of cut-points, i.e., intensity thresholds for SB and PA intensity classification. Since SB refers to any waking behaviour in a reclining posture with requires low related energy expenditure [19], it is important to note that SB time estimations based on cut-points are limited because they are not able to detect changes in posture. Many authors have called for a harmonization of data collection, processing criteria, and selection of cut-points to assess SB time and PA in order to gain comparability between studies [233,271,272]. This harmonization would be of special interest to compare data across studies, especially when the populations assessed are similar. To date, such harmonization and consensus is not available.

Data collection decisions include selecting a device, the body attachment site (i.e., hip or wrist in the majority of studies) and the sampling frequency for the recording (usually between 30-100 Hz) [233]. The traditional hip attachment site is being replaced with a wrist location by some consumer-grade manufacturers (e.g., FitBit, Polar, Garmin, or Up) and by large-scale studies, such as the US NHANES and the UK Biobank. This strategy was undertaken as an effort to obtain a higher wear compliance [94,171,233]. Both hip and wrist attachment sites have been validated for classifying PA intensities [61,69,70,233], and are potentially able to assess energy expenditure during free-living conditions in different populations [95,222], yet due to differences in the protocols used in cut-point validation studies it is unknown how well measures from the hip and wrist compare to each other.

The main purpose of processing criteria is to get a clean estimate of body accelerations by removing gravity acceleration and noise from the acceleration signal. The first commercially available accelerometers coerced researchers into using the manufacturer's activity counts (i.e., accelerations due to body movement) from the

Abbreviations in this page:
ENMO: Euclidean norm minus 1*G*
LPA: light physical activity
MVPA: moderate-to-vigorous physical activity
PA: physical activity
RCT: randomized controlled trial
SB: sedentary behavior
VA: vertical axis
VM: vector magnitude

VA or VM derived from proprietary algorithms. These activity counts were hardly comparable between devices, or even between different models from the same manufacturer [175,199]. However, contemporary accelerometers are capable of storing high-frequency raw accelerations, which are highly comparable between frequently used research-grade devices (i.e., ActiGraph, GENEActiv, and Axivity) [273]. In the last five years, researchers have published open source methods to process raw accelerations in order to obtain alternative acceleration metrics to activity counts [67,147]. ENMO is now widely used and has shown a high agreement between brands [225,273], facilitating data harmonization across studies.

As the process of harmonizing data collection and processing criteria proceeds, it is important to study how different body attachment sites, acceleration metrics, and cut-points affect the final estimations of SB and PA intensities. Rowlands et al. reported a moderate agreement between MVPA estimates derived using different cut-points based on ENMO from wrist accelerations and classical activity counts thresholds based on hip-worn devices [223]. In contrast, other studies comparing cut-points developed independently for different attachment sites and acceleration metrics have reported large differences across MVPA estimates in adolescents [271] and adults [272]. Although there is an increasing interest in the study of SB and LPA [274], previous studies have only focused on MVPA.

Therefore, there is a need to better understand how data collection, processing criteria, and cut-points influence estimations of SB and PA in different populations, including children and those classified as overweight or obese. Thus, this study aimed to examine how cut-points relative to different attachment sites and acceleration metrics affect the final estimations of SB and PA in children with overweight or obesity.

Methods

The present cross-sectional study analysed data from the baseline assessment of the ActiveBrains Project (http://profith.ugr.es/activebrains). A detailed description of the study design and methods has been published elsewhere [63]. Briefly, ActiveBrains is a RCT intended to examine the effect of a 20-week PA intervention on brain structure, function, cognitive performance, academic achievement, and physical and mental health outcomes in overweight or obese children [63]. A total of 110 overweight or obese children (classified based on the WOF cut-points [64,65]) were recruited from Granada (Spain). A final sample of 104 children (10.1 ± 1.1 years of age, 41% girls) met the accelerometry inclusion criteria (more details below). The data were collected between November 2014 and February 2016. We informed the parents or legal guardians about the purpose

of the study, and we obtained written informed parental consent. The ActiveBrains project was approved by the Human Research Ethics Committee of the University of Granada, and was registered as a clinical trial (NCT02295072, http://clinicaltrials.gov).

The participants' anthropometry, SB, and PA were assessed as part of the protocol of the ActiveBrains project [63]. Briefly, we measured the body weight and height to the nearest 0.1 kg and 0.1 cm using an electronic scale (SECA 861, Hamburg, Germany) and a precision stadiometer (SECA 225, Hamburg, Germany), respectively. BMI was calculated as kg/m^2. The participants were also required to concurrently wear two accelerometers (ActiGraph GT3X+, Pensacola, FL, USA) for 7 complete days (24 hours): one on the right hip and the other on the non-dominant wrist. The participants were instructed to wear the accelerometers as many hours as possible and to remove them only for water activities (i.e., shower or swimming), and both at the same time. Concomitantly, the participants reported the time they went to bed and woke-up in a diary log throughout the study.

Abbreviations in this page:
ENMO: Euclidean norm minus 1*G*
PA: physical activity
SB: sedentary behavior
SD: standard deviation
VACounts: activity counts in the vertical axis
VMCounts: activity counts in the vector magnitude

Table 19
Children's age-appropriate cut-points for the estimation of SB and PA intensities

References	Attachment site	Acceleration metric	Epoch length	SB/LPA	LPA/MPA	MPA/VPA
Hildebrand et al.	Hip	ENMO	5 sec	63 m*g*	143 m*g*	465 m*g*
Hildebrand et al.	Wrist	ENMO	5 sec	36 m*g*	201 m*g*	707 m*g*
Hänggi et al.	Hip	VMCounts	1 sec	3 counts	56 c	-
Romanzini et al.	Hip	VMCounts	15 sec	180 counts	757 counts	1112 counts
Chandler et al.	Wrist	VMCounts	5 sec	305 counts	818 counts	1969 counts
Evenson et al.	Hip	VACounts	15 sec	25 counts	574 counts	1003 counts

ActiGraph GT3X+ is a triaxial accelerometer with a dynamic range of +/- 6 G. Both hip- and wrist-worn accelerometers were initialized to capture and store accelerations at 100 Hz. The raw accelerations were then downloaded and converted to ".csv" format using ActiLife v.6.13.3 (ActiGraph, Pensacola, FL, USA). Raw ".csv" files were imported to R software (v. 3.1.2, https://www.cran.r-project.org/) and processed using the GGIR package (v. 1.5-12, https://cran.r-project.org/web/packages/GGIR/). They were also imported and processed in the ActiLife software (ActiGraph, Pensacola, FL, USA) to obtain VMCounts and VACounts using the normal filter developed by ActiGraph. The processing methods involved: 1) Auto-calibration of the data according to the local gravity [66]. 2) Detection of the non-wear time based on the raw acceleration of the three axes [67]. Briefly, each 15-min block was classified as non-wear time if the SD of 2 out of the 3 axes was lower than 13 m*g* during the surrounding 60-min moving window, or if the value range for 2 out of the 3 axes was lower than 50 m*g*. 3) Detection of sustained abnormal high accelerations, i.e., higher than 5.5 g. 4) Calculation of the ENMO. 5) Importation of the VMCounts and VACounts ".csv" files to R to follow the same processing criteria than ENMO. 6) Imputation of detected non-wear time and abnormal high accelerations by means of the acceleration for the rest of the recording period during the same time interval than the affected periods. 7) Identification of waking and sleeping

GGIR vignette
(scan or click here)

Abbreviations in this page:
ANOVA: analysis of variance
BMI: body mass index
ENMO: Euclidean norm minus 1*G*
LPA: light physical activity
MPA: moderate physical activity
MVPA: moderate-to-vigorous physical activity
PA: physical activity
SB: sedentary behavior
SD: standard deviation
VACounts: activity counts in the vertical axis
VMCounts: activity counts in the vector magnitude
VPA: vigorous physical activity

hours using an automatized algorithm guided by the times reported by the participants [40]. Waking and sleeping hours were detected using data from the non-dominant wrist and detected times were then matched to the right hip data for each participant. And, 8) Estimation of SB and PA intensities using different age-appropriate cut-points for ENMO, VMCounts, and VACounts as detailed in **Table 19**.

Mean daily SB and PA intensity levels were then calculated as: (mean of valid weekdays*5 + mean of valid weekend days*2) / 7. The participants were excluded from the analyses if they recorded less than 4 valid days (i.e., ≥ 16 hours/day), including at least 1 weekend day. Out of the 110 participants, 4 children recorded less than 4 days of valid wearing time, 1 accelerometer attached to the non-dominant wrist malfunctioned, and 1 participant was excluded for having mean acceleration values during nights between 6-9 SDs above the group mean. Thus, a final sample of 104 participants was included in the present study.

Table 20
Anthropometry, SB, and PA characteristics of participants

	All (n=104)	Boys (n=61)	Girls (n=43)	*P* sex
Anthropometry				
Age (years)	10.1 ± 1.1	10.2 ± 1.2	9.9 ± 1.1	0.248
Weight (kg)	56.2 ± 10.8	56.8 ± 10.7	55.4 ± 11.1	0.533
Height (cm)	144.3 ± 8.3	144.9 ± 7.9	143.6 ± 8.9	0.443
BMI (kg/m^2)	26.8 ± 3.5	26.9 ± 3.6	26.7 ± 3.5	0.766
Wearing time during waking hours				
Hip device (hours/day)	15.0 ± 0.6	15.1 ± 0.6	15.0 ± 0.6	0.569
Wrist device (hours/day)	14.8 ± 0.6	14.8 ± 0.5	14.8 ± 0.6	0.926
SB (min/day)				
Hip ENMO $_{\text{Hildebrand}}$	817.4 ± 44.7	811.1 ± 42.9	826.3 ± 46.2	0.093
Wrist ENMO $_{\text{Hildebrand}}$	565.1 ± 56.4	560.5 ± 56.3	571.6 ± 56.5	0.327
Hip VMCounts $_{\text{Hänggi}}$	639.1 ± 64.8	634.4 ± 58.3	645.5 ± 73.1	0.412
Hip VMCounts $_{\text{Romanzini}}$	628.3 ± 68.2	623.9 ± 65.7	634.5 ± 71.8	0.445
Wrist VMCounts $_{\text{Chandler}}$	576.4 ± 53.9	577.4 ± 54.7	575.1 ± 53.3	0.828
Hip VACounts $_{\text{Evenson}}$	600.6 ± 70.1	593.0 ± 69.7	611.1 ± 69.9	0.198
LPA (min/day)				
Hip ENMO $_{\text{Hildebrand}}$	65.8 ± 15.8	68.4 ± 15.6	62.1 ± 15.5	**0.043**
Wrist ENMO $_{\text{Hildebrand}}$	282.7 ± 38.5	279.3 ± 37.1	287.4 ± 40.3	0.298
Hip VMCounts $_{\text{Hänggi}}$	176.9 ± 38.0	175.0 ± 33.3	179.5 ± 44.1	0.579
Hip VMCounts $_{\text{Romanzini}}$	198.2 ± 41.5	193.6 ± 39.4	204.5 ± 44.0	0.197
Wrist VMCounts $_{\text{Chandler}}$	239.0 ± 29.5	235.4 ± 29.2	244.0 ± 29.6	0.144
Hip VACounts $_{\text{Evenson}}$	273.1 ± 52.1	276.4 ± 52.0	268.5 ± 52.5	0.452
MPA (min/day)				
Hip ENMO $_{\text{Hildebrand}}$	32.9 ± 13.9	37.5 ± 14.7	26.5 ± 9.6	**<0.001**
Wrist ENMO $_{\text{Hildebrand}}$	47.5 ± 17.4	54.2 ± 18.4	38.1 ± 10.2	**<0.001**
Hip VMCounts $_{\text{Romanzini}}$	53.8 ± 14.4	57.9 ± 14.8	48.0 ± 11.7	**<0.001**
Wrist VMCounts $_{\text{Chandler}}$	81.2 ± 20.1	83.3 ± 22.7	78.4 ± 15.8	0.201
Hip VACounts $_{\text{Evenson}}$	33.8 ± 11.5	37.9 ± 12.2	28.2 ± 7.4	**<0.001**
VPA (min/day)				
Hip ENMO $_{\text{Hildebrand}}$	3.0 ± 2.0	3.7 ± 2.1	2.1 ± 1.4	**<0.001**
Wrist ENMO $_{\text{Hildebrand}}$	7.6 ± 4.4	9.4 ± 4.5	5.0 ± 2.7	**<0.001**
Hip VMCounts $_{\text{Romanzini}}$	37.9 ± 16.1	44.2 ± 16.5	29.1 ± 10.6	**<0.001**
Wrist VMCounts $_{\text{Chandler}}$	6.2 ± 3.6	7.4 ± 3.7	4.6 ± 2.7	**<0.001**
Hip VACounts $_{\text{Evenson}}$	10.7 ± 6.7	12.4 ± 7.6	8.3 ± 4.4	**0.001**
MVPA time (min/day)				
Hip ENMO $_{\text{Hildebrand}}$	36.0 ± 15.3	41.2 ± 16.1	28.6 ± 10.6	**<0.001**
Wrist ENMO $_{\text{Hildebrand}}$	55.1 ± 21.0	63.7 ± 22.0	43.1 ± 11.9	**<0.001**
Hip VMCounts $_{\text{Hänggi}}$	102.4 ± 26.8	110.6 ± 26.4	90.9 ± 23.1	**<0.001**
Hip VMCounts $_{\text{Romanzini}}$	91.7 ± 28.2	102.1 ± 28.7	77.1 ± 20.0	**<0.001**
Wrist VMCounts $_{\text{Chandler}}$	87.5 ± 22.5	90.6 ± 25.4	83.0 ± 16.9	0.071
Hip VACounts $_{\text{Evenson}}$	44.5 ± 16.7	50.2 ± 18.1	36.6 ± 10.3	**<0.001**

Data are presented as mean ± SD. Statistically significant values are shown in bold.

Descriptive statistics were calculated as means and SDs. The time estimates of SB, LPA, MPA, VPA, and MVPA were compared between each pair of estimations (i.e., estimations from each pair of cut-

points) using repeated measures analysis of variance (ANOVA). Additionally, we inspected the distributions of the time spent in MVPA and the prevalence of the study sample meeting the PA guidelines (i.e., at least 60 min/day of MVPA) [8] using different cut-points. All analyses were performed in R. Overall, the significance level was set at $p<0.05$ for all the analyses; however, in order to account for multiple comparisons, significant differences at $p<0.01$ were interpreted as statistically meaningful.

Abbreviations in this page:
ANOVA: analysis of variance
ENMO: Euclidean norm minus 1 *G*
LPA: light physical activity
MPA: moderate physical activity
MVPA: moderate-to-vigorous physical activity
PA: physical activity
SB: sedentary behavior
SD: standard deviation
VACounts: activity counts in the vertical axis
VMCounts: activity counts in the vector magnitude
VPA: vigorous physical activity

Results

The anthropometric characteristics, the time spent in SB, and the various PA intensities (calculated using the different cut-points) are reported in **Table 20**.

The comparisons between SB and PA intensities estimated from the different cut-points are graphically presented in **Figure 16**. The differences expressed in min/day between different cut-point estimates are shown in **Table 21**. Nearly every pairwise comparison was significantly different (all $p < 0.05$) (exceptions are shown in **Table 21**). Overall, the various mean daily estimations differed between 11-252 min/day for SB, 10-217 min/day for LPA, 1-48 min/day for MPA, 1-35 min/day for VPA, and 4-66 min/day for MVPA.

Figure 17 presents the time distributions spent in MVPA for the different cut-points examined. Overall, this figure shows that cut-points based on VMCounts produced higher MVPA time compared to those estimations based on ENMO or VACounts, independently of the attachment site (as reported in **Table 21**).

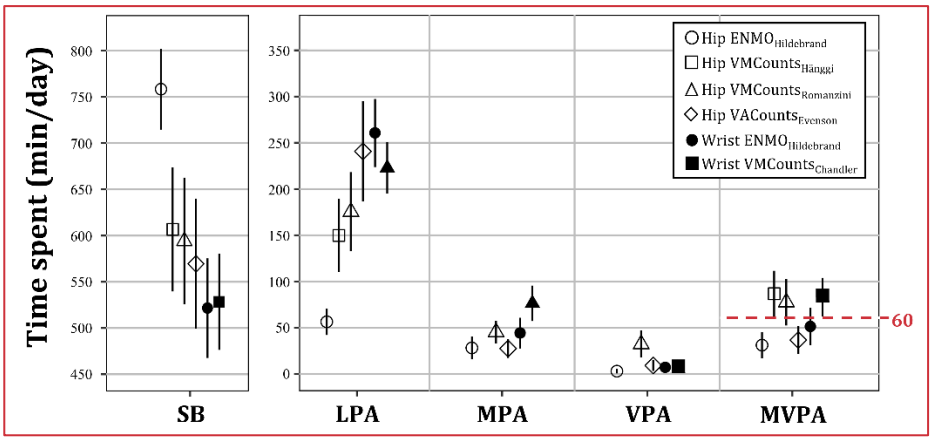


Figure 16
Mean daily time spent (min) and SDs (error bars) in SB and PA intensities considering different attachment sites and metrics.

Figure 18 shows that the sample prevalence meeting the recommended 60 min/day of MVPA per day ranged from 8% to 96% depending on the cut-points applied to the data. Overall, the prevalence of meeting the PA guidelines was higher for boys than for girls using all cut-points except for the Chandler et al. [70] cut-points (i.e., 90% of the boys versus 95% of the girls met the PA guidelines, accordingly).

Discussion

Abbreviations in this page:
CI: confident interval
ENMO: Euclidean norm minus 1*G*
LPA: light physical activity
MPA: moderate physical activity
MVPA: moderate-to-vigorous physical activity
PA: physical activity
SB: sedentary behaviour
VACounts: activity counts in the vertical axis
VMCounts: activity counts in the vector magnitude
VPA: vigorous physical activity

The primary purpose of this study was to provide a clear picture of which cut-points are more and less comparable in free-living conditions in children with overweight or obesity, including traditional (e.g., Evenson cut-points based on VACounts [68]) and recently developed (e.g., Hildebrand cut-points based on ENMO [61,62], Romanzini [69] and Chandler [70] cut-points based on VMCounts) cut-points, and when the accelerometer was attached to the hip and wrist. Contrary to what could have been expected, all cut-points based on VMCounts produced significantly higher estimations of time spent in MVPA than ENMO and VACounts cut-points, regardless of the attachment site. To our knowledge, this is the first study investigating differences across accelerometer-based estimations of SB and PA intensities using a complete set of available cut-points, running from the most traditionally used cut-points for VACounts detected from a hip attachment, i.e., the Evenson et al. [68] cut-points, to the newly developed cut-points for ENMO [61,62] and VMCounts [69,70,108] from both hip and non-dominant wrist attachments.

Table 21
T-tests for the comparison between SB, LPA, MPA, VPA, and MVPA calculated from different cut-points.

	SB (min/day) Difference (95%CI)	LPA (min/day) Difference (95%CI)	MPA (min/day) Difference (95%CI)	VPA (min/day) Difference (95%CI)	MVPA (min/day) Difference (95%CI)
Hip vs. hip					
$ENMO_{Hildebrand}$ - $VMCounts_{Hänggi}$	178 (163 to 194)**	-111 (-119 to -103)**			-66 (-72 to -60)**
$ENMO_{Hildebrand}$ - $VMCounts_{Romanzini}$	189 (173 to 204)**	-132 (-141 to -124)**	-21 (-25 to -17)**	-35 (-38 to -32)**	-56 (-62 to -49)**
$ENMO_{Hildebrand}$ - $VACounts_{Evenson}$	217 (201 to 233)**	-207 (-218 to -197)**	-1 (-4 to 3)	-8 (-9 to -6)**	-9 (-13 to -4)**
$VMCounts_{Romanzini}$ - $VMCounts_{Hänggi}$	-11 (-29 to 8)	21 (10 to 32)**			-11 (-18 to -3)**
$VMCounts_{Romanzini}$ - $VACounts_{Evenson}$	28 (9 to 46)*	-75 (-88 to -62)**	20 (16 to 23)**	27 (24 to 31)**	47 (41 to 54)**
$VMCounts_{Hänggi}$ - $VACounts_{Evenson}$	38 (20 to 57)**	-96 (-109 to -84)**			58 (52 to 64)**
Wrist vs. wrist					
$VMCounts_{Chandler}$ - $ENMO_{Hildebrand}$	11 (-4 to 26)	-44 (-53 to -34)**	34 (29 to 39)**	-1 (-2 to 0)*	32 (26 to 38)**
Hip vs. wrist					
$ENMO_{Hildebrand}$ - $ENMO_{Hildebrand}$	252 (238 to 266)**	-217 (-225 to -209)**	-15 (-19 to -10)**	-5 (-6 to -4)**	-19 (-24 to -14)**
$VMCounts_{Hänggi}$ - $VMCounts_{Chandler}$	63 (46 to 79)**	-62 (-71 to -53)**			15 (8 to 22)**
$VMCounts_{Romanzini}$ - $VMCounts_{Chandler}$	52 (35 to 69)**	-41 (-51 to -31)**	-27 (-32 to -23)**	32 (28 to 35)**	4 (-3 to 11)
$ENMO_{Hildebrand}$ - $VMCounts_{Chandler}$	-241 (-255 to -227)**	-173 (-180 to -167)**	-48 (-53 to -44)**	-3 (-4 to -2)**	-52 (-57 to- 46)**
$VMCounts_{Hänggi}$ - $ENMO_{Hildebrand}$	74 (57 to 91)**	-106 (-116 to -95)**			47 (41 to 54)**
$VMCounts_{Romanzini}$ - $ENMO_{Hildebrand}$	63 (46 to 80)**	-85 (-95 to -74)**	6 (2 to 11)*	30 (27 to 34)**	37 (30 to 43)**
$VACounts_{Evenson}$ - $ENMO_{Hildebrand}$	35 (18 to 53)**	-10 (-22 to 3)	-14 (-18 to -10)**	3 (2 to 5)**	-11 (-16 to -5)**
$VACounts_{Evenson}$ - $VMCounts_{Chandler}$	24 (7 to 41)*	34 (22 to 46)**	-47 (-52 to -43)**	4 (3 to 6)**	-43 (-48 to -37)**

Data are presented as mean differences and 95% of CI.
Hildebrand et al. [61,62], Hänggi et al. [108], Romanzini et al. [69], Chandler et al. [70], and Evenson et al. [68].
* $P < 0.05$
** $P < 0.01$

Since the selection of the different data collection and processing criteria are known to affect SB and PA intensity estimations [233], we applied cut-points specifically developed for the two different attachment sites for use in children. We also followed the same processing criteria (i.e., same acceleration metric and epoch length) as originally used in validation studies. In agreement with recent studies [272,275], our results confirm non-comparable estimates of the time spent in MVPA when using different data collection and processing criteria. However, the present study expands upon this knowledge by additionally comparing estimates of SB and a complete range of PA intensities in a sample of overweight or obese children.

Each of these metrics also displayed non-comparable estimates with large differences between cut-points (see **Table 21** and **Figure 16**). Hildebrand et al. [61,62] developed two sets of cut-points in the same sample to get similar estimations of SB and PA intensities from the hip and the non-dominant wrist. In contrast, herein the estimations for SB and PA for all intensities varied greatly when using the Hildebrand et al. cut-points [61,62] for hip and wrist. This inconsistent result agrees with the Smith et al. findings [271], who reported different estimations derived from two sets of cut-points developed in the same sample and differing only in the acceleration metrics (i.e., VACounts and VMCounts). Our results, together with those from Smith et al. [271], confirm that cut-points from different attachment sites or different acceleration metrics that are comparable in a certain sample could largely differ in others as a result of population-specific features, which may contribute to these differences in SB and PA estimations.

Abbreviations in this page:
ENMO: Euclidean norm minus 1*G*
MVPA: moderate-to-vigorous physical activity
VACounts: activity counts in the vertical axis
VMCounts: activity counts in the vector magnitude

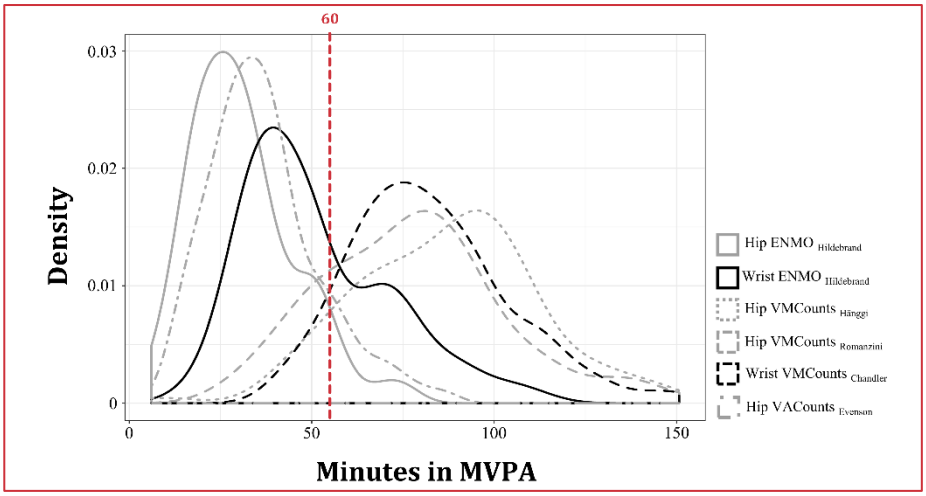


Figure 17
Distributions of the time spent in MVPA intensity (min/day) estimated with different cut-points.

Rowlands et al. [223] looked for ENMO-based cut-points from the non-dominant wrist which could replicate the traditional PA estimations from the Evenson et al. [68] cut-points (applied to VACounts from the hip). Specifically, they reported moderate agreement (ICC = 0.76) and 2 min/day more of MVPA when applying a cut-point of 250 m*g* for ENMO from wrist compared to the Evenson et al. [68] cut-point. Accordingly, we used a lower cut-point for MVPA for ENMO wrist (i.e., 200 m*g* – validated by Hildebrand et al. [61]) and detected 15 min/day more of MVPA from ENMO wrist compared with the Evenson et al. [68] cut-point on hip. Thus, higher values of MVPA can be expected when using the cut-point by Hildebrand et al. [61] for ENMO wrist compared to the MVPA threshold by Evenson et al. [68] for VACounts hip. A more comparable threshold to identify MVPA from ENMO wrist could be 250 m*g* [223].

Taking these findings into consideration, the selection of cut-points to estimate PA intensities with accelerometers is a major obstacle to overcome in objective monitoring since different cut-points could lead to wildly discrepant conclusions. For example, in our

Abbreviations in this page:
ENMO: Euclidean norm minus 1*G*
ICAD: International Children's Accelerometry Database
MVPA: moderate-to-vigorous physical activity
PA: physical activity
SB: sedentary behaviour

sample, the prevalence of boys meeting the 60 min/day of MVPA was higher than that for girls for all the cut-points except for the Chandler et al. [70] cut-points, for which the prevalence was higher in girls than in boys, i.e., 95% vs. 90%. Likewise, **Figure 18** shows large differences in the prevalence of our sample meeting the PA guidelines (i.e., from 8% to 96%), so the fundamental query regarding the prevalence of the population achieving healthful levels of PA is still unresolved. In this regard, Leinonen et al.[276] found moderate-to-high agreement between different methods to classify adults meeting the PA guidelines. It is important to consider that PA guidelines have been developed predominantly using self-reported data, thus, these estimations should be considered with caution. Several authors have proposed reporting PA using a full range of different accelerometer data collection and processing criteria until a consensus is reached [271,272]. However, this is not practical since reporting different and multifactor methodologies could require long explanations and high technical expertise from readers to understand these nuanced inconsistencies. Data pooling and reanalysing raw accelerometer data may be a solution to overcome processing criteria inconsistencies and have been successfully applied (e.g., International Children's Accelerometry Database [ICAD]).

Although estimations of SB and PA intensities are easily understandable for the general population, we suggest that all studies using accelerometers should also report other PA indicators which are not influenced by cut-points, e.g., mean of the acceleration metric per day. As a first step to achieve this, we suggest using research-derived metrics, such as ENMO, which provides a valid estimate of free-living PA from hip and wrist attachments [61,62,67]. Furthermore, in contrast to traditional activity counts, such metrics enable comparability between devices [269,273] and they may be easier to interpret since the acceleration is expressed using a SI unit (i.e., m*g*). In fact, ENMO can be easily implemented in epidemiological studies using the GGIR software implemented in R (https://cran.r-project.org/web/packages/GGIR/). Studies providing normative values for these acceleration metrics will ease the interpretation of findings in the PA measurement field. Furthermore, these normative values could help to identify acceleration values corresponding to meeting the PA guidelines, which could help to obtain a direct measure unaffected by the limitations shown by the cut-points.

Some limitations with this study should be acknowledged: 1) the sample analysed herein was composed of overweight or obese children, and the results may not be generalizable to other populations; 2) the current study did not have a criterion measure for comparison that would allow us to assess the accuracy of each set of cut-points; and, 3) we used 90 accelerometers randomly placed in either

hip or wrist. It could be hypothesized that the use of different accelerometer units is a source of error for the measurement. However, ActiGraph GT3X devices have shown to provide reliable estimations [165], so we assume this source of error is likely to be very small in this study. Furthermore, all the estimates are derived from the same recordings, in case there is a device-related error, this error would be constant in all the estimates presented, and so, it is unlikely this will affect the findings. In contrast, this study's advantages are 1) the use of consistent data processing techniques with all the acceleration metrics (i.e., same calculation of non-wear time, waking and sleeping hours, which allow for a direct comparison between attachment sites, and acceleration metrics); and, 2) that the participants achieved high wearing time compliance, enabling the collection of a complete range of daily living accelerations.

Abbreviations in this page:
LPA: light physical activity
MPA: moderate physical activity
MVPA: moderate-to-vigorous physical activity
PA: physical activity
SB: sedentary behaviour
VPA: vigorous physical activity

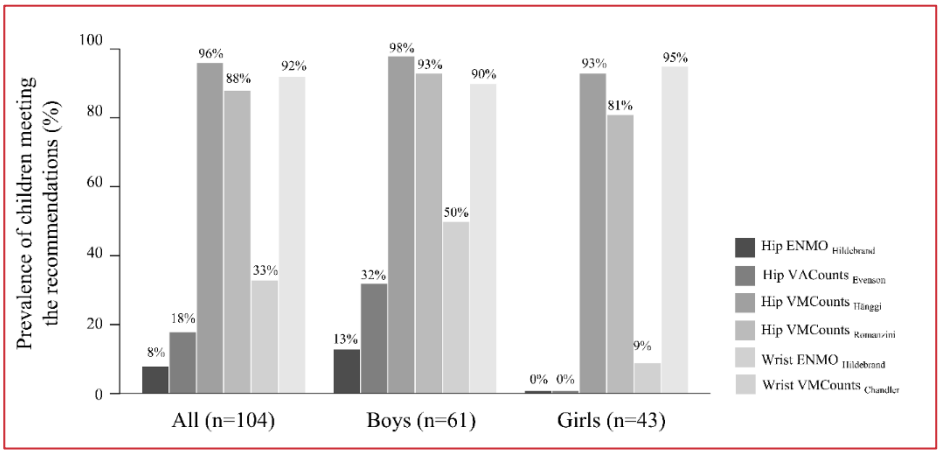


Figure 18
Prevalence of children meeting the PA guidelines (i.e., ≥60 min/day of MVPA) according to different cut-points

In conclusion, this study shows large discrepancies in the time spent in SB and PA intensities across cut-points relative to different body attachment sites and acceleration metrics in overweight or obese children. Furthermore, we provide a comprehensive comparison between available cut-points in order to better understand which cut-points provide comparable results and which ones not. Also, our data clearly showed that it is not currently possible to know the prevalence of a population meeting the PA guidelines based on accelerometer data, with differences from nearly none to nearly everyone meeting the guidelines. Although currently elusive, data harmonization and consensus are essential to comparatively measure and communicate objectively monitored time in SB and various PA intensities across different studies.

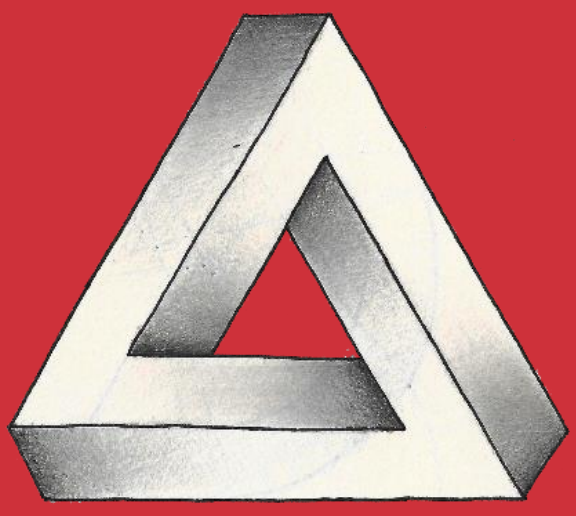
Perspectives

In the present study, we provide a comprehensive overview on the comparability of available cut-points for the classification of SB, LPA, MPA, VPA and MVPA from different accelerometer attachment sites and acceleration metrics in children. This overview allows researchers to know how comparable are their findings with other published studies, for example, it can be expected that SB derived

Abbreviations in this page:
LPA: light physical activity
PA: physical activity

from Hänggi et al. [108] and Romanzini et al. [69] cut-points is comparable, but large differences can also be expected for LPA classified using the same cut-points. The general belief that PA estimations from wrist-worn accelerometers provide higher values than those from hip-worn accelerometers is not supported by the current study. Other factors such as the acceleration metric used, and the cut-points themselves seem to have a higher influence in the final estimations than the accelerometer attachment site. Therefore, our results confirm previous studies and extend their findings to a different sample (children with overweight or obesity) and by using a complete set of published cut-points for this population. Data pooling and harmonization should be performed, as well as meta-analyses using data from cut-points validation studies to propose a consensual set of cut-points to be used in different settings/projects.

Study VI



Step-based metrics and overall physical activity in children with overweight or obesity: cross-sectional study

Migueles JH, Cadenas-Sanchez C, Aguiar EJ, Molina-Garcia P, Solis-Urra P, Mora-Gonzalez J, Garcia-Marmol E, Shiroma EJ, Labayen I, Chillon P, Löf M, Tudor-Locke C, Ortega FB.
JMIR Mhealth Uhealth. 2020 Apr;8(4):e14841
DOI: 10.2196/14841

Contents

Abstract

Background | Best-practice early interventions to increase PA in children with overweight and obesity should be both feasible and evidence based. Walking is a basic human movement pattern that is practical, cost-effective, and does not require complex movement skills. However, there is still a need to investigate how much walking—as a proportion of total PA level—is performed by children who are overweight and obese in order to determine its utility as a public health strategy.

Aims | This study aimed to (i) investigate the proportion of overall PA indicators that are explained by step-based metrics and (ii) study step accumulation patterns relative to achievement of public health recommendations in children with overweight and obesity.

Methods | A total of 105 children with overweight or obesity (mean 10.1 years of age [SD 1.1]; 43 girls) wore hip-worn accelerometers for 7 days. PA volumes were derived using the daily average of VACounts per 15 seconds, categorized using standard cut points for light-moderate-vigorous PA (LMVPA) and MVPA. Derived step-based metrics included volume (steps/day), time in cadence bands, and peak 1-minute, 30-minute, and 60-minute cadences.

Results | Steps per day explained 66%, 40%, and 74% of variance for counts per 15 seconds, LMVPA, and MVPA, respectively. The variance explained was increased up to 80%, 92%, and 77% by including specific cadence bands and peak cadences. Children meeting the WHO recommendation of 60 minutes per day of MVPA spent less time at zero cadence and more time in cadence bands representing sporadic movement to brisk walking (i.e., 20-119 steps/min) than their less-active peers.

Conclusion | Step-based metrics, including steps per day and various cadence-based metrics, seem to capture a large proportion of PA for children who are overweight and obese. Step-based metrics could be useful in discriminating between those children who do or do not achieve MVPA recommendations.

Abbreviations in this page:
LMVPA: light-moderate-vigorous physical activity
MVPA: moderate-to-vigorous physical activity
PA: physical activity
SD: standard deviation
VACounts: activity counts in the vertical axis
WHO: World Health Organization

Key Points

Question
Are step-based metrics enough to describe the overall PA in children with overweight or obesity?

Findings
Step-based metrics, including steps per day and various cadence-based metrics, seem to capture a large proportion of PA for children with overweight or obesity

Meaning
Step-based metrics could be useful in discriminating between those children who do or do not achieve MVPA recommendations.

Introduction

Abbreviations in this page:
PA: physical activity

Decreased PA is associated with increased risk of non-communicable diseases [277–279] and is responsible for approximately 9% of premature mortality [10]. Worldwide PA deficits [9,278,280,281], and inequalities between countries in PA levels [282] require effective counter-active strategies, especially in populations at risk such as people with overweight and obesity. For example, evidence [283,284] suggests that low levels of PA initiated in childhood and perpetuated in adulthood set up adults with overweight or obesity for an increased array of comorbidities during their life span [230]. Best practice early interventions should be both feasible and evidence-based. Walking is a basic human movement pattern that is practical, cost-effective and does not require complex movement skills. Thus, focusing on ambulatory activity could be the most accessible strategy to increase PA levels in children with overweight or obesity [285], who do not engage frequently in sports [286] and present poorer movement skills than normal weight children [287]. However, there is still a need to investigate how much ambulatory activity is performed by children with overweight or obesity (as proportion of total PA level) in order to determine its utility as a public health strategy. Information on what type of PA children with overweight or obesity are more likely to perform could help to plan more effective public health strategies, since intervening on a behaviour which is frequently occurring (e.g., walking) would have a greater impact than generating a new behaviour.

The ability to study health-related influences of PA has advanced in parallel with the increased use of accelerometer-based wearable technologies [41]. Accelerometers are capable of detecting human movement, but are primarily sensitive to ambulatory activity, the most common form of PA performed by adults [34,35]. However, children's movement patterns may be more variable and less is known about how predominant ambulatory activity, primarily walking, is relative to other types of PA behaviours. Time-stamped accelerometers are capable of detecting step-based metrics, including a tally of step accumulation over the day (i.e., volume [steps/day]), the time spent in incremental cadence bands (e.g., time spent walking at 80-99 steps/min) and/or peak 1-, 30- and 60-min cadence indices (i.e., average steps/min of the highest 1, 30 or 60 non-consecutive minutes in a day, respectively) [288–290]. Collectively, these metrics are referred to hereafter as step-based metrics.

Therefore, this study aimed to: 1) investigate the proportion of overall PA that is explained by ambulatory activity (i.e., step-based metrics) in children with overweight or obesity; and 2) study step-based patterns relative to PA guidelines achievement in children with overweight or obesity.

Abbreviations in this page:
BMI: body mass index
LPA: light physical activity
MPA: moderate physical activity
MVPA: moderate-to-vigorous physical activity
PA: physical activity
RCT: randomized controlled trial
SB: sedentary behaviour
SD: standard deviation
VPA: vigorous physical activity
WOF: World Obesity Federation

Methods

Study design and participants

The present cross-sectional analysis included data collected during the baseline assessment of the ActiveBrains project [291]. A detailed description of the study design, inclusion criteria and methods has been published elsewhere [63]. Briefly, ActiveBrains is a RCT intended to examine the effect of a 20-week PA intervention on brain structure, brain function, cognitive performance, academic achievement and physical and mental health outcomes in children with overweight or obesity [63]. A total of 110 children (8.5-11 years old) with overweight or obesity, defined according to the WOF cut points [64,65], were recruited from Granada, Spain. Data were collected from November 2014 to February 2016. Parents or legal guardians were informed of the purpose of the study and written informed parental consents were obtained. The ActiveBrains project was approved by the Human Research Ethics Committee of the University of Granada, and was registered as a clinical trial (NCT02295072).

Table 22
Anthropometry, SB, time-based PA metrics, and step-based metrics of children with overweight or obesity.

	All (n=105)	Boys (n=62)	Girls (n=43)	*P* sex
Age (years)	10.1 ± 1.1	10.2 ± 1.2	9.9 ± 1.1	0.24
Anthropometry				
Weight (kg)	56.6 ± 11.1	57.4 ± 11.1	55.4 ± 11.1	0.38
Height (cm)	144.4 ± 8.3	145.0 ± 7.8	143.6 ± 8.9	0.41
BMI (kg/m^2)	26.9 ± 3.7	27.1 ± 3.9	26.7 ± 3.5	0.53
Awake and wear time				
Awake time (min/day)	919.6 ± 31.5	921.3 ± 28.7	917.0 ± 35.2	0.51
Wear time during waking (min/day)	903.1 ± 35.3	905.1 ± 32.7	900.4 ± 39.0	0.52
SB and PA intensities				
SB (min/day)	600.8 ± 69.6	593.6 ± 69.1	611.1 ± 69.9	0.21
LPA (min/day)	273.2 ± 51.7	276.4 ± 51.2	268.5 ± 52.5	0.44
MPA (min/day)	34.0 ± 11.6	37.9 ± 12.4	28.2 ± 7.4	<0.001
VPA (min/day)	10.7 ± 6.7	12.3 ± 7.5	8.3 ± 4.4	<0.001
MVPA (min/day)	44.7 ± 16.8	50.3 ± 18.2	36.6 ± 10.3	<0.001
Step-based metrics				
Volume (steps/day)	8676.8 ± 2202.9	9257.6 ± 2431.9	7836.9 ± 1485.4	<0.001
Peak 60 min (steps/min)	63.7 ± 13.6	66.3 ± 14.4	59.8 ± 11.4	0.01
Peak 30 min (steps/min)	78.0 ± 14.5	79.7 ± 15.2	75.4 ± 13.2	0.12
Peak 1 min (steps/min)	111.5 ± 13.3	111.2 ± 13.3	111.8 ± 13.3	0.82
Time spent at different cadence bands				
Time at zero cadence (min/day)	346.6 ± 78.1	343.0 ± 79.5	351.7 ± 76.7	0.58
Incidental movement (min/day)	439.0 ± 63.4	434.9 ± 62.2	444.8 ± 65.4	0.44
Sporadic movement (min/day)	71.9 ± 18.2	73.5 ± 19.1	69.7 ± 16.8	0.28
Purposeful movement (min/day)	27.5 ± 9.2	30.3 ± 9.8	23.6 ± 6.6	<0.001
Slow walking (min/day)	15.9 ± 7.9	18.6 ± 8.5	12.0 ± 4.6	<0.001
Medium walking (min/day)	10.2 ± 6.2	11.8 ± 7.0	8.0 ± 3.8	<0.001
Brisk walking (min/day)	6.6 ± 6.0	7.4 ± 7.0	5.5 ± 4.0	0.08
Faster walking (min/day)	1.6 ± 2.6	1.8 ± 2.7	1.3 ± 2.3	0.33

Data are presented as mean ± SD.
Cadence bands represented are the following: incidental movement is 1-19 steps/min, sporadic movement is 20-39 steps/min, purposeful movement is 40-59 steps/min, slow walking is 60-79 steps/min, medium walking is 80-99 steps/min, brisk walking is 100-119 steps/min and faster walking is ≥ 120 steps/min.

Procedures

As part of the protocol of the ActiveBrains project [63], body weight and height were measured to the nearest 0.1 kg and 0.1 cm using an electronic scale (SECA 861, Hamburg, Germany) and a

precision stadiometer (SECA 225, Hamburg, Germany), respectively. BMI (kg/m^2) was calculated. Overweight and obesity were classified based on the cutoffs of the WOF [64].

Participants' overall PA and step-based metrics were measured with an accelerometer (ActiGraph GT3X+, Pensacola, FL, USA) worn on their right hip for 7 complete days (24 hours wear-time protocol). Participants were encouraged to wear the accelerometers as many hours as possible and only remove them for water activities (i.e., shower or swimming). Concurrently, participants logged in a diary for the 7 days the time they went to bed and woke-up. All participants with at least 4 days (including 1 weekend day) with ≥ 16 hours of accelerometer wear time were included in the analyses (n=105).

Data reduction

Raw ".gt3x" files (100 Hz) were loaded and processed with the ActiLife software to obtain activity counts (i.e., metric intended to capture body movement) accumulated in the VA over 15 second epochs and steps accumulation over 60 second epochs using the default filter developed by ActiGraph. Non-wear time was detected based on the raw acceleration values of the three axes using a previously published algorithm [67]. Briefly, each 15-min block was classified as non-wear time if the SD of two out of the three axes was lower than 13 mg during the surrounding 60-min moving window, or if the mean acceleration for two out of the three axes was lower than 50 mg. Likewise, sustained abnormally high accelerations (i.e., higher than 5.5 g; assumed to be related to device malfunction) were detected and labelled as non-wear time. The identified non-wear time (including sustained abnormally high accelerations) was imputed with the mean acceleration value for the corresponding time period over the remaining days of recording. Sleeping hours were identified using an automated algorithm guided by participants' logged times [40] and excluded from analyses. Non-wear time and sleeping hours identification were performed using functions included in the R package GGIR (https://cran.r-project.org/web/packages/GGIR/) [67].

Each 15 second epoch was classified into SB time or time at different PA intensities using the activity count cut-points developed by Evenson et al. [68]. Specifically, these were: SB (≤ 25 VACounts/15s), LPA (26-573 VACounts/15s), MPA (574-1002 VACounts/15s), and VPA (≥ 1003 VACounts/15s). Daily average acceleration (VACounts/15s), time spent at light-moderate-vigorous PA (LMVPA; > 25 VACounts/15s) and time at MVPA intensity (> 573 VACounts/15s) were included in the analyses as indicators of overall PA. Daily average acceleration (VACounts/15s) and MVPA are indicators commonly used to represent overall PA [95,292,293]. LMVPA was also included following the recommendations of the 2018

Abbreviations in this page:
BMI: body mass index
LMVPA: light-moderate-vigorous physical activity
LPA: light physical activity
MPA: moderate physical activity
MVPA: moderate-to-vigorous physical activity
PA: physical activity
SB: sedentary behaviour
SD: standard deviation
VACounts: activity counts in the vertical axis
VPA: vigorous physical activity
WOF: world obesity federation

Abbreviations in this page:
LMVPA: light-moderate-vigorous physical activity
LPA: light physical activity
MVPA: moderate-to-vigorous physical activity
PA: physical activity
SB: sedentary behaviour
VACounts: activity counts in the vertical axis

Physical Activity Guidelines Advisory Committee Scientific Report which acknowledge the importance of any kind of PA for health [51]. Furthermore, LPA could be a stimulus worthy to consider in children with overweight or obesity since they usually engage insufficient MVPA.

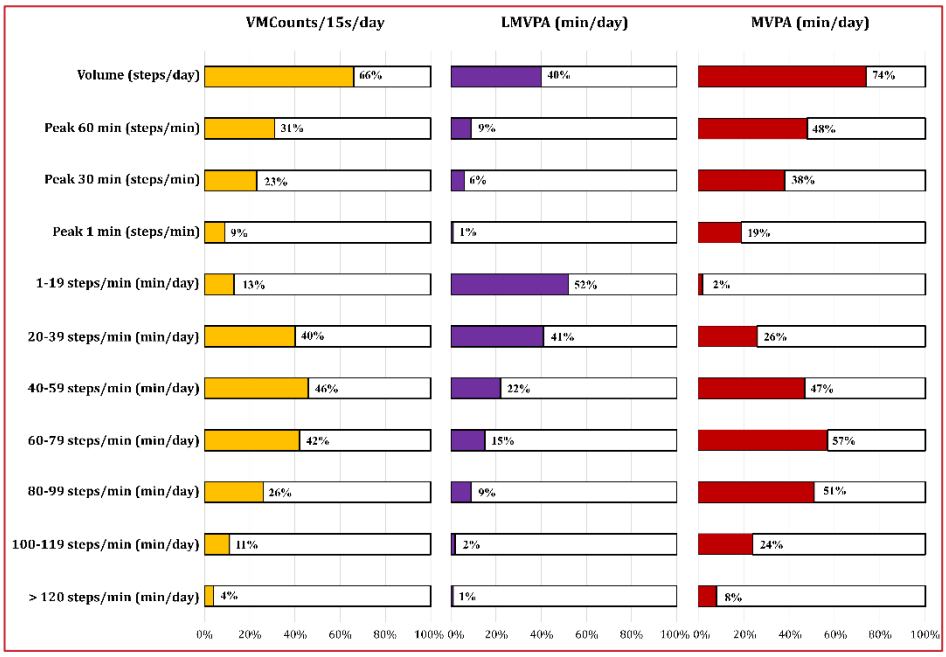


Figure 19
Proportion of variance (r^2) in overall PA indicators which is explained independently by each step-based metric.

Total ambulatory activity volume was derived as the number of steps/day. Furthermore, ambulatory activity cadence patterns were estimated as described previously for adults [294] and children [289]. Briefly, cadences were organized into bands of approximately 20 step/minute increments. These cadence bands has been previously associated with the next behavioural descriptors: incidental movement (1-19 steps/min), sporadic movement (20-39 steps/min), purposeful movement (40-59 steps/min), slow walking (60-79 steps/min), medium walking (80-99 steps/min), brisk walking (100-119 steps/min), and faster walking (≥ 120 steps/min). Time spent in each one of these bands, as well as time at zero cadence, were computed. In addition, the peak 60-min, peak 30-min and peak 1-min cadences were computed by rank ordering each participants' data for each day and then computing the average steps/min for the top 60, 30 and 1 min, respectively. The ActiGraph GT3X+ is valid for counting step [167,295] and its identified cadence bands have been used to describe cadence patterns in large cohorts [289]. Mean daily VACounts/15s, SB time, time-based and step-based metrics were then calculated as: (mean of available weekdays*5 + mean of available weekend days*2) / 7.

Data analyses

Descriptive characteristics of participants were presented as means and SDs. We used simple linear regression models to study the proportion of overall PA indicators explained by each step-based metric, and stepwise regression models to study the proportion

explained by using several step-based metrics as predictors. First, the variable which explained the highest proportion of the outcome variance was introduced. Then, those variables which significantly increased the proportion of variance explained were introduced. If any of the variables presented a variance inflation factor above seven, it was excluded from the model. Time at zero cadence was not included in these models since it represents inactivity. In addition, we identified those participants who achieved the WHO PA recommendations for this age-group [8], i.e., at least 60 min/day of MVPA. Two sample T-tests were then used to compare time spent in different cadence bands and the peak 60-min, the peak 30-min and the peak 1-min cadences of children who accomplished the PA recommendations compared with their peers who did not. All analyses were performed in R (https://cran.r-project.org/). The significance level was set at $P < 0.05$.

Abbreviations in this page:
LMVPA: light-moderate-vigorous physical activity
LPA: light physical activity
MPA: moderate physical activity
MVPA: moderate-to-vigorous physical activity
PA: physical activity
SB: sedentary behaviour
SD: standard deviation
VACounts: activity counts in the vertical axis
VPA: vigorous physical activity

Results

Descriptive characteristics

Table 22 presents anthropometric characteristics, SB, LPA, MPA and VPA, as well as step-based metrics for all participants stratified by sex.

ESM 6 (Table S1) shows the same descriptive characteristics stratified by weight status group (i.e., overweight, mild obesity, severe obesity and morbid obesity).

ESM 6 (scan or click here)

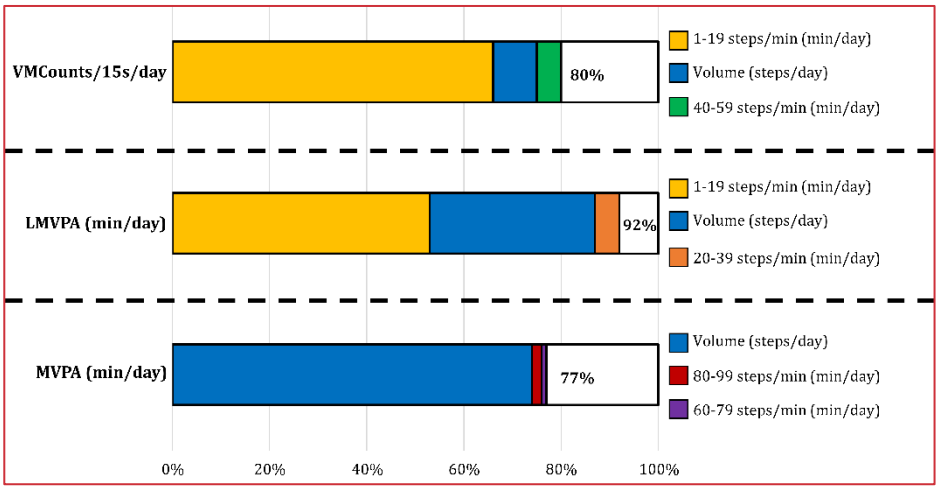


Figure 20
Proportion of variance (r^2) in overall physical activity indicators (i.e., VACounts/15s, LMVPA and MVPA), which is explained by combination of step-based metrics (calculated using stepwise linear regressions).
All predictors presented variance inflation factors < 6 in the selected models.

Proportion of total physical activity explained by step-based metrics

Figure 19 depicts the proportion of variance in indicators of overall PA (i.e., VACounts/15s, LMVPA and MVPA) that each step-based metric explained (r^2) in separate linear regression models (i.e., simple linear regression with each step-based metric as predictor and overall PA metric as outcome). Among the step-based metrics, steps/day explained the highest proportion of VACounts/15s and MVPA (i.e., 66% and 74%, respectively), while time at 1-19 steps/min explained the highest proportion of LMVPA (i.e., 52%). Overall, peak cadence indicators explained a lower proportion of the

Abbreviations in this page:
LMVPA: light-moderate-vigorous physical activity
MVPA: moderate-to-vigorous physical activity
PA: physical activity
SD: standard deviation
VACounts: activity counts in the vertical axis

variance than steps/day in overall PA indicators. Likewise, the shorter time intervals used to calculate the specific peak cadence indicator, the lower explanation capacity of the metric, which is to be expected given the shorter time frame represented (i.e., 60-min > 30-min > 1-min).

Cadence patterns according to physical activity guidelines

Out of the 105 children, 20 achieved the recommended amount of MVPA (i.e., ≥ 60 min/day). Children who performed < 60 min/day of MVPA also had significantly higher values for time at zero cadence (P = 0.004) and less time in cadence bands from 20 to 120 steps/min, compared with those who performed 60 min/day or more of MVPA (all P < 0.02) (**Figure 21**). Likewise, peak 60-min, peak 30-min and peak 1-min cadences were higher in children who achieved the 60 min/day of MVPA.

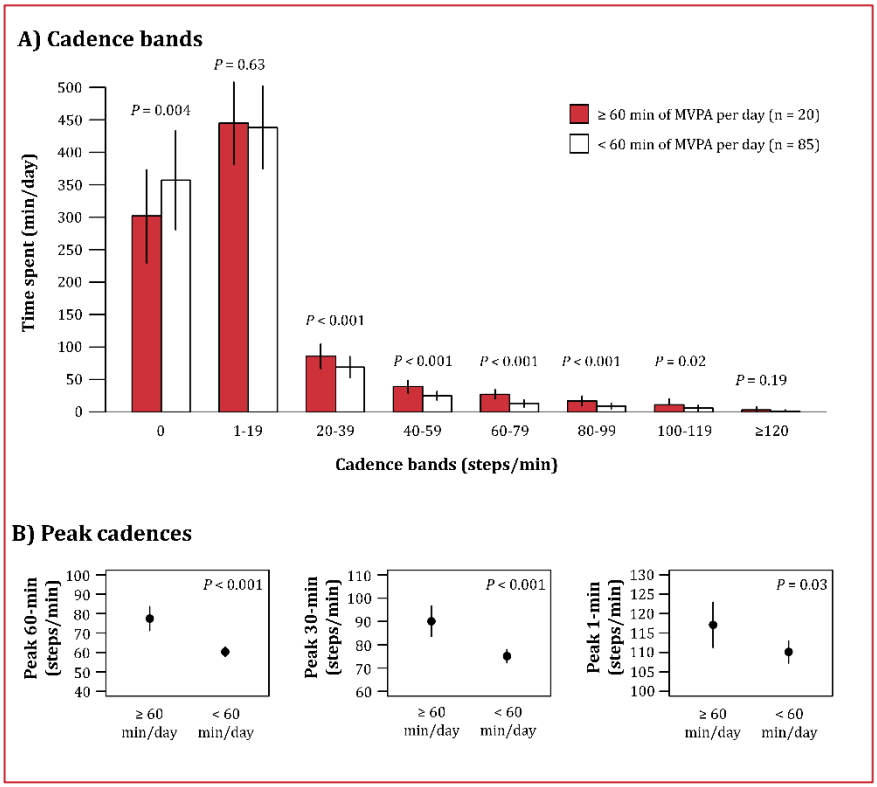


Figure 21
Time spent in each cadence band (panel A) and peak cadences (panel B) across children meeting or not the physical activity guidelines (i.e., at least 60 minutes of MVPA per day).
Error bars represent SDs.

Specifically, participants had to walk around 11000 steps/day to achieve the recommended dose of MVPA (**Figure 22**, panel A). Likewise, they had to spend 105 min/day walking at 20-39 steps/min, 40 min/day at 40-59 steps/min, 25 min/day at 60-79 steps/min, 19 min/day at 80-99 steps/min, 18 min/day at 100-120 steps/min or 10 min/day above 120 steps/min (**Figure 22**, panel B). Finally, their peak cadences had to be higher than 140 steps/min for peak 1-min, 100 steps/min for peak 30-min or 80 steps/min for peak 60-min (**Figure 22**, panel C).

Discussion

The main findings of this study are: 1) steps/day and the 1-19 steps/min cadence band explained the greatest amount of overall PA

(i.e., VACounts/15s [66%], LMVPA [52%] and MVPA [74%]) in children with overweight or obesity. The proportion of variance explained was further improved (by up to 77-92%) by adding other step-based metrics to the models; and, 2) cadence-based step patterns significantly differed between those children with overweight or obesity who achieved the PA recommendations compared with those who did not. Together, these findings seem to point out ambulatory activity as a major source of PA in children with overweight or obesity, as it has been previously reported in adults [34,35]. Further studies with larger and more representative samples should corroborate this finding. This finding can be leveraged to design appropriate PA interventions (i.e., by investigating the amount of walking at a certain intensity needed to meet PA recommendations) as a strategy to lower lifespan health risks in this vulnerable population.

Abbreviations in this page:
LMVPA: light-moderate-vigorous physical activity
MVPA: moderate-to-vigorous physical activity
PA: physical activity
VACounts: activity counts in the vertical axis

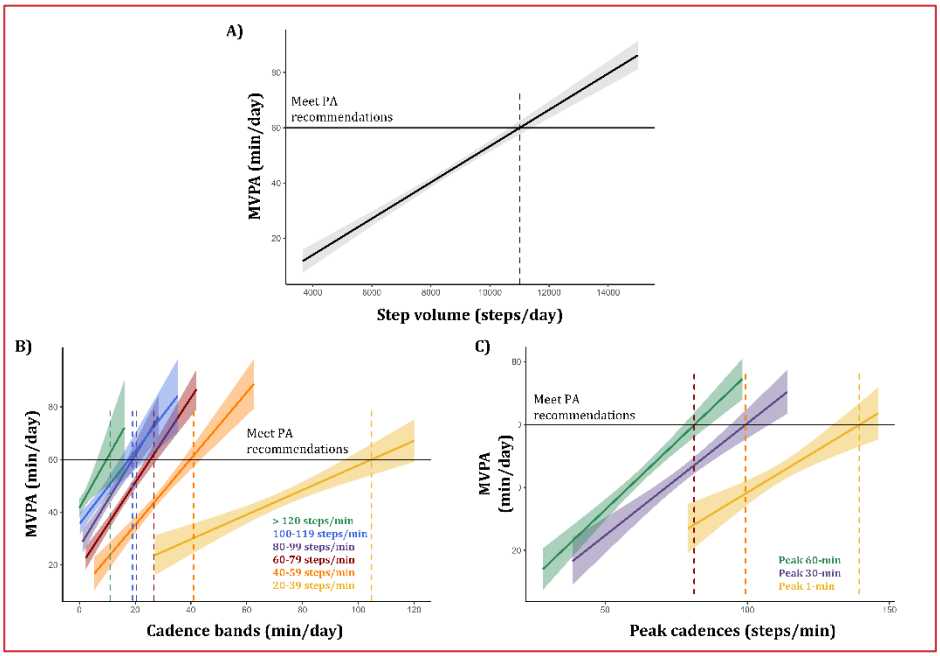


Figure 22
Linear regression lines (and 95% confidence intervals [shaded areas]) for the association between MVPA and step-based metrics, i.e., step volume (panel A), cadence bands (panel B) and peak cadences (panel C).

Steps/day and the 1-19 steps/min cadence band were the best explanatory factors of overall PA. Specifically, more than half of the variation in overall PA could be explained by either steps/day or the 1-19 steps/min cadence band in children with overweight or obesity, depending on the overall PA indicator regressed (i.e., 66% for VACounts/15s [steps/day], 52% for LMVPA [1-19 steps/min] and 74% for MVPA [steps/day]). Furthermore, all of the stepwise models included steps/day to estimate either VACounts/15s, LMVPA or MVPA. Accordingly, our sample was active for 5.3 hours/day (i.e., LMVPA), during which they spent around 2.2 hours/day in ambulatory activity (i.e., from sporadic movement to faster cadences). This presumes around 40% of the time spent in LMVPA, which is similar to the estimation obtained to predict LMVPA from steps/day (i.e., 40%).

However, steps/day was not the only important factor in the prediction of overall PA indicators. Information regarding step accumulation pattern increased the prediction capacity up to 80%, 92%

Abbreviations in this page:
LMVPA: light-moderate-vigorous physical activity
MPA: moderate physical activity
MVPA: moderate-to-vigorous physical activity
PA: physical activity
VACounts: activity counts in the vertical axis

and 77%, for VACounts/15s, LMVPA and MVPA, respectively. These findings support the concept of also considering stepping rate, which has been associated with health-related intensity levels in children [290] in addition to steps/day. The fact that LMVPA explanation increased substantially (i.e., from 40% to 92%) by including more metrics in the stepwise models is noteworthy. This suggests that considering both steps/day and certain cadence bands we can explain around 90% of their active minutes. However, MVPA explanation only increased from 74% to 77%, which means that almost all information on MVPA is already provided by steps/day. Together, it seems clear that step-based metrics are more powerful to explain light-intensity PA than higher intensities.

To this end, Tudor-Locke et al. found that walking at 115 steps/min requires an energy expenditure of approximately 4 METs (i.e., MPA intensity for children) in 9-11-year-old children, measured while walking on a treadmill [296]. However, the cadence-intensity relationship observed under laboratory-controlled conditions may not be generalizable to free-living data from children with overweight or obesity. Likewise, caution is advised since measurement tools differed between studies (i.e., direct observation vs. accelerometers). We observed around 34 min/day classified as MPA intensity by Evenson et al. cut-points [68], and, in turn, around only 2 ± 3 min/day accumulated at a cadence > 115 steps/min, which is indicative of MVPA intensity in this age group as measured in lab conditions [296]. A source for this difference could be the epoch length used to derive MPA [233], i.e., 15 seconds for Evenson et al. cut-points and 60 s for time spent above 100 steps/min. Estimations based on Evenson cut-points could be able to capture short bouts of MPA up to 15 seconds, while step-based estimations of MPA are limited to those bouts lasting at least 1 minute. We decided to use 60 s epochs for cadence to maintain consistency with previous studies, to ease comparability of findings and because there are no studies examining the cadence measured in 15 s epochs and intensity to date (making more difficult to interpret the findings). It could be also argued that most of the MPA performed by our sample was not related to ambulatory activity, which seems unlikely because step-based metrics explained almost 80% of the variance in MVPA. We must also acknowledge that metabolic intensity is indirectly inferred from detected movement signals and is not a clear indicator of metabolic cost, so there are likely to be measurement differences attributable to differential definitions. Therefore, further research is needed to understand how free-living cadence bands relate to energy expenditure and accelerometer signals.

Peak cadence indices and cadence bands have been previously used as proxy indicators for ambulatory activity intensity and

pattern, respectively, in children [289] and adults [294,297]. In congruence with Barreira et al. [289], we found that most of the day was spent in low intensity or SB. Specifically, we found around 10 hours/day of SB time using Evenson et al. [68] cut-points, or, in regards to step-based metrics, a time at zero cadence value of 5.8 hours/day and 7.3 hours/day in incidental movements (1-19 steps/min). Barreira et al. reported similar step accumulation patterns in 6-11-year-old children from the US NHANES 2005-2006. Notably, only 38% of the NHANES population-based sample had overweight or obesity [289]. In contrast, our sample of children who had overweight or obesity accumulated more time at zero cadence and also time in incidental movements (1-19 steps/min), and less time in cadence bands from sporadic movement to faster walking (20-120 steps/min). Likewise, differences in accelerometer models, study design and socio-environmental context should be considered when comparing these studies.

According to the Evenson et al. cut-points definition of MVPA [68], 20 children (i.e., 19%) from our sample met the PA recommendation of at least 60 min/day of MVPA [8]. This finding should be interpreted with caution since quantification of time-based PA with accelerometers is notoriously challenging and is dependent on a variety of data collection and processing decisions [233], including those related to selecting appropriate analytical cut-points [266]. We have previously reported that changing cut points can derive extremely different estimations of the proportion of children meeting PA recommendations in this sample [266]. It is also important to consider that PA recommendations are mainly based on self-reported data, which could bias interpretation over objective data. Additionally, when compared to normative values from NHANES 2005-2006 [298], our sample can be considered "below average" for steps/day for 8-9 years old children (i.e., 7,647-9,398 steps/day) or "average" for 10-11 year-olds (i.e., 8,504-10,066 steps/day). Likewise, our sample presented "below average" values for peak 60-min cadence (i.e., 62-71 steps/min). Nevertheless, a large proportion of the count-based MVPA performed was related to step-based metrics, which suggests that limited ambulatory behaviours could be responsible of the low prevalence of children meeting the PA recommendations. Furthermore, we found significant differences in ambulatory activity intensity (i.e., time spent in almost every band cadence was significantly different) between those who met and did not meet the PA recommendations. Specifically, children who met the recommendations spent around 55 min/day less in time at zero cadence, 17 min/day more in sporadic movement (20-39 steps/min), 14 min/day more in purposeful movement (40-59 steps/min), 13 min/day more in slow walking (60-79 steps/min), 8 min/day more

Abbreviations in this page:
MVPA: moderate-to-vigorous physical activity
NHANES: National Health And Nutrition Examination Survey
PA: physical activity
SB: sedentary behaviour

Abbreviations in this page:
MVPA: moderate-to-vigorous physical activity
PA: physical activity
SB: sedentary behaviour

in medium walking (80-99 steps/min) and 5 min/day more in brisk walking (100-119 steps/min). Additionally, peak 1-min, peak 30-min and peak 60-min cadences seem to be able to discern between children achieving or not the recommended dose of MVPA per day (i.e., 60 min/day).

Findings of this study have several practical implications to consider. For example: 1) as a large proportion of overall PA identified by accelerometers is explained by step-based metrics in children with overweight or obesity, these measures could be used to describe and compare PA patterns in this population; 2) it could be assumed that increasing ambulatory activity volume and intensity is a feasible form of PA which can increase the chances of meeting PA recommendations in this population. This is especially important to consider as ambulatory activity is a feasible PA strategy that may lead to several health benefits, for example, improved body composition and mental health, in children with overweight or obesity [299–301]. Notably, walking does not require complex movement skills and so, can be performed with most populations, including children with overweight or obesity who frequently do not engage in sports because of their low physical competence [287].

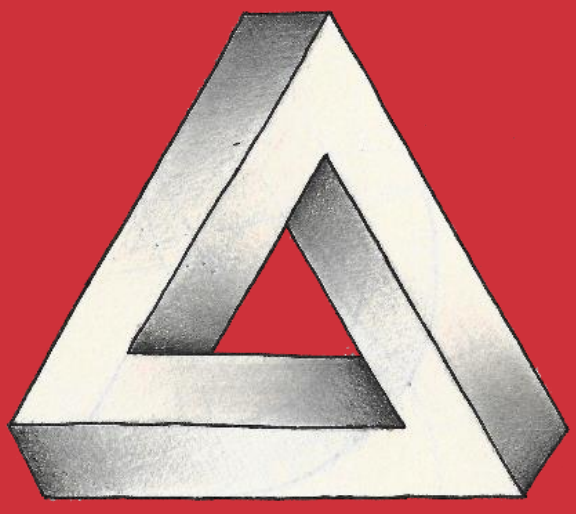
Several limitations should be acknowledged. First and foremost, accelerometer measurements of PA are influenced by a variety of data collection and processing decisions [233]. This means that it cannot be considered a gold-standard for overall PA measurement and that changes in the quantification of PA could change the findings observed in the present study. However, we were as consistent as possible regarding the measurement of both overall PA and step-based metrics. Both outcomes come from the same hip-worn accelerometer and cut-points used are based on the VA acceleration, which is consistent with the ActiGraph procedures to detect steps. This would reduce the methodological inconsistencies between the overall PA and the step-based metrics estimations, which, in turn, can be considered as a strength of the present study. Note that epoch length discrepancies between count-based and step-based metrics may be partially responsible of the differences observed. However, our findings should be interpreted with caution since overall PA refers to accelerometer-determined PA (which is not a gold-standard and could ignore certain activities such as swimming). Note that step-based metrics derived from pedometers could vary the findings from this study and their relationship with overall PA should be investigated. Likewise, another strength to highlight is that we are focusing on a population who may benefit greatly from increases in ambulatory activity; for example, this study demonstrates that they could have substantially increased chances of meeting PA recommendations by only focusing on ambulatory activity.

Conclusion

In conclusion, step-based metrics including steps/day and various cadence-based intensity indicators seem to capture the majority of PA in children with overweight or obesity. Step-based metrics could be useful for discriminating between those children who do or do not achieve MVPA recommendations. Further studies should test whether step-based metrics derived from pedometers are similarly useful for that purpose.

Abbreviations in this page:
MVPA: moderate-to-vigorous physical activity
PA: physical activity

Study VII



The GRANADA consensus on analytical approaches to assess associations with accelerometer-determined physical behaviours in epidemiological studies

Migueles JH, Aadland EC, Andersen LB, Brønd JC, Chastin SF, Hansen BH, Konstabel K, Kvalheim OM, McGregor DE, Rowlands AV, Sabia S, van Hees VT, Walmsley R, Ortega FB.
Submitted to Br J Sports Med

Contents

Abstract

Background | Research on the interrelationships between PA, SB and sleep (collectively defined as physical behaviours) has given birth to the field of 'physical behaviour epidemiology'. Each of these physical behaviours has been associated with health in epidemiological studies, but their co-dependency and interactions in relation with health need to be further explored and accounted for in data analysis. Modern accelerometers present the major advantage of capturing continuous movement through the day; which, in turn, comes with the intrinsic challenge of how to best utilize the richness of the data collected. In the past few years, statistical procedures first applied in other scientific fields have been applied to physical behaviour epidemiology. A comprehensive description, discussion, and consensus regarding the strengths and limitations of analytical approaches used in the field of 'physical behaviour epidemiology' will help researchers decide which approach to use in different situations.

Aims | Thus, we aimed to provide a comprehensive description and discussion on: 1) the analytical approaches (i.e., from generation to statistical modelling of descriptors) currently used in the scientific literature on physical behaviour, highlighting their strengths and limitations and providing practical recommendations on their use; and 2) current gaps and future research directions around the analysis and use of accelerometer data.

Methods | In this context, a scientific Workshop entitled "International Workshop: A focus on statistical methods to analyse accelerometer-measured physical activity" followed by an Expert Meeting, were held in Granada on October 21st and 22nd 2019, respectively. These events brought together an international panel of researchers with expertise on the above topics.

Impact | Advances in analytical approaches to accelerometer-determined physical behaviours in epidemiological studies are expected to influence the interpretation of current and future evidence and ultimately impact on future physical behaviour guidelines.

Abbreviations in this page:
PA: physical activity
SB: sedentary behaviour

Key Points

Question
What accelerometer data descriptors and statistical model should be used to the study of physical behaviours and health associations?

Findings
A comprehensive description of the accelerometer data descriptors and statistical procedures most-frequently used in the field is provided

Meaning
Researchers can follow the decision tree diagram included in this article for some guidance in the decision of the analytical approach

Introduction

Abbreviations in this page:
PA: physical activity
SB: sedentary behaviour

Interrelationships between PA, SB and sleep, collectively described as physical behaviours [14], have drawn the attention of many researchers worldwide, unifying previously separated research fields into 'physical behaviour epidemiology'. Traditionally, self-report methods have been used for physical behaviour research and subsequent guidelines. Accelerometers (movement sensors) are increasingly being used in physical behaviour research, but no consensus exists in the literature on how to use the time series produced by these sensors when examining associations (either cross-sectional or longitudinal) between physical behaviour and health outcomes. The data-analytical approach to link physical behaviour exposure with health outcomes usually includes these steps: 1) the reduction of the acceleration time series into meaningful descriptors of physical behaviours; 2) the adequate mathematical treatment of these descriptors if needed; and 3) the selection of a statistical model to investigate the epidemiological associations of physical behaviours with health.

In this context, a meeting titled "International Workshop: A focus on statistical methods to analyse accelerometer-measured physical activity" was held in Granada on October 21st-22nd 2019. This event brought together an international panel of researchers with expertise in the analysis of data collected with accelerometers, in order to have a comprehensive and constructive discussion about the strengths and limitations of currently used analytical approaches. The present article aims to provide a comprehensive description and discussion on: 1) the most-frequently used analytical approaches (i.e., from descriptors to statistical modelling) currently used in the field, highlighting their strengths and limitations and providing practical recommendations on their use; and 2) current gaps and future research directions around the analysis and use of accelerometer data in physical behaviour epidemiology. This article focuses on modelling physical behaviours as exposure variables and health indicators as outcomes. Our focus is on PA, although we also consider SB and sleep given their interrelationships with PA. Data collection protocol decisions are outside the focus of this document. However, it is important to note that decisions on the accelerometer body attachment site [233,302], number of days recorded [303], treatment of weekdays and weekend days [15,304], seasonality [305], among others, affect the validity of accelerometer data as a measure of a person's typical physical behaviours. For example, regarding the body attachment site, hip-worn accelerometers have typically been more accurate than wrist-worn accelerometers at classifying PA intensity [68], although new algorithms have reached good validity with wrist-worn accelerometer data [61]; sleep assessment seems to be more

Abbreviations in this page:
ENMO: Euclidean Norm Minus 1*G*
MAD: mean amplitude deviation
MIMS: Monitor Independent Motion Summary
PA: physical activity
SB: sedentary behaviour

feasible from wrist-worn accelerometer data [38,40]; and recently, methods to classify PA types have been greatly improved using thigh-worn accelerometers [36]. A recent consensus report discussed best practices on these decisions [306].

Overview of descriptors and statistical procedures currently used

Accelerometer data descriptors

Modern accelerometers collect raw accelerations (measured in *G*'s) at sample frequencies typically varying from 30 to 100 Hz. As an example, raw data from a thigh-worn accelerometer is presented in **Figure 23**. This raw signal is usually filtered and aggregated to remove the gravitational acceleration and the noise effects on the signal [67]. Examples of common accelerometer data aggregation metrics are activity counts (brand-specific and proprietary aggregation metrics), ENMO, MAD, Monitor Independent Motion Summary (MIMS) units, or Activity Index, among others (hereinafter we refer collectively to them as 'acceleration metrics'). With regard to MIMS it should be noted that the claim that it is accelerometer brand independent has so far not been demonstrated, only sensor from the Actigraph brand were used in the study by Dinesh and colleagues [307]. Further, no evidence exists that other metrics like MAD and AI_0 are not brand independent. MIMS applies a narrow frequency filter by which its potential lack of sensitivity to differences in the monitor comes at the cost of lower sensitivity to movements in the low- and high frequency range.

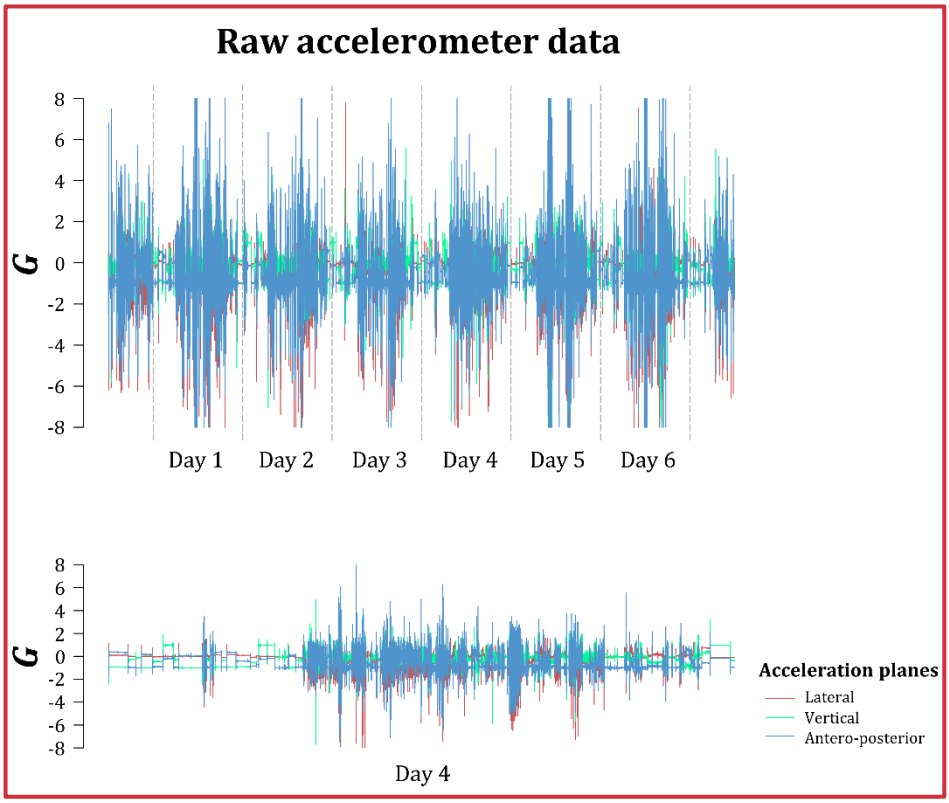


Figure 23
Sample raw accelerometer data recording from a thigh-worn accelerometer. Accelerometer model: Axivity AX3, sampling frequency: 30 Hz, body attachment site: thigh; 24h/day recording protocol.

In-depth discussions about the influence that these aggregation metrics on the final estimates have been published elsewhere [67,153,233,302]; we focus our discussion on the conversion of such acceleration metrics to descriptors at a day or person level. **Table 23** presents a brief description and example references on the descriptors included in this section. Given the numerous versions of accelerometer data descriptors presented in the literature, we decided to focus on those descriptors representative of PA volume, type, and intensity since they are the most frequently-used in public health guidelines.

Abbreviations in this page:
LPA: light physical activity
MPA: moderate physical activity
PA: physical activity
PAEE: physical activity-related energy expenditure
SB: sedentary behaviour
VPA: vigorous physical activity

Table 23
Description of accelerometer-based descriptors of physical behaviours

Descriptor	Brief description	Examples
Average acceleration	Arithmetic average of the processed acceleration throughout the measurement period or per day.	[274,292,308]
Time-use behaviours	Estimates of time spent in physical activity intensities (e.g., LPA, MPA, VPA), types (e.g., walking, running, cycling), or SB, optionally expressed in bouted and un-bouted behaviour. These estimates can be derived with heuristic methods or machine learning.	[266,274,309]
Intensity spectrum	The intensity spectrum is an extension of cut-points which attempts to provide a much more detailed description of the physical activity intensity pattern. Instead of using cut-points representative of SB, LPA, MPA or VPA, the cut-points are arbitrarily selected to obtain a wider range of intensity bins.	[310,311]
Intensity gradient	The intensity gradient describes the negative curvilinear relationship between physical activity intensity and the time accumulated at that intensity during the 24 h day.	[292,308]
MX metrics	The acceleration above which a person's most active X minutes/time (MX) are accumulated, to focus on a person's most active periods of the day.	[312]
Acceleration functions	Description of the accelerometer data with a function rather than with a scalar. Functions seek a more detailed description of the accelerometer data without making a priori assumptions.	[313–315]

Average acceleration

Average acceleration over a 24 h period is directly derived from the processed acceleration and can be used as a proxy for total daily PAEE [316]. This single estimate indicates the overall activity level and/or the volume of activity. It is usually expressed in m*g* or a manufacturer-provided acceleration metric (usually counts). Average acceleration usually has a moderate correlation with PAEE ($r \sim 0.3$-0.5), which can be improved by considering body weight, body composition, and activity type in the models [200,237]. Given that the correlation is not high, it is often used as a direct measure of movement, without making inferences about PAEE.

Time-use descriptors

Various descriptors quantify the daily time spent in a set of behaviours, e.g., time spent in certain activity intensities (e.g., LPA, MPA, or VPA) or types (e.g., sitting, standing, walking). In this regard, cut-points represented one of the first developed and most frequently used methods for assessing the time spent in SB and in LPA,

Abbreviations in this page:
LPA: light physical activity
MET: metabolic equivalent
MPA: moderate physical activity
PA: physical activity
PAEE: physical activity-related energy expenditure
ROC: receiver operating characteristic
SB: sedentary behaviour
VPA: vigorous physical activity

MPA, and VPA using the acceleration metric [37]. The identified linear association between acceleration and energy expenditure was used to determine cut-points based on standard METs thresholds (e.g., SB, ≤1.5; LPA, >1.5 and <3.0; MPA, ≥3.0 and <6.0; VPA, ≥6.0 [317]). **Figure 24** graphically represents a cut-point-based classification of the acceleration recorded during one day without any definition of bouts. Cut-points can be derived with standard statistical procedures such as linear regression or ROC curves, which assume a linear relationship between magnitude of acceleration and METs. However, non-linear approaches have also been used. Otherwise, classification of activity types usually relies on thresholds applied to the device angle variability, usually from thigh- or wrist-placed accelerometers [36,40]. Similarly, thresholds have been applied to acceleration metrics and accelerometer angles to detect sleep from the accelerometer signal [38–40]. More sophisticated models have used the acceleration signal to detect whether the activity performed is locomotion or not, and then applied specific regression models for each activity type (locomotion vs. not locomotion) [158]. Machine learning methods have gained momentum to classify both activity intensities and types from an accelerometer time series [318]. Classifying behaviours or estimating energy expenditure using a supervised machine learning approach requires data labelled with 'true' intensity or type (as measured with indirect calorimetry or heart rate monitors, among others) [115,118,127,221,319], which is used to iteratively improve classification/estimation. Alternatively, unsupervised machine learning methods can be used to define "states" in the accelerometer signal pattern that can be interpreted as specific behaviours [320].

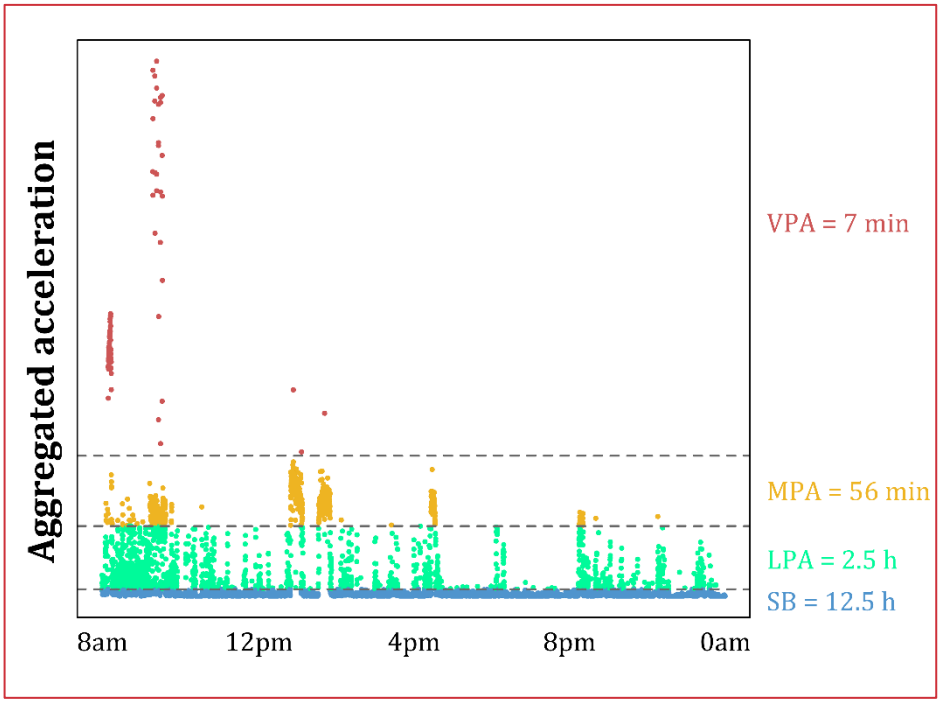


Figure 24
Graphical representation of cut-point-based metrics without bout-specification. Accelerometer model: ActiGraph GT3X+, sampling frequency: 100 Hz, body attachment site: hip; only awake time represented.

Independently of the method used, these descriptors estimate daily time devoted to a specific behaviour. Descriptors of time spent in different PA intensities were first developed to assess objectively the information gained from questionnaire data (the source of most knowledge on the benefits of PA). Use of these time estimates in recent research has confirmed the benefits of PA for health and demonstrated stronger effects of PA than observed with self-report [321].

Abbreviations in this page:
PA: physical activity
PAEE: physical activity-related energy expenditure

Intensity spectrum

The intensity spectrum is also quantified as daily time spent in certain categories, so it is a time-use descriptor. Specifically, time acceleration metric signal over time is classified based on increasing acceleration bins (e.g., time spent from 0-50, 50-100, 100-150, [...] counts or m*g*). Thus, the intensity spectrum uses a wider range of narrower equally-sized bins for increased resolution of the data [311]. The definition of the bin size is arbitrary, might not directly relate to PAEE and does not make any assumption on the behaviour underlying the intensity bin (its purpose is purely descriptive). It can also be regarded as a discretisation of a functional representation of the intensity distribution. The idea behind this approach is to avoid exaggerated aggregation of data (into only 3-4 categories) leading to loss of information. Thus, the number of bins should be large enough to incorporate all essential features in the accelerometer signal.

Intensity gradient

The intensity gradient describes the negative curvilinear shape of the intensity spectrum (i.e., the higher the intensity the less time spent at this intensity) [308]. The regression coefficient from a linear regression of time spent in an intensity bin on intensity, both on a logarithmic scale, is used as a scalar descriptor of this curvilinear relationship. It is always negative, reflecting the drop-in time accumulated as intensity increases; a more negative (lower) gradient reflects a steeper drop with a large proportion of time accumulated at lower intensities, while a less negative (higher) gradient reflects a shallower drop with time accumulated at higher intensities (**Figure 25**).

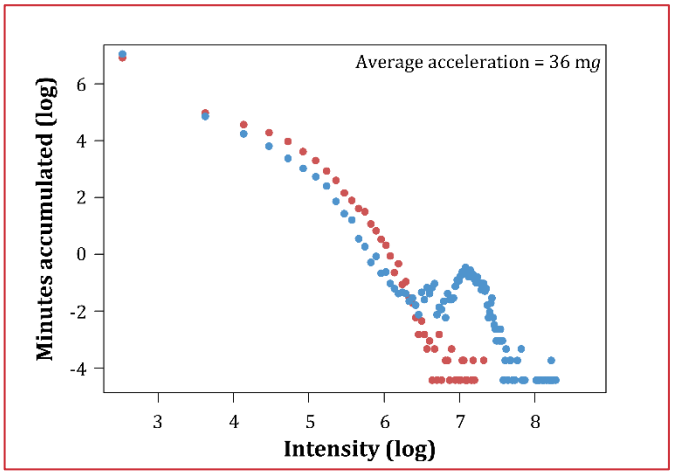


Figure 25
Example of intensity gradients from different participants with a similar average acceleration but discordant intensity distribution (i.e., intensity gradient). Accelerometer model: ActiGraph GT3X+, sampling frequency: 100 Hz, body attachment site: non-dominant wrist.

Abbreviations in this page:
IG: intensity gradient
MX: person's most active X min
MVPA: moderate-to-vigorous physical activity
PA: physical activity
PAEE: physical activity-related energy expenditure
VPA: vigorous PA

MX metrics - acceleration values corresponding to a set of percentiles

Time-use descriptors were based on the time accumulated in a series of a priori defined behaviours/bins. An alternative is to turn this approach on its head and describe the acceleration intensity distribution in terms of standardized periods of time or fractions of the 24 h day (percentiles). The acceleration for each epoch during the day is ranked in descending order to obtain the acceleration above which the person's most active X (MX) minutes are accumulated [308]. Therefore, instead of reporting the minutes above a given acceleration threshold, the minimum acceleration achieved for a given duration is reported (the unit of measurement is often m*g* or counts). MX, where X refers to the duration, e.g. M30, refers to the minimum acceleration for the most active 30 min (~percentile 98th) of the day. The active minutes may be accumulated in any way across the day. For example, if a child had an M60 value of 230 m*g*, the child accumulated 60 min of PA at accelerations (intensity) greater than 230 m*g* across the day. Similarly, the periods with the lowest recorded activity can be defined.

A range of MX metrics covering short to long time durations can be used to aid interpretation of the volume and intensity of the 24 h profile of physical activity. Using the MX metrics facilitates interpretation in terms of time spent in indicative activities (e.g., brisk walking) or above cut-points for different intensities of activity, e.g., MVPA or VPA. Plotting a broad range of MX variables on a radar plot illustrates the intensity and volume of the 24h activity profile (**Figure 26**), facilitating e.g., translation of results from analyses investigating the relative contributions of average acceleration and intensity gradient to markers of health, and/or comparisons between and within groups. For example, the M120, M60, M45, M30, M15, M10, M5 and M2 illustrate the more active periods of the day, while M8h refers to the most active 8 h of the day.

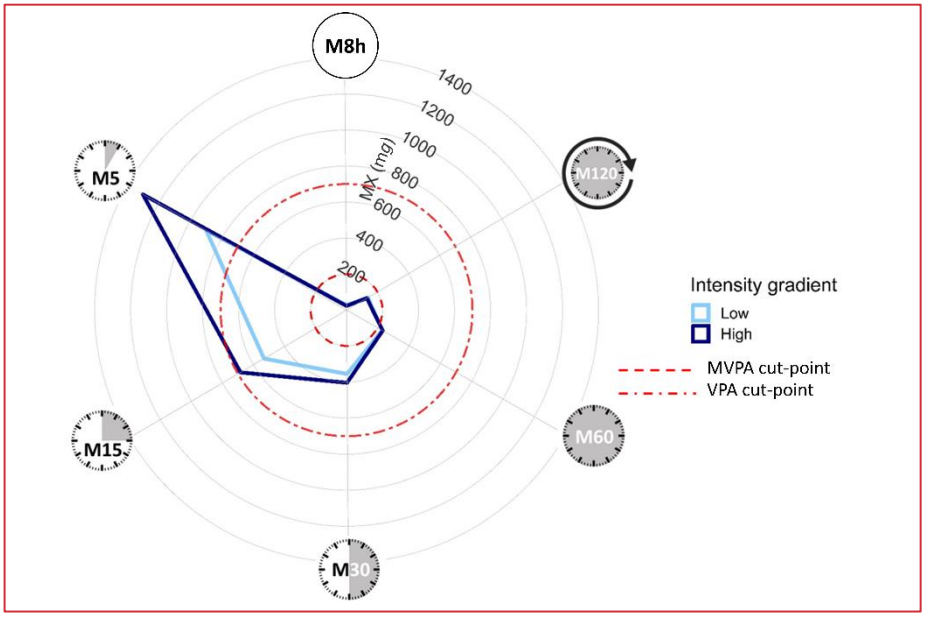


Figure 26
MX metrics example from two participants with similar average acceleration but different intensity gradient. Accelerometer model: ActiGraph GT9X, sampling frequency: 100 Hz, body attachment site: non-dominant wrist. Adapted from Rowlands et al. [312] with the permission from the publisher.

Acceleration functions

While the abovementioned descriptors are represented by scalar numbers, acceleration can also be described using a function. For example, the intensity gradient (described above) can be defined by its function instead of only reporting the beta coefficient. Other functions of interest could be the acceleration over time of the day [322] or the acceleration distribution (**Figure 27**) [315]. Acceleration functions seek a more detailed description of behaviours without making a priori assumptions. For example, while time in light activities assumes that all of the data between two cut-points (e.g. 0.05 to 0.10 *g*) relates similarly to health outcomes, analysis of acceleration functions could detect that a group tend to do more activities at acceleration less than 0.05m*g* or more activities at acceleration above 0.07*g*.

Abbreviations in this page:
IS: interdaily stability
IV: intradaily variability
PA: physical activity
SB: sedentary behaviour

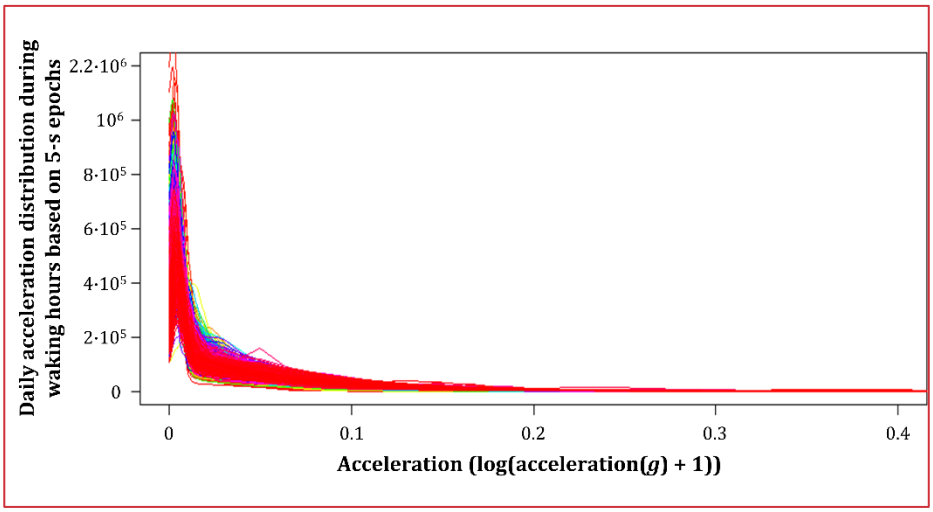


Figure 27
Sample of accelerometer-based distribution as a function of acceleration and time. Accelerometer model: GeneActiv, sampling frequency: 85.7 Hz, body attachment site: non-dominant wrist; 24h/day recording protocol

Indicators of movement behaviour patterns and quality

All the above-mentioned descriptors are time-based (time-use behaviours and intensity spectrum) or acceleration-based (average acceleration, MX metrics, acceleration functions) descriptors. That is, they either measure time in a given behaviour or acceleration in a certain time interval. Other descriptors of movement behaviour quality and patterns can be obtained thanks to the time-stamped data derived from accelerometers. Time-stamped accelerometer data can be used to derive certain characteristics of the PA and SB patterns throughout the day, such as the time accumulation in bouts of PA intensities or types. Time-stamped data also provides insight on timing of behaviours, domain (school/work or leisure), and circadian rhythmicity. For example, fragmentation of PA and sleep, sedentary breaks, intradaily variability (IV), interdaily stability (IS), sleep efficiency, or waking periods after sleep onset are frequently used in the field to assess the quality and patterns of PA, SB, and sleep.

Mathematical treatment

This section focuses on mathematical treatments to account for the specific singularities of the descriptors presented above. Time-

use behaviours and the intensity spectrum consist of a set of components that represent parts of some finite total. This total may be explicit (e.g., complete 24-hour data) or it may arise through interpretation of the data as proportions (e.g., waking day data). Therefore, these descriptors can be considered as compositional data. Each part is called a component and the proportional distribution is called composition. So, for a composition with *i* components:

$$\sum_i Component_i = 1 = 100\% = Whole$$

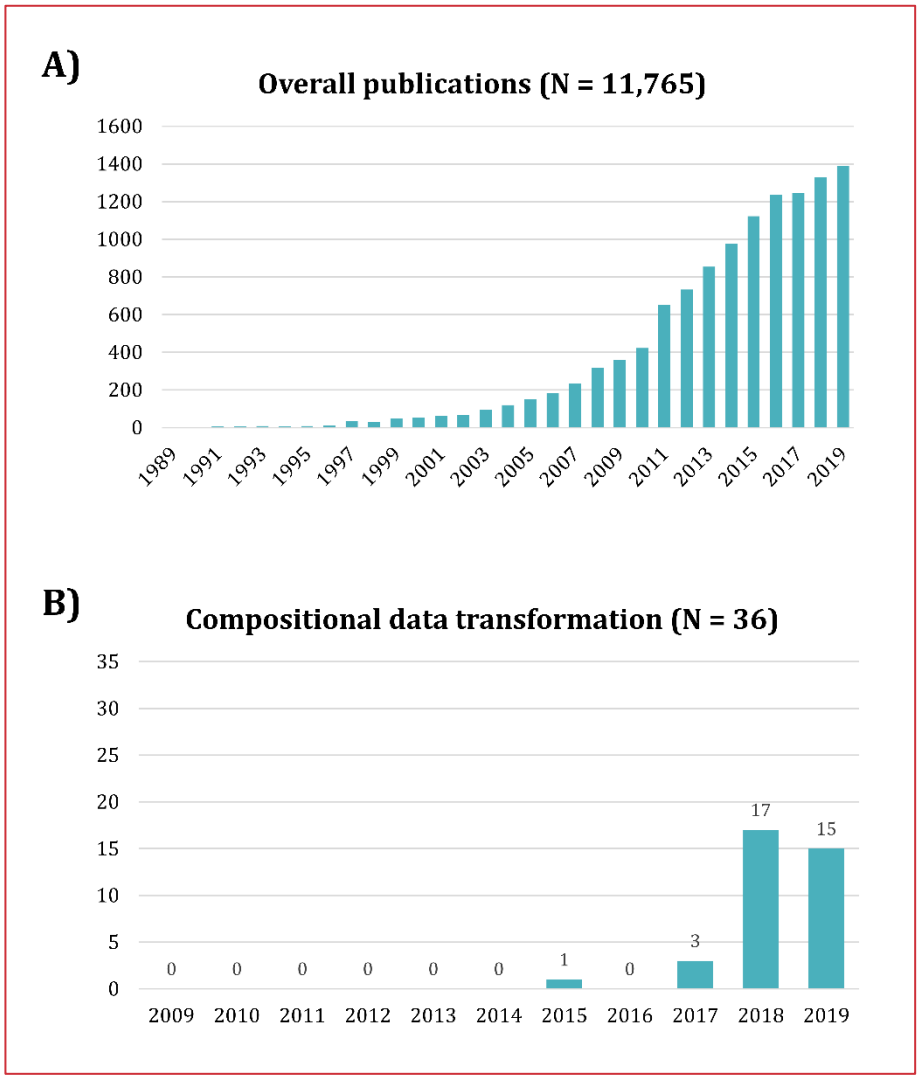


Figure 28
Overall number of publications using accelerometer-determined PA (panel A) and number of publications using compositional data transformations from inception to December 31st, 2019.
Search syntax introduced in the Web of Science: Panel A: (((((("physical activity")) OR "sedentary")) AND ((acceleromet*) OR actigraph*)); Panel B: (((((("physical activity")) OR "sedentary")) AND ((acceleromet*) OR actigraph*) AND ("compositional data analysis")).

Compositional data analysis is an approach to analyse compositional data. Its birth is often attributed to Pearson's paper on spurious forms of correlation in ratio data [323]. Arguably the father of compositional data analysis is John Aitchison, who developed comprehensive statistical frameworks to deal with compositional data [324]. Compositional data analysis is an established branch of statistics and has been used in many fields of research such as geosciences, nutrition, the study of the microbiome and gene sequencing. In the last five years, it has been applied in the field of 'physical behaviour epidemiology' to study the association between daily time use and health (**Figure 28**) [325–327].

Compositional data transformation

Time-use descriptors of physical behaviours are by nature compositional when they describe a time or energy budget (**Figure 29**). Hence the sum of time spent in each behaviour will be the period of interest (24 hours, waking period, week, wear time) and the proportions will sum to 100% of this period. In this example, the composition is made of four components over 24 hours: sleep, SB, LPA and MVPA.

Abbreviations in this page:
LPA: light physical activity
MVPA: moderate-to-vigorous physical activity
SB: sedentary behaviour

$$t_{sleep} + t_{SB} + t_{LPA} + t_{MVPA} = 24\ hours$$

This is also true if we consider part of the day, such as the composition of movement behaviours during the waking day. Though waking hours are typically not fixed, we can still carry out a compositional data analysis of the proportions.

$$t_{SB} + t_{LPA} + t_{MVPA} = waking\ hours$$

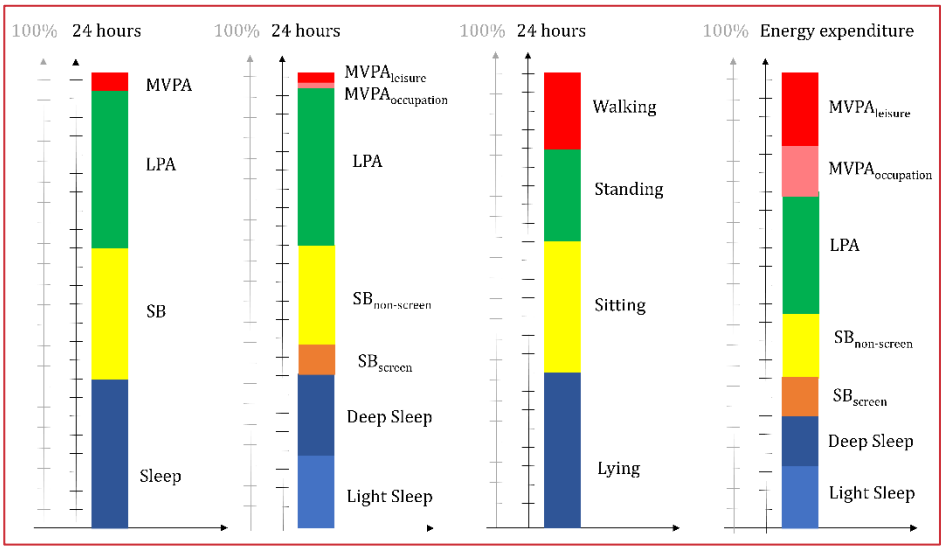


Figure 29
Visualization of the compositional nature of physical behaviour data

A composition can have an unlimited number of parts that can be defined by intensity band, activity type, context information or a combination of those, provided they are mutually exclusive. As a consequence of the fact they describe mutually exclusive components of a time or energy budget, each part only contains relative information rather than an absolute value and, then, the interpretation of compositional data is in terms of relative time spent in the different behaviours. If the data is regarded as a composition; mathematical transformation of the data is required prior to introducing the variables in a statistical model. For some applications, the absolute time may be important, in which case it would not be appropriate to apply the compositional transformation.

Compositional data transformations are simple and rely on logarithmic transformations. The purpose of this transformation is to resolve the difficulties around co-dependency and spurious

Abbreviations in this page:
LPA: light physical activity
MVPA: moderate-to-vigorous physical activity
SB: sedentary behaviour
VPA: vigorous physical activity

correlation associated with the compositional nature of these descriptors. Statistical models can, therefore, be adjusted for all physical behaviour components without incurring perfect collinearity. Specifically, the data transformations that have been used so far in 'physical behaviour epidemiology' are the centred log ratio [328,329] and the isometric-log ratio [326,330–332]. Using the centred log ratio, each component is centred according to the mean logarithm of all the components [324], which is mathematically expressed as:

$$z_i = \ln \frac{t_i}{\sqrt[D]{\prod_{j=1}^{D} t_j}} \text{ with } i \text{ indicating each component}$$

The sum of the D (number of components) centred log ratio-transformed variables is 0. This fixed sum means they are singular, and cannot be used in regression models. However, we can apply an additional transformation to the components to obtain a D-1 dimensional space without this constraint. This is referred to as the isometric log ratio-transformation when the new space uses an orthonormal basis. There are multiple such bases, however the most common approach in physical behaviour epidemiology research is shown below (e.g., SB, LPA, MVPA and sleep):

$$z_{SB} = \left(z_1 : \sqrt{\frac{3}{4}} ln \frac{SB}{(LPA \cdot MVPA \cdot Sleep)^{1/3}}, z_2 : \sqrt{\frac{2}{3}} ln \frac{LPA}{(MVPA \cdot Sleep)^{1/2}}, z_3 : \sqrt{\frac{1}{2}} ln \frac{MVPA}{Sleep} \right) (1)$$

$$z_{LIPA} = \left(z_1 : \sqrt{\frac{3}{4}} ln \frac{LPA}{(MVPA \cdot Sleep \cdot SB)^{1/3}}, z_2 : \sqrt{\frac{2}{3}} ln \frac{MVPA}{(Sleep \cdot SB)^{1/2}}, z_3 : \sqrt{\frac{1}{2}} ln \frac{Sleep}{SB} \right) (2)$$

$$z_{MVPA} = \left(z_1 : \sqrt{\frac{3}{4}} ln \frac{MVPA}{(Sleep \cdot SB \cdot LPA)^{1/3}}, z_2 : \sqrt{\frac{2}{3}} ln \frac{Sleep}{(SB \cdot LPA)^{1/2}}, z_3 : \sqrt{\frac{1}{2}} ln \frac{SB}{LPA} \right) (3)$$

$$z_{Sleep} = \left(z_1 : \sqrt{\frac{3}{4}} ln \frac{Sleep}{(SB \cdot LPA \cdot MVPA)^{1/3}}, z_2 : \sqrt{\frac{2}{3}} ln \frac{SB}{(LPA \cdot MVPA)^{1/2}}, z_3 : \sqrt{\frac{1}{2}} ln \frac{LPA}{MVPA} \right) (4)$$

Thus, the isometric log ratio produces a set of coordinates for each component (i.e., z_1, z_2 and z_3 in each component of the example above) that should be introduced together as covariates in any statistical model (see below considerations on the statistical model selection). The main difficulty associated with these transformations is in interpreting the results; this is a problem similar to (for example) in standard linear regression when a variable is log-transformed. For compositional data, a solution is to find an appropriate graphical representation of the results, keeping in mind the co-dependence of the parts and using model predictions rather deriving the estimate directly from model coefficients. Another difficulty arising from these mathematical transformations is related to having zeros or values close to zero in any of the components. This can happen in certain populations which may not perform VPA or even MVPA. Considering

very low values in a composition could lead to spurious correlations [333], usually, these values are either ignored in the analysis or imputed to stabilize the models [326].

Abbreviations in this page:
MVPA: moderate-to-vigorous physical activity

Statistical modelling

The third and last step of the analytical process relates to the decisions on how to model the associations between the selected descriptor(s) (with or without mathematical transformations) and health. As far back as the 1950's [2,334], many studies have investigated the epidemiological associations of physical behaviours with health outcomes. The use of accelerometers confirmed some of these associations, and allowed a better characterisation of the dose-response curve overcoming the cognitive biases of self-reports. However, most studies have solely focused on basic descriptors of one behaviour in isolation (e.g., MVPA). Out of the 11,765 publications identified in a search in the Web of Science on physical activity and accelerometers (**Figure 28, Panel A**), only 125 studies explored the interdependencies among physical behaviours using isotemporal substitution models, multivariate pattern analysis or functional data analysis (**Figure 30**) [335]. This consensus group believes that now is the right time to move to more detailed and informative studies on the combined effects and interactions across physical behaviours on health outcomes. **Table 24** contains a brief definition of each of the statistical models described in this section.

Table 24
Brief description of approaches to analyse associations between physical behaviours and health outcomes.

Statistical model	Brief description	Examples
Standard regression modelling	Traditional models establishing the relationship between a set of explanatory variables and an outcome (i.e., health outcome). Exposure is usually limited to a single time use behaviour. Interpretation is in terms of increasing time in one behaviour.	[336,337]
Isotemporal substitution model	Isotemporal substitution models examine the theoretical effects of displacing a fixed duration of time between behaviours. Given the fixed and finite duration of a day, increasing time in one movement behaviour (e.g., LPA) will result in a net equal and opposite change in other movement behaviours (e.g., SB). Interpretation is in terms of substituting one behaviour for other behaviours.	[338,339]
Multivariate Pattern Analysis	A regression approach/analysis that can handle an unlimited number of multicollinear explanatory variables by using latent variable modelling. Models are cross-validated to optimize predictive ability. Interpretation is based on the complete pattern of associations among the explanatory variables in relation to the outcome.	[328,340–343]
Functional data analysis	Functional data analysis is an extension of scalar regression where the exposure or outcome is defined as a function rather than a scalar variable. The function can describe the full distribution of intensity of acceleration or the time-series of acceleration over the day. The function can be included in standard regression analysis through dimensional reduction techniques. Interpretation is in terms of certain accelerometer trace shapes.	[210,314,322,344,345]
Machine learning	ML entails a broad range of techniques to automate the learning of high-dimensional and/or non-linear patterns in data with predictive ability (supervised machine learning) or data reduction (unsupervised machine learning) as its core priority.	[346–348]

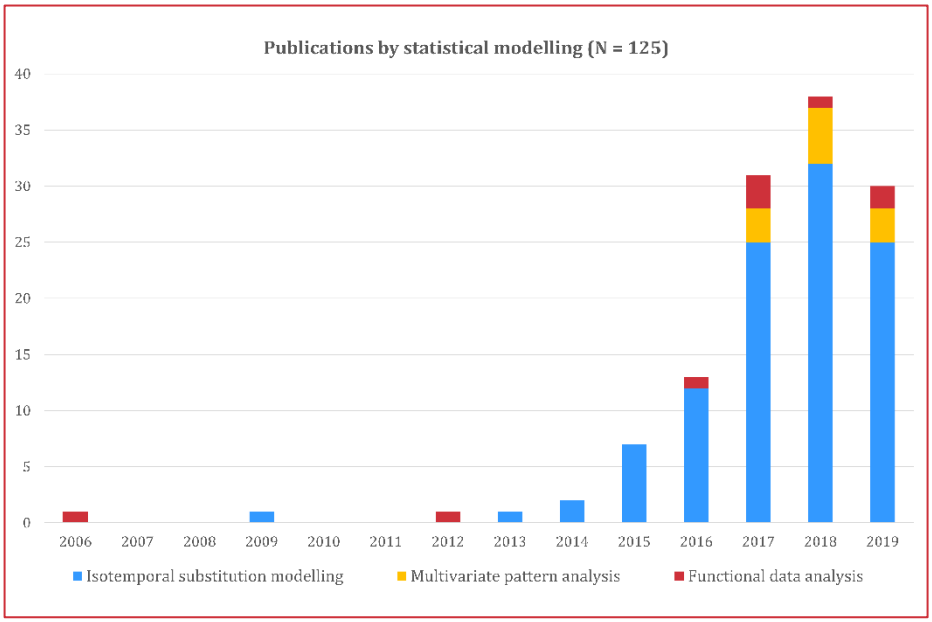


Figure 30
Number of publications using some of the approaches described in the present document from inception to December 31st, 2019.
Search syntax introduced in the Web of Science: isotemporal substitution models: (((((("physical activity")) OR "sedentary")) AND ((acceleromet*) OR actigraph*) AND ("isotemporal substitution")); multivariate pattern analysis: (((((("physical activity")) OR "sedentary")) AND ((acceleromet*) OR actigraph*) AND ("Physical activity signature" OR "multivariate pattern analysis")); functional data analysis: (((((("physical activity")) OR "sedentary")) AND ((acceleromet*) OR actigraph*) AND ("Physical activity signature" OR "functional data analysis")).

Abbreviations in this page:
LPA: light physical activity
MVPA: moderate-to-vigorous physical activity
SB: sedentary behaviour

Standard regression modelling

Standard regression is the most frequently used statistical model in the field, often including the physical behaviour descriptor as a continuous exposure variable in a linear, logistic or Cox regression (depending on the outcome of interest). Standard linear regression models are interpreted in terms of the (theoretical) effect of increasing the explanatory variable on the outcome, under a linear relationship. Standard linear regression models are usually adjusted for the covariates that could influence the association of interest. Highly correlated explanatory variables result in multicollinearity, which is a phenomenon in which redundant information carried by predictors leads to erratic estimation of the models [349].

Standard regression models can also be used with compositional ILR-transformed descriptors, which may eliminate that part of the collinearity which arises from the fixed sum (or closure) constraint [326,327]. In this case, the model coefficients are interpreted in terms of time replacements across behaviours. For example, the estimate for the z_1 coordinate of the z_{SB} equation presented above represents the effect of increasing SB while proportionally reducing the time in LPA, MVPA and sleep. The dose-response association between a specific behaviour and the health outcome is assumed to be logarithmic (curvilinear) using compositionally-transformed descriptors. Likewise, the regression model predictions (using compositional data) can be used to estimate the time replacement between pairs of behaviours (e.g., reallocating time from SB to MVPA). This results in a similar interpretation to the isotemporal substitution models. When examining longitudinal associations, advanced regression models (e.g., survival analysis using Cox regression) may be used with either absolute descriptors [274,321,350] or compositionally transformed descriptors [331].

Isotemporal substitution models

The isotemporal substitution modelling framework considers potential outcomes of increasing one behaviour at the expense of another and whether the strength of the association is dependent on the behaviour being displaced. Isotemporal substitution models are standard regressions in which all-but-one of the time-use behaviours are introduced as the exposure (together with the pertinent covariates) and the health outcome is the dependent variable. These models examine the estimated effects of replacing time spent in one behaviour (the missing behaviour in the model) with an equal amount of time spent in another, while keeping monitor wear time constant. They do so by dropping the behaviour of interest from the model (otherwise, the model would suffer from perfect collinearity). The linear effects of the pair-wise reallocations are then estimated from the model coefficients. Similar interpretations of time replacement between pairs of behaviours can be obtained from applying standard linear regression over compositional data.

Abbreviations in this page:
PA: physical activity

Multivariate pattern analysis

Multivariate pattern analysis can handle completely collinear explanatory variables by combining the data into orthogonal latent variables [351]. Thereby, this method tackles collinearity as a dimension reduction problem, rather than a data transformation (as compositional data analysis does). Multivariate pattern analysis is especially well-suited to analyse a wide range of collinear descriptors, such as the intensity spectrum, without requiring any data transformation [310,311], although transformations can be done to make distributions within bins more normal and linearly associated with the outcome. Another important feature is that the models are optimized for predictive ability by Monte-Carlo resampling whereby half of the data are repeatedly used for modelling and half for prediction [340]. In this way, the optimal number of latent variables can be determined and only relevant features in the descriptor retained.

Multivariate pattern analysis uses partial least squares regression modelling [351], or other latent-variable regression models [342], to determine the multivariate association pattern. Partial least square regression decomposes the explanatory variables into orthogonal linear combinations (partial least square components), while simultaneously maximizing the covariance with the outcome variable. Similar procedures to reduce the data can be observed in factor analysis, principal component analysis, or joint and individual variance explained models [352]. Multivariate pattern analysis differs from these others by creating components that maximize the covariation with the outcome, not internally among the explanatory variables. Joint and individual variance explained models seek to maximize the variance explained across explanatory variables

Abbreviations in this page:
ENMO: Euclidean norm minus 1*G*
MAD: mean amplitude deviation
PA: physical activity

assuming that they come from different dimensions (e.g., PA, sleep, and circadian rhythms) and improving the within and between dimension representation [352]. The procedure for obtaining the multivariate patterns is completely data-driven, with no assumptions on variable distributions or degree of collinearity among variables. Selectivity ratios are calculated to express and rank each single explanatory variables' association with the outcome [353,354]. The selectivity ratio represents each explanatory variable's ratio of explained to residual variance in relation to the outcome (**Figure 31**). By replacing residual variance with total variance in the denominator, a straight-forward measure of explained variance can be obtained [343]. Multivariate pattern analysis has been applied with time-use descriptors and intensity spectrum in both their absolute scale and with the compositional centred log ratio transformation [328]. Since multivariate pattern analysis can handle singular data (e.g., CLR-transformed data), the isometric log ratio transformation is not necessary if modelling compositional data.

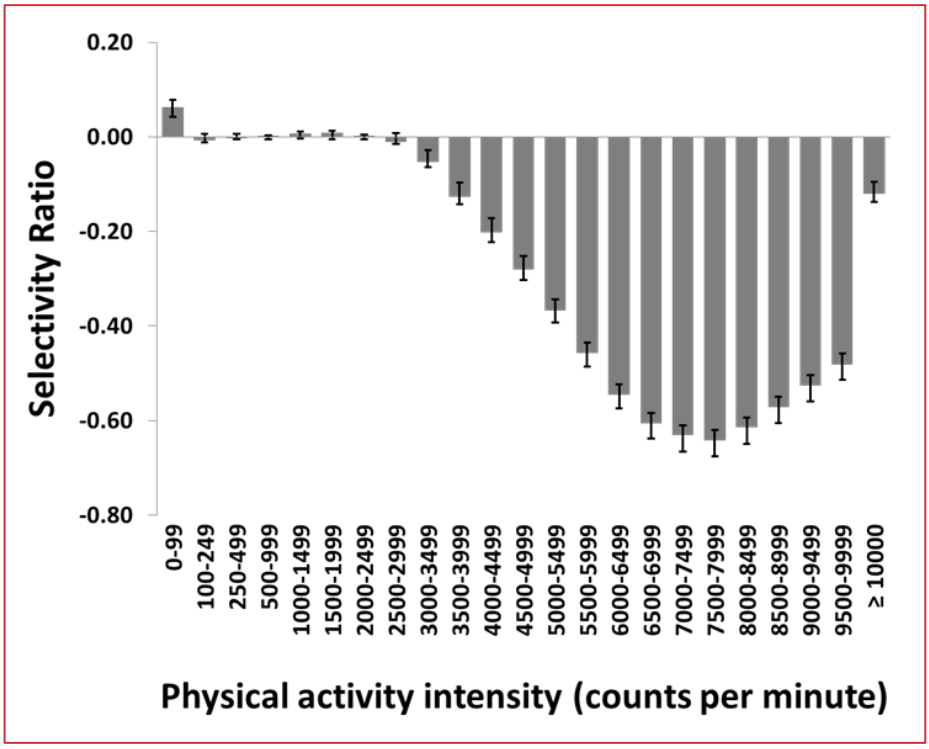


Figure 31
Multivariate pattern analysis example. Accelerometer model: ActiGraph GT3X+, sampling frequency: 30 Hz, body attachment site: right hip; awake time recording protocol. Selectivity ratio represents the explained-to-total outcome variance ratio. Taken from Aadland et al. [328] with permission from the publisher

Functional data analysis

Functional data analysis is an extension of standard regression analysis where the exposure or the outcome (or both) is a function instead of a scalar [344,355,356]. In physical behaviour epidemiology, the rationale of functional data analysis in the context of accelerometer data comes from the availability of moment-by-moment acceleration data allowing the use of the entire range of accelerations, whatever the aggregated metric used (e.g., counts, ENMO, MAD) [313,314]. The acceleration functions described in above can be used in functional data analysis. A first step often consists in

smoothing the function of interest so that the smoothed function can then be used in functional data analysis, although some approaches do not smooth the data at subject level and rather pool the data across subjects to avoid the loss of information from the accelerometer signal. For example, when the interest is in the distribution of acceleration over time of the day, one can reduce data into 10 minute epochs as the objective is to assess when individuals are more or less active at each time of the day [357]. When the function of interest is the acceleration density distribution, Gaussian Kernel smoothing methods can be used (**Figure 32**) [358]. In that case, careful attention should be given to the number and place of nodes for acceleration values: a higher number of nodes should be present in the acceleration range where most of the time is spent. Then, the smoothed function of interest can be used for further analysis as an outcome variable (Function-on-scalar analysis), an exposure (Scalar-on-function analysis), or both (Function-on-function analysis) using functional data analysis regression techniques.

Abbreviations in this page:
PA: physical activity

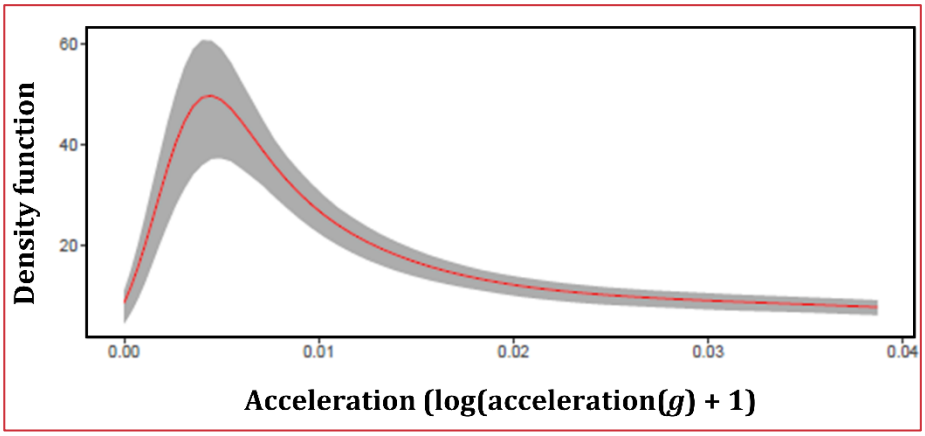


Figure 32
Smooth mean and interquartile acceleration density function
Red curve represents the mean density function of the study population and the grey area the interquartile range.

Machine learning for epidemiological analysis

Machine learning methods provide a broad range of techniques to identify patterns in data. Although it has been increasingly used to derive descriptors from raw accelerometer data [318], machine learning has rarely been applied to the study of the associations of accelerometer data descriptors (examples of machine learning for health association analysis using physical behaviour data include [347,348]). As machine learning methods typically emphasise prediction or data reduction, they are most often relevant for hypothesis generation and data exploration. While there is no clear distinction between conventional statistical methods and machine learning, there is typically a different emphasis, and so they can be difficult to apply directly to problems requiring statistical inference. Bi et al. discuss possible epidemiologic applications of a wide range of machine learning methods in detail [346]. Examples of machine learning methods which could be applied to health association analysis using accelerometer data include Decision Trees/ Random Forests, Support Vector Machines and Neural Networks.

Analytical approaches: discussion and practical considerations

Abbreviations in this page:
CPM: counts per minute
FDA: functional data analysis
ISO: isotemporal substitution model
MPA: multivariate pattern analysis
MVPA: moderate-to-vigorous physical activity
PAEE: physical activity-related

This section describes and discusses the different analytical approaches' applicability in various situations, including the combination of accelerometer descriptors and statistical models with and without mathematical (compositional) transformation. We provide practical considerations in relation to (1) how informative each analytical approach is for public health messaging and (2) what context or type of research question is more appropriate for the use of each analytical approach. Additionally, the performance of these approaches in relation to potential risks of closure or collinearity, relationship assumptions between physical behaviours and health outcomes, and their interpretation for public health guidelines are presented in **Table 25**.

Table 25. Summary of analytical approaches (including descriptor, mathematical transformation and statistical model) strengths and limitations in relation to closure, collinearity, relation-shape assumptions and interpretation relative to public health guidelines.

Descriptor	Compositional transform	Statistical modelling	Risk of closure[a]?	Risk of collinearity?	Handles closure?	Handles collinearity?	Relationship assumptions	Allow investigation of longitudinal associations (e.g., Cox regression)	Interpretation relative to guidelines? (e.g., 150 min/week of MVPA)
Average acceleration	**No**	**Standard**	No	No	-	-	Linear	Yes	No
Time-use descriptors	**No**	**Standard**	Yes	Yes	No	No	Linear	Yes	Yes
	Yes	**Standard**	Yes	Yes	Yes	In part[b]	Log-linear	Yes	Yes
	No	**ISO**	Yes	Yes	Yes	No	Linear	Yes	Yes
	No	**MPA**	Yes	Yes	No	No	Linear	Not at the moment	Yes
	Yes	**MPA**	Yes	Yes	Yes	Yes	Log-Linear	Not at the moment	Yes
Intensity spectrum	**No**	**Standard**	Yes	Yes	No	No	Linear	Yes	Yes[c]
	Yes	**Standard**	Yes	Yes	Yes	In part[b]	Log-linear	Yes	Yes[c]
	No	**ISO**	Yes	Yes	Yes	No	Linear	Yes	Yes[c]
	No	**MPA**	Yes	Yes	No	No	Linear	Not at the moment	Yes[c]
	Yes	**MPA**	Yes	Yes	Yes	Yes	Log-linear	Not at the moment	Yes[c]
Intensity gradient	**No**	**Standard**	No	No	-	-	Linear	Yes	No
	No	**FDA**	No	No	-	-	Fewer assumptions than other models	Yes	Yes[d]
MX metrics	**No**	**Standard**	Yes	Yes	No	No	Linear	Yes	Yes[c]
	No	**MPA**	Yes	Yes	No	Yes	Linear	Not at the moment	Yes[c]
Other acceleration functions	**No**	**FDA**	No	No	-	-	Fewer assumptions than other models	Yes	Yes[d]

[a]Closure refers to whether a certain descriptor is a specific part of the daily time constraint (i.e., it is measured in time per day).
[b]Indicates that it solves the collinearity due to the closure, but collinearity can still exist across the CoDA-transformed variables.
[c]Indicates that the interpretation is made through a post-hoc application of validated cut-points to identify the PA intensity (e.g., MVPA).
Indicates that more work is needed on the interpretation of functional data analysis, an example can be found elsewhere [315].
ISO: isotemporal substitution models; MPA: multivariate pattern analysis; FDA: functional data analysis; ML: machine learning.

Average acceleration and standard regression

The average acceleration provides the simplest estimate of the overall movement and the simplest proxy for total daily PAEE.

Abbreviations in this page:
MVPA: moderate-to-vigorous physical activity
PA: physical activity
PAEE: physical activity-related energy expenditure
SB: sedentary behaviour

However, it usually has only moderate correlations (with total daily PAEE), so we do not recommend average acceleration as a direct estimator of PAEE. Statistical interpretation of findings using standard linear regression is straightforward since there is a single variable representing the overall activity volume. Thus, co-dependence with other explanatory variables is not usually a concern with average acceleration. This consensus group believes the average acceleration is useful as a covariate to remove the confounding effect of overall PA in a given association analysis, or as the main exposure in cases where it explains a large proportion of the PAEE in a certain cohort. Beyond this, the average acceleration on its own is not very informative relative to the associations of specific physical behaviours with health outcomes, limiting its applicability for public health messaging. A recent study has proposed the minimum clinically informative difference for average acceleration derived from wrist-worn accelerometers [359], which may ease the interpretation of this descriptor, but further studies are needed in this regard. Nevertheless, the average acceleration may be the best descriptor to test the "move more" message reported in several guidelines.

Time-use behaviours or intensity spectrum and standard linear regression

Among the time-use descriptors, time spent in PA intensities is the most frequently-used in the field of 'physical behaviour epidemiology', while PA types have gained momentum in the last years. These descriptors are often introduced in standard linear regression models to test the association of time spent in a certain behaviour with health outcomes. It is useful for comparing estimates with other cohorts that have already used this approach. The intensity spectrum is an extension of PA intensities with higher resolution and without assumptions on energy bands. Therefore, we discuss their usefulness in the analytical approaches together. When using time-use behaviours, it is important to consider requirements for bouts in these behaviours. We observe that there is no consensus in the literature on how a bout should be calculated, including the definition of acceptable grace period and the definition of the minimum and maximum duration. However, usually we see 30-min bouts are used for SB and 10-min bouts for MVPA often allowing short time intervals outside the behaviour of interest [360]. It is unclear how much these choices are driven by a desire for harmonisation, by public health guidelines, or by evidence.

Time-use behaviours (or intensity spectrum) are co-dependent on each other and standard linear regression does not handle closure and collinearity among explanatory variables. Therefore, when using time-use behaviours or the intensity spectrum, standard linear regression adjusted for all physical behaviour components may show multicollinearity between variables [349]. Variance inflation factors

Abbreviations in this page:
LPA: light physical activity
MVPA: moderate-to-vigorous physical activity
SB: sedentary behaviour

are frequently used to quantify multicollinearity in linear regression models. However, variance inflation factors are unable to explain inconsistencies between standard linear regression models sequentially excluding a behaviour from the explanatory variables [327]. Therefore, it has been suggested that variance inflation factors are not an acceptable diagnostic indicator for the interdependency between time-use descriptors [327]. Additionally, standard linear regression assumes linearity in the association, while PA intensities are known generally not to be linearly related to health outcomes. This consensus group recommends moving forward to other analytical approaches more suitable for studying the co-dependencies among the time-use behaviours. In this regard, transforming the time-use behaviours using the compositional data isometric log ratio transformation has been suggested. Using the isometric log ratio transformation, each variable indicates the time spent in a given behaviour (e.g., MVPA) relative to the time spent in the rest of behaviours of the composition (e.g., SB, LPA and sleep time). In other words, it quantifies the effect of increasing the time in a behaviour while proportionally reducing the time in the rest of behaviours included. Pair-wise reallocation of time can also be interpreted from standard regression predictions on specific time compositions arising from hypothetical reallocations of time rather than from regression coefficients (as it is done in isotemporal substitution models). Although indicators on data quality cannot be analysed compositionally, they can be used as covariates in the models [361]. Once the variables have been transformed, the co-dependency among the time-use descriptors relative to their time closure is already solved (i.e., it accounts for the reallocation of time among variables). It is however important to consider that the transformed variables can still be collinear, and collinearity should be investigated because standard linear regression cannot handle collinearity, regardless of its source. This is especially problematic when analysing the intensity spectrum since it provides a wide range of variables (usually more than 10) that are highly correlated, even if using isomeric log ratio-transformed variables [328].

This consensus group believes the field of 'physical behaviour epidemiology' should move to studying the combined effects and interactions of physical behaviours on health. Using isometric log ratio-transformed time-use behaviours and standard regression would be a feasible alternative to study the combined effects and interactions of physical behaviours on health [326,327]. Additionally, this approach would be informative for public health messaging by investigating what combinations of behaviours (considering every behaviour that occurs in the day) are more beneficial for health. Clustering groups of people based on their behaviours is also an alternative to investigate the interactions between behaviours. However,

compositional data analysis means it is possible to study the variables on a continuous scale (which carries more information than comparing groups with certain characteristics) while accounting for the interdependencies between them. Otherwise, the use of standard regression models to analyse the intensity spectrum variables (either transformed or not) is not recommended because of the high risk of collinearity issues in the transformed variables [328]. Other statistical models should be used for this purpose. An advantage of standard regression models is that they can be used with more advanced regression models to investigate longitudinal associations, either with absolute [274,321,350] physical behaviour data or with compositional transformations [331]. As a challenge, standard regression models with compositional data may need appropriate graphical representation of the results to interpret the magnitude of the association, keeping in mind the co-dependence of the parts and using model predictions rather than deriving the result directly from model coefficients.

Time-use behaviours or intensity spectrum and isotemporal substitution models

Isotemporal substitution modelling carries forward the main limitations of the standard linear regression, that is, it cannot handle multicollinearity and it assumes linearity of physical behaviours with health (as the magnitude of the association is derived from regression coefficients). These important limitations preclude us from recommending the use of isotemporal models with time-use descriptors. However, it is important to note that this approach provides broadly similar findings to compositional isometric log ratio transformation of time-use behaviours and standard regression [362]. Public health messaging can be complemented with information on the effect of reallocating the daily time across behaviours (either with isotemporal substitution models or with standard regression over compositional data). The intensity spectrum has not been analysed with isotemporal substitution models at this time. We do not recommend such an analysis since the large number of variables in the intensity spectrum would complicate the analysis and interpretation.

Time-use behaviours or intensity spectrum and multivariate pattern analysis

Multivariate pattern analysis provides an alternative approach that fully handles the collinearity among explanatory variables using latent variable modelling. As such, collinearity is approached as a dimension reduction problem in which the parts of the explanatory variable that share more variance with the outcome are retained for the model. Multivariate pattern analysis does not directly quantify the effect of reallocation of time across behaviours on health, but it

Abbreviations in this page:
MVPA: moderate-to-vigorous physical activity
MX: person's most active X min
PA: physical activity

rather describes the pattern of associations for the behaviours with the outcome accounting for the correlated structure of the data. Therefore, the associations with health are interpreted for each descriptor (each PA intensity or bin in the intensity spectrum) considering its co-dependency with the rest, but without quantifying time exchange between descriptors. A limitation of this analytical approach is that partial least squares regression models cannot be adjusted as usual. If covariates are included in the partial least squares model, they will contribute their shared variance with PA and the outcome (and possibly stabilize the model), but the association between PA and the outcome will not be independent of the covariates. The association pattern for PA will in many cases be similar, but the model fit (explained variance) will differ if covariates are associated with the outcome. Aadland et al. proposed to obtain residuals for the outcome from a linear regression model including confounders as explanatory variables, prior to entering the outcome variable in the PLS model [310,311,328]. This challenge is particularly important for the analysis of categorical or time-dependent outcomes (e.g., mortality).

Likewise, time-use behaviours or the intensity spectrum could be transformed as compositional data before introducing them in the model. Since multivariate pattern analysis can handle singular data, the use of isometric coordinates is not necessary. Aadland et al. recently compared the use of raw and centred log ratio-transformed time-use and intensity spectrum descriptors with respect to associations with metabolic health using multivariate pattern analysis [328]. While associations appeared to differ substantially, the interpretation of associations, considering the absolute and relative interpretation, were partly equivalent. However, the interpretation of the centred log ratio-transformed variables may not be very informative for public health messaging as they represent the effect of time exchange from the geometric mean of the descriptor distribution to a specific descriptor (e.g., MVPA or any intensity spectrum bin).

The main strength of multivariate pattern analysis is that it can fully handle collinearity among explanatory variables by reducing the dimensions of the descriptors and keeping only the parts which share more variance with the outcome. Other similar alternatives that try to reduce dimensionality of the data while retaining relevant information by increasing covariance among descriptors include factor analysis, principal component analysis, or joint and individual variance explained models [352]. Thus, this consensus group recommends considering this approach to analyse many explanatory variables (e.g., intensity spectrum, MX metrics) in relation to health [363]. Since the intensity spectrum variables and the MX metrics are strongly correlated [328], and therefore do not contribute uniquely to explain the outcome, multivariate pattern analysis is more

Abbreviations in this page:
PA: physical activity
PAEE: physical activity-related energy expenditure

meaningful than other approaches to overcome these potential collinearity issues. For example, Aadland et al. used an intensity spectrum composed of 16 bins to study the association of the PA intensity pattern with metabolic health in children [311]. There is no clear recommendation on the number of bins (or number of explanatory variables) to generate for this analytical approach, though previous studies have used from 16 (uniaxial data) [311] to 102 (triaxial data) [310] intensity bins. Resolution may influence the relationship with the outcome and depend on the sample characteristics; thus, further research is needed.

Intensity gradient and standard regression

The intensity gradient was developed to: 1) capture the entire intensity distribution, 2) avoid the reliance on calibration protocols (that are, by nature, population and protocol-specific) and 3) provide information that complements the average acceleration. The latter enables the intensity gradient to be used alongside average acceleration to more fully describe the 24 h movement profile by capturing both the volume and intensity of PA. This means the intensity gradient and average acceleration can be used together in standard linear regression models to investigate independent, additive and interactive associations of volume and intensity of PA with health. Otherwise, it does not allow direct estimation of the population attainment of current public health guidelines on physical behaviours. Since the intensity gradient is not affected by any kind of cut-points (which usually depend on the population characteristics), its use is also recommended for surveillance and to compare PA differences across populations without making assumptions on their PAEE.

Intensity gradient or intensity distribution and functional data analysis

Although the intensity gradient can be summarised in a scalar (using the linear regression coefficient between time and intensity both on the logarithmic scale), its defining function can be directly used in functional data analysis. Likewise, the acceleration distribution over time of the day or the acceleration density can be defined by a function and used in functional data analysis. Any of these acceleration functions can be used as an explanatory variable in regression models such as linear [314], logistic or Cox regression models, using scalar-on-function data analysis [344]. For example, in the case of the acceleration density function as explanatory variable, the association with the event of interest is described along the acceleration range [314]. This shows acceleration sections that are associated with the outcome of interest by accounting for the full distribution of acceleration, allowing identification of cut-point above which proportion of time spent above this acceleration cut-point is associated with the outcome of interest. Once acceleration above or below which an association with the outcome of interest is found, it is

Abbreviations in this page:
MVPA: moderate-to-vigorous physical activity
MX: person's most active X min
PA: physical activity
PAEE: physical activity-related energy expenditure

possible to estimate differences (odds ratio or hazard ratio depending on the regression model) in the outcome of interest by allocating time below this cut-point to time above this cut-point [315].

This analytical approach is not affected by multicollinearity since it handles the data continuity instead of using several reduced variables. Once the association of certain acceleration intensities or patterns with health is described, functional data analysis can also test the effects of reallocation time from sections of the acceleration range to other sections relevant for health; thus, it can also consider the closure of the data. Another advantage is that functional data analysis models relax the assumptions of linearity in particular behaviours made by other statistical models by not splitting the continuous acceleration into categories. Likewise, functional data models can detect the sections of the accelerometer data that are important for (i.e., associated with) a certain health outcome. Among its main limitations, the acceleration functions included usually carry much information that is not relevant for the outcome, but it is considered in the analysis. However, its main drawback is the difficulty in translating the findings into useful and straightforward public health messages. Investigation of how to make the conclusions of functional data analysis relevant for public health guidelines is highly encouraged by this consensus group (see [315] for an example).

MX metrics and standard regression

A major advantage of using the MX metrics is that the analysis is not affected by cut-points assumptions on energy expenditure, while cut-points may be post-hoc applied to ease the public health messaging. For example, if the M60 of a child is 230 m*g*, this can be compared to an MVPA cut-point, e.g., 200 m*g* [61], showing that the child is meeting the 60 min daily MVPA recommendation. However, if compared to a more stringent 250 m*g* MVPA cut-point, the child does not quite reach the recommendation. An advantage of this approach is that the intensity of physical activity for the specified duration is captured regardless of how inactive a person is - no one scores zero. This makes it particularly suited to describing the physical activity of less active populations where an intensity cut-point may never be exceeded. Regarding statistical modelling, as with time-use descriptors or the intensity spectrum, the MX metrics are usually described for a wide range of variables as explained in the 'MX metrics' section. These MX metrics are likely to be co-dependent as they are time-use descriptors. This co-dependency may result in multicollinearity in the standard linear regression models. Likewise, whether several MX metrics are derived from the accelerometer recording to describe the PA pattern, each MX metric would carry partial and relative information on the pattern. Thus, a compositional transformation of the MX metrics would also be interesting when using

standard linear regression models, although this approach has not been tested at the time.

The MX metrics provide novel information on PA patterns that should be further investigated. The fact that analyses using these metrics are cut-point free could be important for surveillance and comparisons across populations with different characteristics. It is also noteworthy that the post-hoc implementation of cut-points allow those individuals meeting the pertinent public health guidelines to be identified.

MX metrics and multivariate pattern analysis

The usefulness of the MX metrics with multivariate pattern analysis has not been investigated yet. However, since one of the limitations of the MX metrics with standard linear regression is the collinearity among the explanatory variables, multivariate pattern analysis could provide new insights by reducing the dimension of the explanatory variables and overcoming the collinearity.

Multiple descriptors and machine learning

Machine learning describes a broad range of techniques to automate finding patterns in data with a focus on predictive ability (supervised machine learning) or data reduction (unsupervised machine learning). Although machine learning methods have been widely applied to derive accelerometer descriptors [318], they have rarely been applied to the study of associations of accelerometer data descriptors with health [348]. Different machine learning approaches have different strengths and limitations. In general, strengths of machine learning methods for health association analysis include their usefulness for data-driven hypothesis generation, their capacity to handle multi-dimensional data, their ability to find non-linear patterns, and the possibility of training a model in one dataset and updating it in another. However, it can be difficult to interpret how results are obtained and their significance for public health guidelines. Machine learning methods can also be data-hungry and computationally intensive. Overfitting and sensitivity to (potentially unknown) biases in the training data are risks.

In some ways, multivariate pattern analyses and the other dimension reduction methods can be considered machine learning methods. However, as machine learning represents a broad range of methods, with individual strengths and drawbacks, general statements about usefulness and relevance in physical behaviour epidemiology should not be made before a wider range of these methods have been applied and tested in this field. The Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD) Initiative developed a set of recommendations for the reporting of studies developing, validating, or updating a

Abbreviations in this page:
MVPA: moderate-to-vigorous physical activity
MX: person's most active X min
PA: physical activity
TRIPOD: Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis

Abbreviations in this page:
MVPA: moderate-to-vigorous physical activity
PA: physical activity
SB: sedentary behaviour

machine learning-based prediction models for diagnostic or prognostic purposes [364].

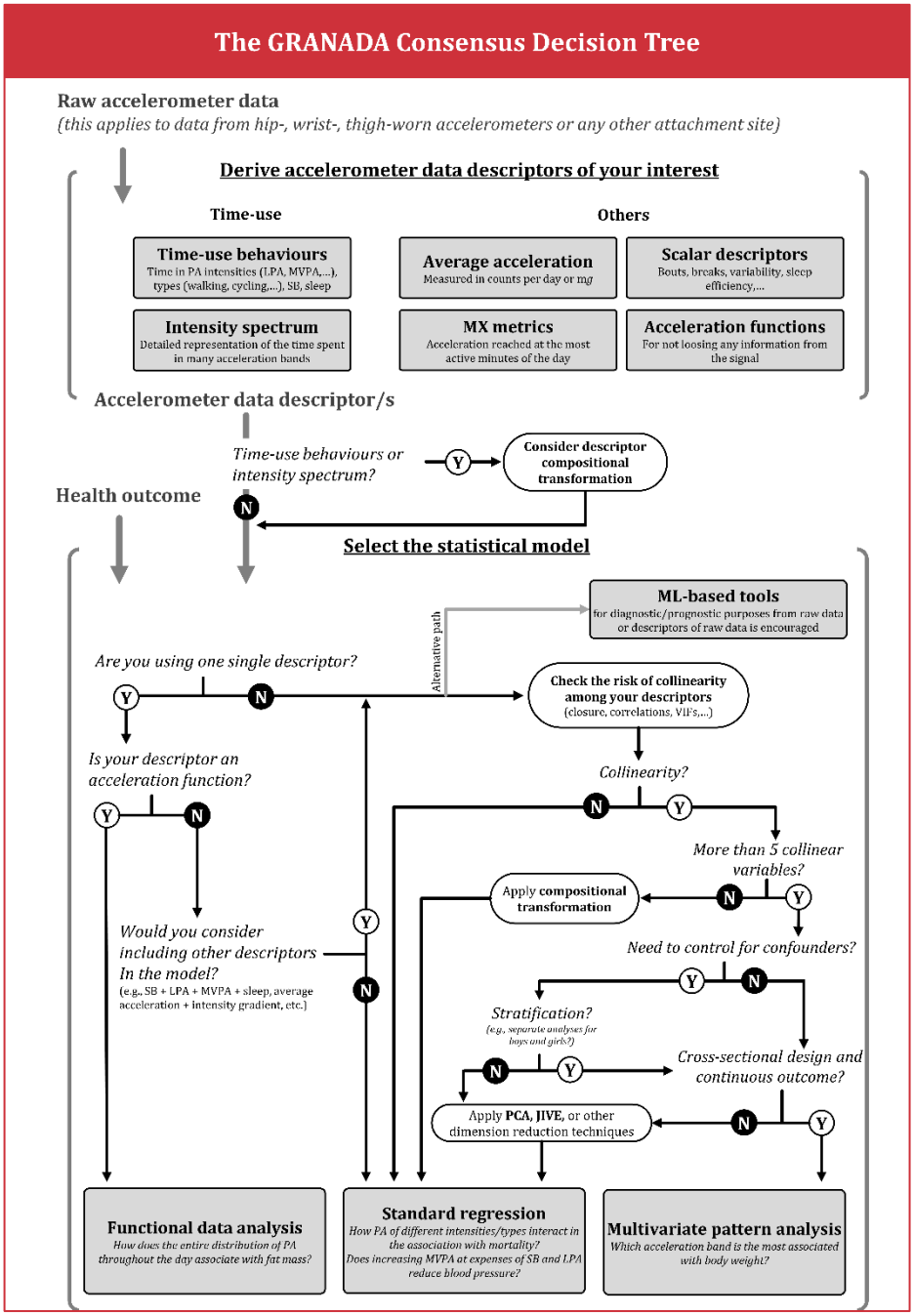


Figure 33
The GRANADA consensus decision tree and research question examples to assist in the selection of an analytical approach in the field of "physical behaviour epidemiology"

Future research directions

The international workshop held in Granada ("International Workshop: A focus on analytical methods to analyse accelerometer-measured physical activity"), as well as the later expert meeting and the work developed in the following months by this author group, has initiated a discussion on analytical approaches, and their usefulness for developing public health guidelines. In 2008, the American public health guidelines on PA recommended 150 min/week of MVPA in adults and older adults, and 60 min/day of MVPA in children as beneficial PA levels to improve health [7], which was also recommended by other national agencies and by more recent international

guidelines (https://www.who.int/publications-detail/global-action-plan-on-physical-activity-2018%E2%80%932030). Further studies have extensively confirmed this recommendation. We propose future research directions based on the research gaps identified, i.e., the uncertainty regarding the accelerometer data descriptors to use and about what analytical approaches best suit given research questions. The authors of this consensus article agree that investigations determining associations between physical behaviours and health should be extended to understand the interplay of 'physical behaviours' in their relationship with health. Measurement and processing capacity are now richer than when these guidelines were developed, thus, the field would benefit from further information on how different intensities and types of PA interact to improve health. At the same time, the focus on translation of findings to meaningful information for interpretation in practice cannot be lost when using advanced analytical models. The main implications for the analysis of accelerometer data proposed and agreed by the authors of this consensus manuscript are presented below.

Abbreviations in this page:
LPA: light physical activity
MVPA: moderate-to-vigorous physical activity
PA: physical activity
SB: sedentary behaviour

Short-term agenda

- Clear communication on the rationale for the use and limitations of each analytical approach in studies is important for a meaningful interpretation of the findings. Practical recommendations for this are provided above and a decision tree was developed (**Figure 33**) to assist researchers with decision making.
- Investigation of the associations of physical behaviours with health using different analytical approaches is encouraged. In an ideal world, the field of physical behaviour epidemiology needs to draw consistent conclusions independently of the applied analytical approach. To do so, clear reporting on the interpretation of findings derived from each analytical approach is crucial to understand 'a priori' inconsistencies across methods and consider their results in a proper manner. Triangulation of results from different analytical approaches is currently the best solution to quantify the associations of physical behaviours with health outcomes. Additionally, using the best-suited analytical approaches for a given research question is crucial (see **Figure 33**).
- Some accelerometer data descriptors include a wide range of variables. However, very few studies report adjusting for multiple testing in their analyses. Methods to adjust for multiple testing should be applied when the number of comparisons requires it.
- Machine learning-based approaches for diagnostic/prognostic purposes (associations with health outcomes) are worth

Abbreviations in this page:
MVPA: moderate-to-vigorous physical activity
PA: physical activity
SB: sedentary behaviour
TRIPOD: Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis

Box 1. Consensus points from the GRANADA report

1. The study of the association between physical behaviours (i.e., PA, SB and sleep) and health should move to a more thorough investigation of the interactions and co-dependencies between different behaviours (or physical activity intensities) and health. Several analytical approaches are provided in this consensus document, although none of them is free from limitations.
2. We recommend investigating more detailed physical activity intensities than the typically studied (i.e., SB and MVPA). Examples include light physical activity of different intensities or the more fine-grained intensity bands as described in this document.
3. Public health guidelines on physical behaviours should acknowledge that behaviours are co-dependent and this may affect the guidelines as traditionally understood.
4. Further investigation in functional data analysis and machine learning is needed concerning the associations of physical behaviours with health.
5. There is not a gold-standard able to test which analytical approach is the best for a given research question. Thus, we cannot make a strong recommendation on a single analytical approach. Instead, we provide some practical recommendations to select analytical approaches well-suited for a given research question. Triangulation across findings from different analytical approaches is currently the best solution.

implementing in the field, although they have been barely explored so far. We encourage their use and a transparent reporting of the resulting tools by following the TRIPOD initiative checklist.

- Translating study findings to meaningful information for guidelines and practitioners should always be a priority. Accurate reporting of study findings, interpretation, and practical implications is highly encouraged.

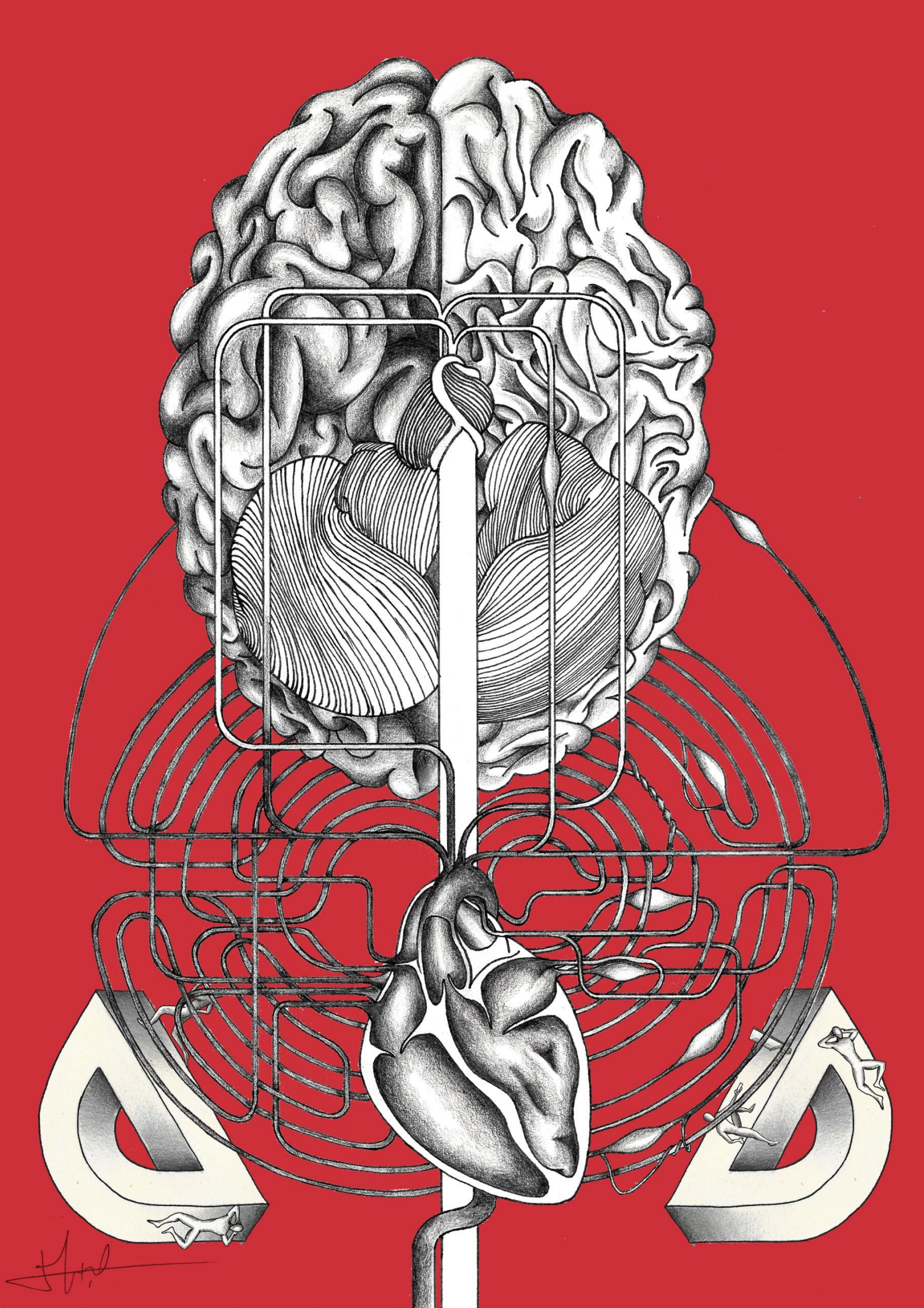
Long-term agenda

- How to conveniently adjust for confounders in multivariate pattern analysis should be investigated and its application should be extended to time-dependent outcomes (e.g., survival analysis with mortality outcomes).
- Further efforts are needed in order to translate functional data analysis and other advanced analytical approaches outputs into meaningful information for public health guidelines and practitioners.
- It will be a matter of interest in the future to evaluate whether the information gathered from these metrics and approaches can result in complementary information for public health guidelines. Such complementary information may result in more specific recommendations for certain health outcomes or populations, or even in their implementation at population level through movement sensors using evidence-based goals on PA intensity, duration, timing or type among others.

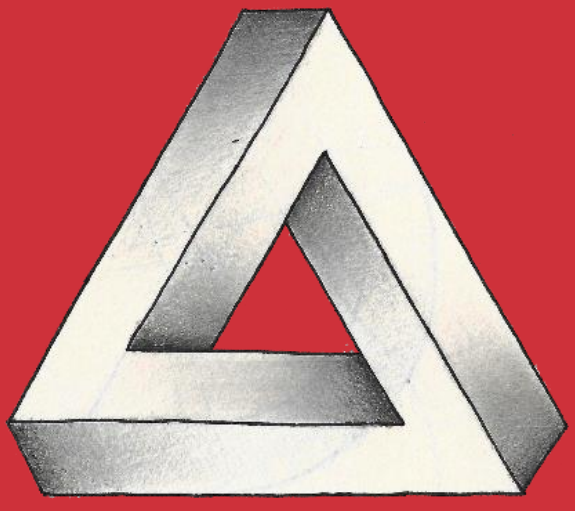
Conclusions

Authors from this group agreed on a number of consensus points and research needs which are relevant for physical behaviour epidemiology (see **Box 1**). This consensus article will increase researchers' understanding of different analytical approaches (from generation of accelerometer data descriptors to the examination of health associations) that have been used in recent epidemiological studies of physical behaviours. Furthermore, this article and the decision tree provided aims to assist researchers with the selection of analytical approaches based on their research questions and available data. This will ultimately have an impact on the scientific evidence and, therefore, on future public health guidelines on physical behaviours. Additionally, short- and long-term agendas are defined to promote best practices for data analysis and reporting in 'physical behaviour epidemiology'.

SECTION II



Study VIII



Associations of objectively-assessed physical activity and sedentary time with hippocampal grey matter volume in children with overweight or obesity

Migueles JH, Cadenas-Sanchez C, Esteban-Cornejo I, Torres-Lopez LV, Aadland E, Chastin SF, Erickson KI, Catena A, Ortega FB.
J Clin Med. 2020 Apr;9(4):1080.
DOI:10.3390/jcm9041080

Contents

Abstract

Background | Previous studies have not specifically investigated the association between SB, PA, and hippocampal GMV in paediatric obesity. Children with overweight or obesity engage in more SB, perform less PA, and present poorer brain health.

Aims | This study investigated PA and SB in relation to hippocampal GMV in paediatric overweight/obesity.

Methods | Ninety-three children (10 ± 1 year) were classified as overweight, obesity type I, or type II-III. PA was assessed with non-dominant wrist accelerometers. GMV was acquired by MRI.

Results | Neither PA nor SB associated with GMV in the hippocampus in the whole sample ($p > 0.05$). However, we found some evidence of moderation by weight status ($p < 0.150$). MVPA (MVPA) positively associated with GMV in the right hippocampus in obesity type I ($B = 5.62$, $p = 0.017$), which remained when considering SB, LPA, and sleep using compositional data ($\gamma = 375.3$, $p = 0.04$). Compositional models also depicted a negative association of SB relative to the remaining behaviours with GMV in the right hippocampus in overweight ($\gamma = -1838.4$, $p = 0.038$). Reallocating 20 min/day of SB to MVPA was associated with 100 mm^3 GMV in the right hippocampus in obesity type I. Multivariate pattern analysis showed a negative-to-positive association pattern between PA of increasing intensity and GMV in the right hippocampus in obesity type II-III.

Conclusion | Our findings support that reducing SB and increasing MVPA are associated with greater GMV in the right hippocampus in paediatric overweight/obesity. Further studies should corroborate our findings.

Abbreviations in this page:
GMV: grey matter volume
LPA: light physical activity
MVPA: moderate-to-vigorous physical activity
PA: physical activity
SB: sedentary behaviour

Key Points

Question
Are PA and SB associated with GMV in the hippocampus in children with overweight or obesity?

Findings
Our findings indicate that PA and SB were not associated with GMV in the hippocampus in children with overweight or obesity. However, some evidence of moderation by weight status in the associations

Meaning
Reducing SB and engaging in more MVPA by 20 min/day was associated with presenting 100 mm^3 of GMV in the right hippocampus in children with obesity type I.

Introduction

Improving brain health during childhood is important to enhance brain development, achieve academic goals, and improve cognition [51]. Within the brain, the hippocampus is crucial for short- and long-term memory [56,57], being a determinant of academic success and cognition in children [54,365]. Furthermore, the hippocampus presents a high degree of plasticity [366,367] (i.e., its capacity to change and reorganize in response to internal and/or external influences) [368]. Among the processes related to this plasticity, neurogenesis and angiogenesis can stimulate changes in the GMV. GMV in the hippocampus can be amplified by a variety of lifestyle factors [58]; among them, aerobic exercise has been widely investigated [55,368,369]. Aerobic exercise consists in structured and organized PA sessions aimed to improve aerobic fitness. Aerobic fitness is the integrated ability of the organism systems to perform PA, and it is a powerful marker of health in children [23]. Otherwise, PA stands for any movement produced by skeletal muscles which increases the BMR [8].

Abbreviations in this page:
BMR: basal metabolic rate
GMV: grey matter volume
LPA: light physical activity
MVPA: moderate-to-vigorous physical activity
PA: physical activity
SB: sedentary behavior

Aerobic fitness is associated with GMV in the hippocampus of children [54,55,370], which makes PA a potential resource to target hippocampal GMV. However, associations of PA with GMV in the hippocampus are inconclusive [370,371]. Herting et al. used a whole-brain approach to test associations between self-reported PA and GMV in 34 male adolescents [371]. Higher self-reported PA was associated with greater GMV in the right pericalcarine, right cuneus, and left precuneus, but it was not associated with GMV in the hippocampus [371]. However, self-report measures of PA are limited because of their low accuracy and social desirability bias, especially in youth [32]. To overcome these limitations, Ruotsalainen et al. used accelerometers to assess PA but found no association with GMV in the hippocampus in 60 adolescents [370]. They reduced the accelerometer data into MVPA [370], while other PA intensities remain unstudied. Likewise, SB time, defined as awake time spent sitting or reclining with low energy requirement [19], has not been studied in relation to GMV in the hippocampus of children to the best of our knowledge.

Accelerometer-determined SB and PA data have certain features that should be considered. PA is usually monitored for seven days, for which the information is averaged to obtain daily estimates of SB, LPA, and MVPA together with sleep time [233]. This results in a set of interdependent (i.e., multicollinear) variables as they are constrained to 24 h (i.e., sleep + SB + LPA + MVPA = 24 h). In other words, increasing time in any of these behaviours would reduce the time in at least one of the others, a characteristic usually referred to as

Abbreviations in this page:
GMV: grey matter volume
PA: physical activity
SB: sedentary behavior

'closure' [326,327]. Multicollinearity and closure have not been appropriately handled in previous studies on the association between SB, PA, and GMV in the hippocampus of children [370,371]. Thus, studies using appropriate analytical approaches are needed to study the association between PA and SB with GMV in the hippocampus in children. The rate of hippocampal neurogenesis sharply declines during childhood and continues to decline during adulthood [372]. Therefore, it is crucial to find strategies to stimulate GMV in the hippocampus at young ages to ensure future healthy brains. Promoting PA is a promising strategy which needs further study.

Previous studies have not specifically investigated the association between SB, PA, and hippocampal GMV in pediatric obesity. Children with overweight or obesity engage in more SB [373], perform less PA [373], and present poorer brain health [374]. Thus, the study of the associations between PA and hippocampal GMV in children with overweight or obesity could provide meaningful information for public health messaging, as well as to appropriately design interventions targeting both physical and brain health in pediatric obesity. Therefore, this study aims to investigate associations of objectively measured SB and PA with GMV in the hippocampus using analytical approaches able to deal with the closed structure and strong multicollinearity of data obtained from accelerometry in children with overweight or obesity. Based on previous research on aerobic exercise [55,368,369], we hypothesized that lower SB and higher PA would associate with greater hippocampal GMV in children with overweight or obesity.

Material and methods

Participants and study design

We used baseline data from the ActiveBrains project (Identifier: NCT02295072) [63] collected from November 2014 to February 2016 in Granada (Spain). Initially, 110 children enrolled in the ActiveBrains project. Those with valid accelerometer and brain data at baseline were included in this cross-sectional analysis (n = 93, 10 ± 1 years of age, 37 girls). More information about the study can be found elsewhere [63]. Briefly, all participants met the inclusion criteria: 1) overweight or obesity based on the WOF cut-off points [64,65]; 2) 8–11 years old; 3) no physical disabilities or neurological disorders that affect physical performance; and 4) in the case of females, were not menstruating at the time of the baseline assessment.

Parents or legal guardians were informed of the purpose of the study and provided written informed consent. The ActiveBrains project was approved by the Ethics Committee on Human Research of the University of Granada.

Accelerometer data collection and processing

Accelerometer data collection and processing criteria are described elsewhere [266,302]. In brief, participants were required to wear accelerometers ActiGraph GT3X+ (ActiGraph, Pensacola, FL, USA) on their non-dominant wrist for 7 consecutive days, and to complete a sleep log with information on time to bed and time out of bed every day. Parents were suggested to supervise their children in the fulfillment of the sleep logs. Accelerometers were initialized to record accelerations at 100 Hz with a dynamic range of ±6 G. Raw accelerations were downloaded via the ActiLife v.6.13.3 software (ActiGraph, Pensacola, FL, USA) and processed in the R package GGIR (v.1.5.12) [67,200]. Non-wear time and abnormal high accelerations related to malfunctioning of the accelerometers were imputed by average acceleration during the same time interval from the rest of the days [67]. Sleep time was identified using an automated algorithm guided by the time reported by the participants [40,202]. Finally, SB (<35 m*g*) and intensity-specific PA (LPA: 35–200 m*g*; MVPA: >200 m*g*) were calculated using previously-proposed acceleration thresholds for the non-dominant wrist in children [61,62]. Additionally, the intensity spectrum was defined using time spent in 33 acceleration bands of increasing intensity by 25 mg (i.e., time spent in 0–25 m*g*, 25–50 m*g*, 50–75 m*g*, and so on). Only awake time was used to calculate the intensity spectrum variables since sleep and SB can occur at similar acceleration bands, which would confound the interpretation of findings. The average daily values of time spent in each category were calculated as: (weekdays x 5 + weekends x 2) / 7. The participants were excluded if they recorded less than 4 valid days (≥16 h/day), including at least 1 weekend day [233].

Abbreviations in this page:
BMI: body mass index
LPA: light physical activity
MVPA: moderate-to-vigorous physical activity
PA: physical activity
SB: sedentary behavior
SD: standard deviation

Table 26
Descriptive characteristics of participants.

	All (n = 93, 37 girls)	Overweight (n = 23, 9 girls)	Obesity I (n = 41, 15 girls)	Obesity II-III (n = 29, 13 girls)
Age (years)	10.01 (1.12)	10.13 (1.08)	10.29 (1.04)	9.51 (1.14)
Peak height velocity (years)	-2.31 (0.97)	-2.36 (1.04)	-2.1 (0.93)	-2.58 (0.91)
Weight (kg)	55.67 (10.69)	46.32 (7.30)	56.92 (9.63)	61.74 (9.31)
Height (cm)	143.95 (8.10)	142.16 (8.80)	146.59 (7.78)	141.84 (7.08)
BMI (kg/m^2)	26.74 (3.63)	22.64 (1.41)	26.26 (2.06)	30.68 (2.36)
Total brain volume (mm^3)	1202.1 (106.67)	1210.0 (99.41)	1221.0 (94.93)	1169.5 (122.50)
Parental university level, %				
Neither parent	68	57	59	90
One parent	16	17	22	7
Both parents	16	26	19	3
Grey matter volume				
Left hippocampus (mm^3)	3468.73 (371.48)	3387.17 (348.96)	3572.49 (346.4)	3386.71 (397.67)
Right hippocampus (mm^3)	3597.99 (382.49)	3568.46 (420.35)	3709.9 (354.77)	3463.19 (352.41)
Physical activity				
SB (min/day)	561.39 (60.85)	534.25 (71.12)	559.35 (50.59)	585.78 (57.52)
LPA (min/day)	275.36 (39.75)	277.85 (40.42)	273.16 (43.31)	276.49 (34.83)
MVPA (min/day)	54.61 (20.91)	61.76 (26.79)	53.84 (19.66)	50.05 (16)

Data are presented as mean (SD) or percentages.

Abbreviations in this page:
BMI: body mass index
FIRST: FMRIB's Integrated Registration and Segmentation Tool
FSL: FMRIB's Software Library
MPRAGE: magnetization-prepared rapid gradient-echo

Magnetic resonance imaging data acquisition and processing

All images were collected on a 3.0 Tesla Siemens Magnetom Tim Trio scanner (Siemens Medical Solutions, Erlangen, Germany) with a 32-channel head coil. High-resolution, T1-weighted images were acquired using a 3D MPRAGE (magnetization-prepared rapid gradient-echo) protocol. The acquisition parameters were the following: repetition time = 2300 ms; echo time = 3.1 ms; inversion time = 900 ms; flip angle = 9°; field of view = 256 × 256; acquisition matrix = 320 × 320, 208 slices; resolution = 0.8 × 0.8 × 0.8 mm; and scan duration = 6 min and 34 s.

Hippocampal volumetric analyses were conducted using FMRIB's Software Library (FSL) version 5.0.7. (FMRIB analysis group, Oxford, UK). Specifically, we used FMRIB's Integrated Registration and Segmentation Tool (FIRST) in FSL. FIRST is a semi-automated model-based subcortical segmentation tool which uses the Bayesian framework from shape and appearance models obtained from manually segmented images from the Centre for Morphometric Analysis, Massachusetts General Hospital (Boston, MA, USA) [375]. Briefly, FIRST runs a two-stage affine registration to a standard space template (i.e., Montreal Neurological Institute -MNI- space) using 12 degrees of freedom and uses a subcortical mask to exclude voxels outside the subcortical regions. Second, subcortical regions, including the hippocampus, are segmented for both hemispheres separately. Manual volumetric region labels are parameterized as surface meshes and modelled as a point distribution model. In addition, the hippocampus segmentation from FIRST was then split based on the centre of gravity of the region into anterior and posterior sub-regions for each hemisphere separately. This resulted in separate anterior and posterior hippocampal segmentation for each participant, for each hemisphere [376,377]. The final segmentations were visually inspected for quality. The volume of each region was obtained from FIRST in mm^3.

Confounders

Participants' body mass, height, peak height velocity, and parental education level were obtained as part of the protocol of the ActiveBrains project [63]. Weight and height were measured twice consecutively with an electronic scale (SECA 861, Hamburg, Germany) and a stadiometer (SECA 225, Hamburg, Germany), respectively, and averaged values were used in analyses. BMI was calculated as weight (kg) divided by squared height (m^2). Children were classified as having overweight, obesity type I, and obesity type II–III using the sex- and age-specific BMI cut-offs proposed by the WOF [64,65]. Peak height velocity was derived from standing or seated height as a continuous measure of maturational status using the Moore et al. equations: for boys, –8.13 + (0.007 × (age × seated

height)); for girls, –7.71 + (0.004 × (age × height)) [92]. Parents reported their highest completed level of education. Parental education level was categorized as both of them, one of them, or neither of them reaching university-level education. Total brain volume was derived from FreeSurfer software version 5.3.0 (Laboratory for Computational Neuroimaging, Athinoula A. Martinos Center for Biomedical Imaging, Harvard Medical School, Boston MA, USA) as the sum of total white matter volume and total GMV.

Abbreviations in this page:
ESM: electronic supplementary material
GMV: grey matter volume
LPA: light physical activity
MVPA: moderate-to-vigorous physical activity
PA: physical activity
SB: sedentary behavior

Statistics

Participants' descriptive characteristics were summarized as mean and SD or percentages. Bivariate correlations among PA and SB indicators and between these variables and GMV in the right and left hippocampi were performed. Then, associations of PA and SB (explanatory/independent variables) with GMV in the hippocampus (outcome/dependent variable) were analysed using different analytical approaches (i.e., multiple standard linear regression using absolute and compositional data and multivariate pattern analysis with absolute data). After testing the potential confounding effect on the associations, the same set of confounders was accounted for in all analyses (i.e., sex, peak height velocity, parental education level, and total brain volume). Interactions between weight status (i.e., overweight, obesity type I, or obesity type II–III) and PA were tested because of the moderator effect shown in previous studies [55,378]. Using multiple linear regression with absolute data, a moderation effect was found in the association of LPA and MVPA with GMV in the right hippocampus ($p < 0.15$). Thus, the analysis was stratified by obesity category. The analytical approaches were implemented as follows.

Multiple linear regression models using absolute PA and SB data were performed to compare associations with previously-published findings. Separate models were performed for each PA intensity and SB. Findings from these models should be interpreted as incrementing time spent in a behaviour in isolation (i.e., without considering the remaining behaviours).

Multiple linear regression with compositional data [326,327] was used to study the relative association of PA and SB with GMV in the hippocampus. Compositional data analysis accounts for the relative nature of physical behaviour by quantifying the effect of incrementing time in each behaviour by reducing the time spent in the rest (i.e., closure). Since time exchange can also occur with sleep time, detected sleep period time (i.e., time from going to bed to waking up) was included in compositional analyses. Isometric log-ratios were firstly calculated and then introduced in multiple linear regression models as previously proposed [326] (see **ESM 7** for a detailed explanation of the models). Gamma (γ) coefficients with their respective 95% interval inform of the strength and direction of the

ESM 7 (scan or click here)

Abbreviations in this page:
CI: confidence interval
ESM: electronic supplementary material
GMV: grey matter volume
LPA: light physical activity
MVPA: moderate-to-vigorous physical activity
PA: physical activity
SB: sedentary behavior

association. For an accurate estimation of the effect size, isotemporal substitution plots were computed to investigate the effect of increasing LPA and MVPA in the detriment of SB. Findings from compositional models should be interpreted as incrementing time spent in a behaviour relative to time spent in the remaining behaviours (or pair-wise time exchange between behaviours in the reallocation plots).

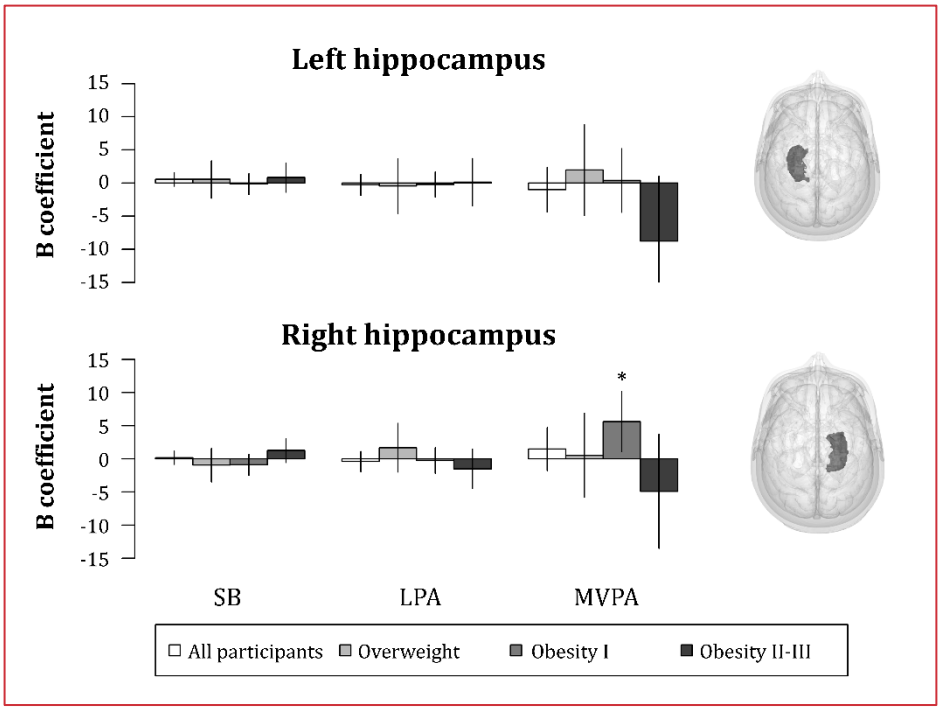


Figure 34
Regression non-standardized B coefficients and 95% CIs (i.e., error bars) for the association of SB, LPA, and MVPA with GMV in the hippocampus adjusted for sex, peak height velocity, parental university level, and total brain volume.
* $P < 0.05$

Multivariate pattern analysis with absolute PA and SB data was used to further understand the associations depicted by previous models. Partial least squares regression was performed since it can handle completely collinear variables through the use of latent modelling [311,353]. Models were cross-validated using Monte Carlo resampling [379] with 1000 repetitions by repeatedly and randomly keeping 50% of the subjects as an external validation set. For each validated partial least square regression model, a single predictive component was subsequently calculated through target projection [342,353] to express all the predictive variance in the PA intensity spectrum related to GMV in the hippocampus in a single intensity vector. Selectivity ratios with 95% confidence intervals were obtained as the ratio of this explained predictive variance to the total variance for each PA intensity variable [310]; see **ESM 7** for an in-depth description of selectivity ratio interpretation. Briefly, the selectivity ratio has a range of −1 to 1 and the negative or positive sign informs the direction of the association with the outcome. Associations from the partial least square regression should be interpreted as each intensity variable's importance for predicting the outcome, while simultaneously taking into account all intensity bands in one joint model. Thus, the model provides the total association pattern between PA intensity and hippocampal GMV.

ESM 7 (scan or click here)

All analyses were performed in R (v. 3.6.2), except for the multivariate pattern analysis, which was performed in Sirius v.11.0 (Pattern Recognition Systems AS, Norway).

Results

Participants' sociodemographic and anthropometric characteristics, hippocampal GMV, PA, and SB are reported in **Table 26**. Children spent around 39% of the day in SB, 19% in LPA, and 4% in MVPA, with the remaining 38% spent in bed. SB increased and MVPA decreased with more adverse weight status, while LPA was relatively constant across weight status groups. SB, LPA, and MVPA were correlated in this study sample (r ranging from 0.3 to 0.5, $p < 0.001$; **ESM 8, Table S1**).

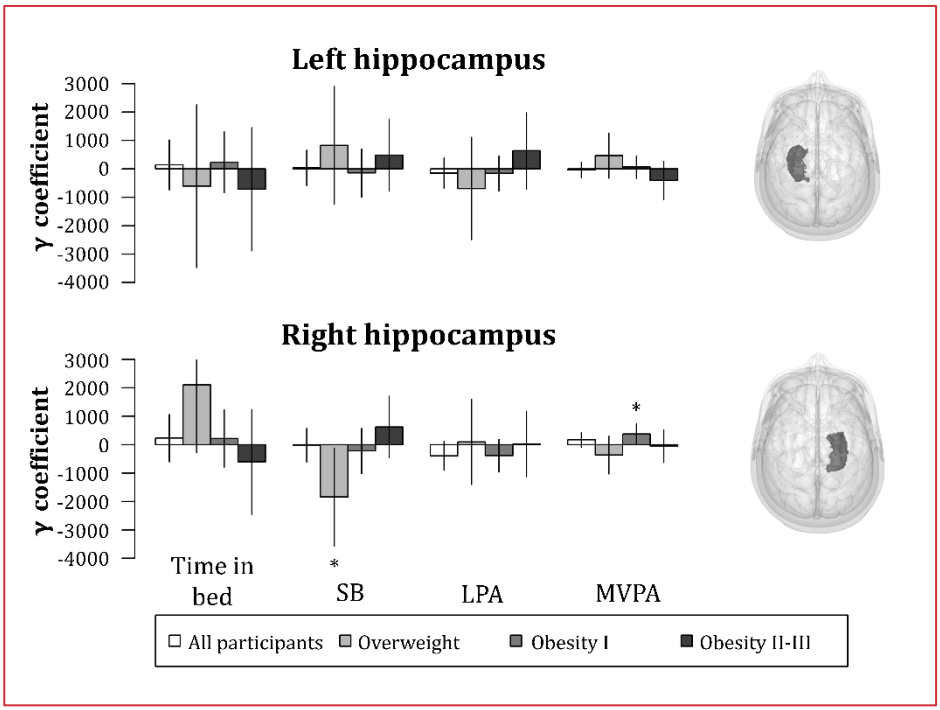


Figure 35
Compositional models γ coefficients and 95% confidence intervals (i.e., error bars) for the association of SB, LPA, and MVPA with GMV in the hippocampus adjusted for sex, peak height velocity, parental university level, and total brain volume.
Each bar represents the association of the pertinent behavior (e.g., SB) relative to the remaining behaviors (e.g., LPA, MVPA, and sleep) with GMV in the hippocampus.
* $P < 0.05$

Bivariate correlations of PA and SB with GMV in the left and right hippocampi stratified by weight status are presented in the supplementary material (**ESM 8, Table S2**). Non-standardized beta coefficients with their respective 95% confidence intervals from the multiple linear regression models with absolute PA and SB data are shown in **Figure 34**. Overall, neither SB nor PA were associated with GMV in the left or right hippocampi in the whole study sample (n = 93, $p > 0.05$). Separate analyses in weight status groups depicted that MVPA was positively associated with GMV in the right hippocampus in children with obesity type I (n = 41, $p = 0.017$).

Figure 35 shows γ coefficients from compositional models with their respective 95% confidence intervals. The γ coefficients represent the direction and strength of association between the isometric log-ratio (this is, the association of each behaviour relative to the remaining behaviours) and GMV in the left and right hippocampi. Consistent with the standard multiple regression models, SB and PA

Abbreviations in this page:
ESM: electronic supplementary material
GMV: grey matter volume
LPA: light physical activity
MVPA: moderate-to-vigorous physical activity
PA: physical activity
SB: sedentary behavior

ESM 8 (scan or click here)

Abbreviations in this page:
GMV: grey matter volume
LPA: light physical activity
MVPA: moderate-to-vigorous physical activity
PA: physical activity
SB: sedentary behavior

were not associated with either left or right hippocampi in the whole sample (n = 93, p > 0.05). The association of MVPA relative to SB, LPA, and sleep with GMV in the right hippocampus was significant in the sub-sample of children with obesity type I (n = 41, p = 0.040). Likewise, SB relative to LPA, MVPA, and sleep was negatively associated with GMV in the right hippocampus in children with overweight (n = 23, p = 0.038). MVPA was not associated with GMV in the sub-sample of children with obesity type II–III using compositional models.

The hypothetical effect of increasing either LPA or MVPA in the detriment of SB on GMV in the right hippocampus is presented in **Figure 36**. The subsample of children with obesity type I presented a significant positive effect of reallocating time from SB into MVPA on GMV in the right hippocampus. Since neither SB nor PA were associated with GMV in the left hippocampus (**Figure 35**), isotemporal reallocations were not depicted for this region.

Finally, a multivariate pattern analysis with partial least squares regression was performed to investigate the association of the absolute PA pattern with GMV in the hippocampus. Similar to previous analyses, the PA pattern was not associated with GMV in the hippocampus in the whole sample. Regarding the stratified analyses for weight status, we found that the absolute PA pattern was associated with GMV in the right hippocampus in those children with obesity type II–III (**Figure 37**). Negative selectivity ratios were found with low acceleration bands (representative of SB and LPA), while positive selectivity ratios were observed in high acceleration bands (indicators of MVPA). The most negative association was found in the acceleration band of 25–50 mg, which is an indicator of SB (selectivity ratio = –0.855, which means this band explains ~85% of the 30% explained by the latent components, i.e., ~25%), while the most positive was found in the 350–375 mg band, an indicator of MPA (selectivity ratio = 0.404). No associations were found in other weight groups with the left or the right hippocampi using multivariate pattern analysis.

Discussion

The main finding of this study was the lack of association between SB, LPA, and MVPA with hippocampal GMV in children with overweight or obesity. This lack of association persisted after performing the compositional data analysis and multivariate pattern analysis models, which take into account the relative nature and closure characteristics of the accelerometer-determined SB and PA data. SB, LPA, and MVPA were correlated in this study sample, which confirms our decisions on using analytical approaches to handle this co-dependency. Further studies using these analytical approaches will corroborate our findings. Nonetheless, we found that

associations were potentially moderated by weight status, which could be hiding any association in certain weight groups; thus, we performed separate analyses for weight status categories (i.e., overweight, obesity type I, and obesity type II–III). In this regard, we found a positive association of MVPA with GMV in the right hippocampus in children with obesity type I (using multiple regression with standard and compositional data) and in obesity type II–III (using multivariate pattern analysis). Likewise, we found that longer time in SB relative to LPA, MVPA, and sleep was associated with lower GMV in the right hippocampus in children with overweight (only in compositional data models). Otherwise, neither of the analyses performed depicted significant associations between PA or SB with GMV in the left hippocampus.

Abbreviations in this page:
GMV: grey matter volume
LPA: light physical activity
MVPA: moderate-to-vigorous physical activity
PA: physical activity
SB: sedentary behavior

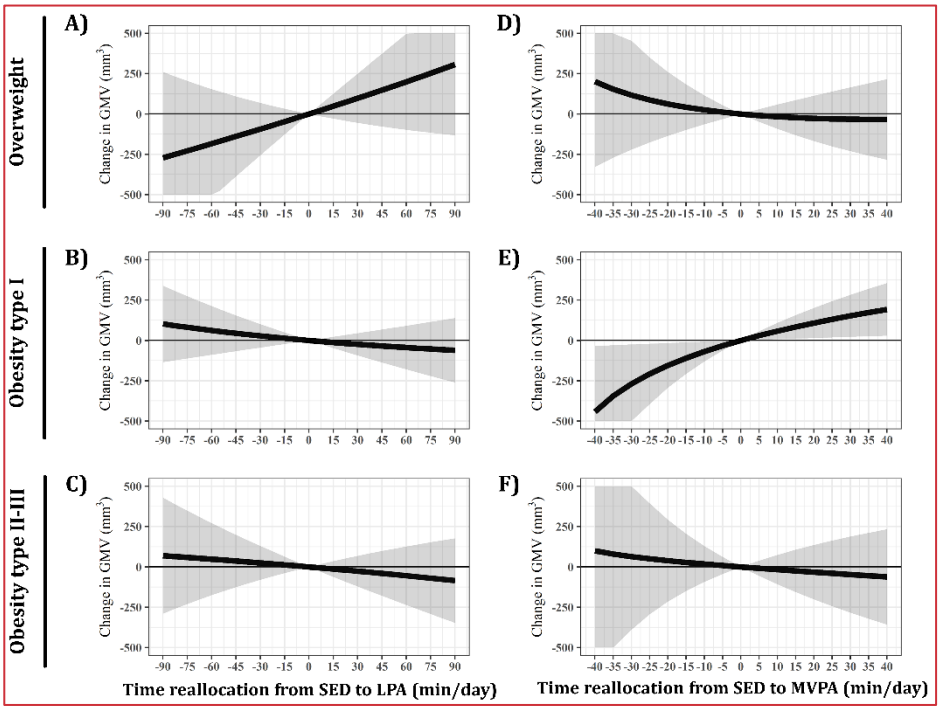


Figure 36
Effect of reallocating time from SB to LPA (Panels A, B, and C) and to MVPA (Panels D, E, and F) on the association with GMV in the right hippocampus using compositional models.

Relative to the moderation effect by weight status, it should be considered that our sample sizes in each subgroup are limited and these findings should be cautiously interpreted. We used the WOF categories [64,65] because: 1) they have been extensively related to both physical [380] and brain health [378]; and 2) these cut-off points were developed as sex- and age-specific in pediatric ages to connect at the age of 18 years with the adults BMI worldwide accepted cut-off points (i.e., 25 for overweight, 30 for obesity type I, and ≥35 for obesity type II–III). A previous study described a moderation effect of weight status on the acute effects of walking on memory in children [381]. Specifically, they found a single bout of walking to be effective in children with overweight or obesity to substantially improve word recognition memory performance, while it was not effective in children with normal weight [381]. The authors proposed circulating inflammatory markers to be tested as responsible for this moderation effect. In brief, obesity is characterized by an unhealthy

Abbreviations in this page:
ESM: electronic supplementary material
GMV: grey matter volume
LPA: light physical activity
MVPA: moderate-to-vigorous physical activity
PA: physical activity
SB: sedentary behavior

inflammatory response and PA has demonstrated higher anti-inflammatory and neuroprotective effects in obesity-induced brain inflammation [382–386]. Based on this, it would be expected that a larger association of PA with GMV in the hippocampus as the weight status is worse, but we did not find this linear trend. In this regard, further studies with larger sample sizes should deeply study this moderation effect with larger sample sizes.

Relative to our separate analyses for weight status groups, PA appears to be positively associated with GMV in the right hippocampus. Equally significant, we found that using appropriate analytical approaches to account for the data singularities of accelerometer-determined PA (i.e., closure and multicollinearity) is needed to elucidate the pattern of associations. In this regard, it appeared to be a negative non-adjusted association between MVPA and GMV in the left hippocampus in obesity type II–III (**ESM 8, Table S2**), which disappeared in compositional models and turned positive in multivariate pattern analysis. Since sample sizes were relatively small in our analyses, we could be under-powered to detect small-to-medium associations, so there is a risk of spurious associations in our findings. Therefore, further appropriately-powered studies should corroborate and extend our findings. The associations differed between analytical approaches. The compositional data analysis found an association with obesity type I and the multivariate pattern analysis with obesity type II–III. In this regard, compositional analysis is interpreted in terms of increasing a behaviour in exchange with others; thus, we found that increasing MVPA relative to decreases in SB, LPA, and sleep was positively associated with GMV in the right hippocampus in obesity type I. Otherwise, the multivariate pattern analysis is interpreted in terms of absolute changes in a certain behaviour fully considering multicollinearity among PA intensities. In this sense, we found that MVPA is positively associated with GMV in the right hippocampus in obesity type II–III. Our models with compositional data required the inclusion of three extra covariates to account for the relative nature of the data, which could imply that even larger sample sizes are needed to investigate associations with compositional models. The multivariate pattern analysis fully considers multicollinearity among PA variables and it is less affected by sample size, which could explain why the MVPA association was found using this approach in obesity type II–III even with its limited sample size. We do not have a large enough sample size to elucidate why associations differed across analytical approaches; thus, we suggest considering these associations with caution. Additionally, we call for further studies with larger sample sizes to apply these analytical approaches (i.e., compositional data analysis and multivariate pattern analysis) that

ESM 8 (scan or click here)

are more suitable than standard linear regression to accelerometer-determined PA and SB data.

Hippocampal plasticity across the lifespan has been previously confirmed [366,367]. However, the rate of hippocampal neurogenesis sharply declines during childhood and continues to decline during adulthood [372]. Thus, it is crucial to find strategies to stimulate hippocampal plasticity at young ages to ensure future healthy brains. Aerobic fitness is among the factors associated with hippocampal GMV [54,55,370], which makes PA a potential resource to target hippocampal GMV. In this study, we found lower SB and higher MVPA to be associated with the GMV in the right hippocampus in children with overweight, but no associations were found for the left hippocampus. Although both left and right hippocampi are related to episodic memory in humans, they have differential functions with the left being involved in verbal and linguistic memory and the right in non-verbal and visuospatial memory [387,388]. Hippocampal structures do not follow similar maturational trajectories [389]. It is plausible that the left and the right hippocampi show differential plasticity, especially in youths' developing brains, which could explain why we found associations only with the right hippocampus. Therefore, it seems that reducing SB and incrementing the time devoted to PA may be advised to stimulate higher GMV in the hippocampus in children with obesity type I. However, since we cannot conclude that a causal relationship exists, further RCTs that are appropriately powered to test the moderating role of weight status should be carried out.

Abbreviations in this page:
GMV: grey matter volume
MVPA: moderate-to-vigorous physical activity
PA: physical activity
RCT: randomized controlled trial
SB: sedentary behaviour

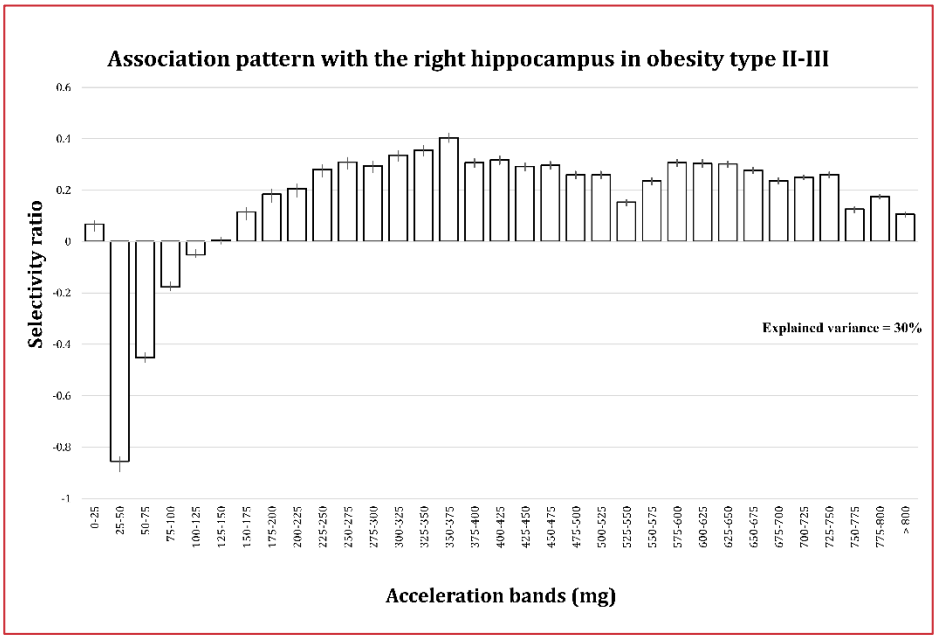


Figure 37
Association pattern of the PA spectrum with GMV in the right hippocampus in children with obesity type II–III.

We decided to focus our analyses on the hippocampus given that it is a brain region highly sensitive to PA in older populations [24,376]. Thus far, evidence in youths is limited with only two previous studies investigating the relationship of PA with GMV in

Abbreviations in this page:
GMV: grey matter volume
MVPA: moderate-to-vigorous physical activity
PA: physical activity
SB: sedentary behavior

adolescents [370,371]. None of these studies directly focused on the hippocampus, but rather used a whole-brain approach [371] or regional analyses [370] and did not find associations between PA and GMV in the hippocampus. These studies presented several limitations such as the use of self-reported tools to estimate PA [371], the reduction of PA data into one single variable (i.e., MVPA), hardly representative of the whole PA pattern [370,371], and the use of standard analytical approaches to test associations without considering the singularities of PA data (i.e., closure and multicollinearity) [370,371]. Furthermore, both studies had relatively small sample sizes (i.e., 34 and 60 participants) and were focused on adolescence, a period in which hippocampal neurogenesis might not be sensitive to external factors [372], such as PA or SB. This study overcomes previous limitations by using accelerometers to estimate PA, SB, and sleep. Considering the singularities of PA data with appropriate analytical approaches [311,326] applied in a sample of nearly a hundred children, our findings support the general public health recommendations on reducing SB and increasing PA to benefit brain health, specifically GMV in the right hippocampus. The overweight or obesity condition of our sample is important since these children usually have poorer physical and brain health profiles [374], thus, the study of potential interventions to improve their health status is a global need. In this regard, and similar to some previous studies [55,378], we found a potential moderator effect of weight status on the association between PA and GMV in the hippocampus that should be further corroborated with larger sample sizes. No less important is the fact that MVPA was associated in obesity type I using compositional analysis, and in obesity type II using multivariate pattern analysis. Although a moderation effect of weight status has been previously reported [55,378], it would be expected that the magnitude of the association increases as does the weight status [381]. The lack of this increasing size of the association could be partly explained because of our limited sample size (n = 23 and 29, respectively), which should be corroborated with further well-powered studies. Although previous studies failed at finding an association between PA and GMV in children [370,371], the positive association of aerobic fitness and GMV in several brain regions (including the hippocampus) has been widely reported in children [54,55,370]. Aerobic fitness could be an indicator of PA level since it is linearly associated with MVPA (standardized β around 0.3–0.4, $p < 0.01$ in this specific sample) [390]; however, the direct study of the behaviour (i.e., PA and SB) is important for public health for various reasons: first, aerobic fitness is partially explained by genetic factors, which are not modifiable by PA; second, although PA is effective at improving aerobic fitness, there could be other physiological changes related to health

behaviour (PA) but not to aerobic fitness; third, the interpretation and applicability of aerobic fitness to public health is not straightforward (i.e., the general population is not familiar with the interpretation of aerobic fitness values, nor with the strategies that should be followed to increase aerobic fitness); and fourth, in contrast, knowledge on how much time should be spent in certain activity types/intensities to improve brain health is more easily understandable by the general population. As an example, in our sub-sample of children with obesity type I, reallocating 20 min/day from SB to MVPA was associated with 100 mm^3 (3%) increase in grey matter in the right hippocampus.

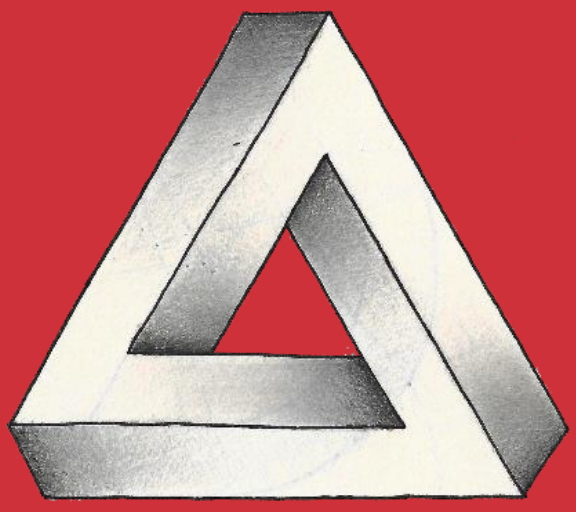
The main limitations of this study were: the cross-sectional design of the study, which does not allow causal interpretation of findings; although our study involves a larger sample size than previous studies, even more powerful studies are needed to confirm or contrast our findings; and sample sizes in the weight groups were asymmetric. We could have used the median split or terciles to match sample sizes but decided to use evidence-based and standard cut-offs. On the other hand, strengths of this study include: the use of accelerometers to objectively assess PA, SB, and sleep; the inclusion of sleep in compositional models to test its relative effect on the associations of PA and SB with GMV in the hippocampus; the use of MRI for the quantification of GMV in the hippocampus; the use of modern analytical approaches to analyse accelerometer data, which allows appropriate conclusions by handling the PA data singularities (i.e., closure and multicollinearity); and the focus on children with overweight or obesity, which is a harmful condition for both physical and brain health in children.

Abbreviations in this page:
GMV: grey matter volume
MVPA: moderate-to-vigorous physical activity
PA: physical activity
SB: sedentary behavior

Conclusions

Our findings indicate that PA and SB were not associated with GMV in the hippocampus in children with overweight or obesity. However, we found some evidence of moderation by weight status in the associations, so that reducing SB and engaging in more MVPA were associated with greater GMV in the right hippocampus. Specifically, reallocating 20 min/day from SB to MVPA would be associated with 100 mm^3 more GMV in the right hippocampus in children with obesity type I. We performed an in-depth data analysis by using compositional data and multivariate pattern analysis on accelerometer-determined PA data. These findings should be further confirmed by future studies.

Study IX



Associations of sleep with grey matter volume and their implications for academic achievement, executive function, and intelligence in children with overweight or obesity

Migueles JH, Cadenas-Sanchez C, Esteban-Cornejo I, Mora-Gonzalez J, Rodriguez-Ayllon M, Solis-Urra P, Erickson KI, Kramer AF, Hillman CH, Catena A, Ortega FB.
Pediatr Obes. 2020 Aug 13;e12707
DOI: 10.1111/ijpo.12707

Contents

Abstract

Background | Children with overweight/obesity have poorer sleep and smaller GMV than normal-weight children. No studies have investigated the associations of objectively-assessed sleep and GMV in children with overweight or obesity, or their implications for academic and cognitive outcomes.

Aims | To explore the associations of sleep behaviours with GMV in the whole brain and particularly the hippocampus as a region of interest independent of SB and PA; and to assess whether GMV in the associated regions was related to academic achievement, executive function and IQ.

Methods | Ninety-six children with overweight or obesity (10 ± 1 year) were included. Sleep behaviours were assessed with accelerometers. GMV was acquired by magnetic resonance imaging. Academic achievement, executive function and IQ were assessed with separate tests. Analyses were adjusted for sex, peak height velocity and parent education as well as SB and PA.

Results | Earlier wake time, less time in bed, WASO and WASO occurrences were associated with higher GMV in eight cortical brain regions (k:56-448, P's < .001). Longer total sleep time, higher sleep efficiency and less WASO time were associated with higher GMV in the right hippocampus (β:0.187-0.220, P's < .05). The inferior temporal, fusiform, supramarginal, and postcentral gyri, the superior parietal cortex, precuneus and hippocampus associated with academic achievement and/or IQ. Associations remained after adjustments for SB and PA.

Conclusion | Sleep behaviours are associated with GMV in multiple cortical regions including the right hippocampus in children with overweight or obesity, which in turn, were associated with academic achievement and IQ.

Abbreviations in this page:
GMV: grey matter volume
IQ: intelligence quotient
PA: physical activity
ROI: region-of-interest
SB: sedentary behaviour
WASO: wakening after sleep onset

Key Points

Question
Is sleep associated with GMV in children with overweight or obesity?

Findings
Sleep behaviours are associated with GMV in cortical and subcortical regions, including the hippocampus in children with overweight or obesity.

Meaning
These associations appeared to imply positive effects on children's academic achievement and IQ, but not on executive function.

Introduction

Childhood academic achievement and cognition are affected by insufficient sleep duration (i.e., <8h), a public health concern reported worldwide [391,392]. GMV is a measure of the volume of tissue in the brain region being examined. It represents all tissue properties contained in grey matter including vasculature, glial cells, and neuronal cell bodies. GMV contributes to the process of information in the brain. Further, greater GMV in the developing brain is positively associated with brain health outcomes such as academic achievement or cognition [54], being a crucial characteristic for children to success in school and in general life. Likewise, several studies have reported that sleep behaviours, i.e., those behaviours related to sleep that can be measured in free-living conditions, such as total sleep time or total bed time, are associated with academic achievement [393–396], executive function [397,398] and IQ [399]. The study of the sleep behaviours and their associations with GMV could provide insights into the mechanisms underlying these associations. To date, previous research suggests a positive link between sleep behaviours and GMV in several cortical regions in 14-year-old adolescents [400] and with the hippocampal GMV in 5-18-year-olds[401]. Special attention should be paid to the hippocampus, which is important for memory consolidation during sleeping [402,403].

Abbreviations in this page:
GMV: grey matter volume
IQ: intelligence quotient

Most of the previous findings on the relationship between sleep behaviours and academic achievement, executive function and IQ have focused on sleep onset and wake-up timing and total sleep time. In this regard, later sleep onset and wake-up times are associated with lower GMV in cortical regions [400] and poorer school grade average [395,400], as well as mathematics, reading and social sciences [396,404]. Total sleep time showed a small effect on school grade average [394]. Previous studies hypothesized impaired attention during school hours to explain the link between reduced sleep and poorer academic achievement [393,400]. Likewise, previous studies have found modest associations between later sleep onset, shorter total sleep time, and lower sleep quality with poorer executive function in adolescents [397] and young adults [398], i.e., cognitive flexibility, inhibition and working memory [405,406]. Lastly, sleeping more than 8 hours seems to be associated with higher IQ in male adolescents [399].

It is noteworthy that previous research derives timing and duration variables from self-reported sleep onset and wake-up times (i.e., asking participants what time they go to bed and wake on average) [400,401]. Self-report methods have been found to be influenced by subjective inaccuracies and social desirability [31,407,408]. As an example, total sleep time has been usually derived from the

Abbreviations in this page:
GMV: grey matter volume
IQ: intelligence quotient
PA: physical activity
ROI: region-of-interest
SB: sedentary behaviour

difference between sleep onset and wake-up times reported, which would indicate total bedtime rather than total sleep time. Thus, previous findings based on self-reported information should be complemented using objective assessments. Several algorithms have demonstrated that wrist-worn accelerometers can provide valid assessment of sleep behaviours [38,40,202]. Another important limitation of previous studies is the use of school grades as indicator of academic achievement [400], which are affected by teachers' subjectivity and inter-school variability [409]. The use of accelerometer-derived sleep behaviours and standardized tests for the measurement of academic achievement are needed to investigate sleep associations with academic achievement.

Children with overweight or obesity have been characterized as having poorer sleep behaviours [410] and smaller GMV than children with a normal weight [374]. Furthermore, these children are often involved longer in SB and insufficient PA [411], which may also be associated with poorer sleep behaviours [51,412]. SB and PA behaviours coexist with sleep in the 24-hour cycle. All these behaviours can potentially affect each other, since increasing the time spent in one would reduce the time devoted to another or others of the remaining behaviours. Therefore, when studying associations between sleep behaviours and GMV, it is crucial to study the potential influence of SB and PA behaviours on these associations. This scenario requires investigation on sleep behaviours and GMV, academic achievement, and cognition in children with overweight or obesity, as well as how SB and PA may influence these relationships.

Therefore, this study explored: 1) the associations of sleep behaviours with GMV in children with overweight or obesity using a whole-brain volumetric approach, as well as the specific associations between sleep behaviours and GMV in the hippocampus using a ROI approach, independently of SB and PA; and, 2) whether GMV in those regions associated with sleep behaviours are also related to academic achievement, executive function, and IQ. To our knowledge, there is not a clear hypothesis guiding which brain regions might underlie the association of sleep with academic achievement, executive function, or IQ. However, a main candidate could be the hippocampus [401–403]. Thus, we chose to use a whole-brain approach and further investigate the specific association with the hippocampus using a ROI approach. Based on previous research [393–398,413], we hypothesized that sleep behaviours would be associated with GMV, and that GMV in some of these regions would be associated with academic achievement, executive function and/or IQ.

Methods

Participants and study design

This study used baseline data from the ActiveBrains project (http://profith.ugr.es/activebrains). Of the 110 children who enrolled, 96 provided valid accelerometry and brain data at baseline (10 ± 1 years, 38 girls) and were included in this cross-sectional study. Since the ActiveBrains project is a RCT aimed to find effects of exercise on brain, cognition and academic achievement in children with overweight or obesity, the sample size was primarily calculated to detect medium changes in brain outcomes after the intervention [63]. The detailed rationale and inclusion criteria are described elsewhere [63]. Briefly, inclusion criteria included: 1) overweight or obesity based on the WOF cut-off points [64,65]; 2) eight to eleven years old; 3) no physical disabilities or neurological disorders that affect physical performance; and, 4) in the case of females, not to have started menstruation at the time of the baseline assessment. Data were collected from 2014 to 2016 in Granada (Spain) in three different waves evaluated during the months of November, December, January and February (always during school time in the three waves). Parents were informed of the purpose of the study and parental written informed consent was obtained. The ActiveBrains project was approved by the Ethics Committee on Human Research of the University of Granada.

Abbreviations in this page:
RCT: randomized controlled trial
WASO: wakening after sleep onset
WOF: World Obesity Federation

Sleep behaviours, sedentary behaviour and physical activity

Participants wore accelerometers ActiGraph GT3X+ (ActiGraph, Pensacola, FL, USA) on their non-dominant wrist for 7 consecutive days and reported information on time to go to bed and time to get off the bed every day. Raw accelerations were downloaded via the ActiLife v.6.13.3 software (ActiGraph, Pensacola, FL, USA) and processed in the R package GGIR [200] (v. 1.5.12, https://www.cran.r-project.org/). Detailed information on accelerometer data processing can be found elsewhere [266]. Identification of sleep onset and wake-up times were determined by an automatized algorithm guided by participants' reported times[40]. First, the algorithm examined potential sleep occurrences (i.e., at least 5 minutes with low variability in arm angle, i.e., <5 degrees) throughout the 24 hours. Next, the first and last epochs classified as sleep after and before the reported times were considered the definitive sleep onset and wake-up times. Finally, the algorithm developed by Sadeh et al. [38] was applied within the bedtime defined to classify every 1-min epoch as 'asleep' or 'awake'. Sleep behaviours included indicators of sleep timing (i.e., wake-up time and sleep onset), total time in bed, total sleep time and sleep patterns (i.e., sleep efficiency, wakening after sleep onset [WASO] time and number of WASO).

Abbreviations in this page:
BMI: body mass index
IQ: intelligence quotient
LPA: light physical activity
MPRAGE: magnetization-prepared rapid gradient-echo
MVPA: moderate-to-vigorous physical activity
SB: sedentary behaviour
SD: standard deviation
WASO: wakening after sleep

Total bedtime is the time difference between wake-up and sleep onset times. Total sleep time represents the sum of all minutes classified as sleep within total bedtime. Sleep efficiency is the percentage of time classified as sleep over the total bedtime. Cut points for the non-dominant wrist proposed by Hildebrand et al. [61,62] were used to classify SB, LPA and MVPA. A total of 104 met the pre-requisite of recording 4 valid days (i.e., ≥ 16 hours/day); including at least 3 weekdays and 1 weekend day were required. Specifically, to consider a day valid, participants should accumulate 2/3 of the waking hours and 2/3 of night hours as wear time and altogether accumulate at least 16 hours of wear time. Sleep-related variables, SB, LPA and MVPA daily values were averaged as follows: ((school-day average * 5) + (weekend day average * 2)) / 7.

Magnetic resonance imaging

All images were collected with a 3.0 Tesla Siemens Magnetom Tim Trio scanner (Siemens Medical Solutions, Erlangen, Germany) with a 32-channel head coil. High-resolution, T1-weighted images were acquired using a 3D MPRAGE protocol. The acquisition parameters were the following: repetition time = 2,300 ms; echo time = 3.1 ms; inversion time = 900 ms; flip angle = 9º; field of view = 256 x 256; acquisition matrix = 320 x 320, 208 slices; resolution = 0.8 x 0.8 x 0.8 mm; and scan duration = 6 min and 34 s.

Table 27
Descriptive characteristics of participants

	All (N=96)	Boys (N=58)	Girls (N=38)
Physical characteristics, mean (SD)			
Age (years)	10.02 ± 1.13	10.16 ± 1.15	9.79 ± 1.09
Peak height velocity (years)	-2.3 ± 0.96	-2.65 ± 0.79	-1.77 ± 0.96
BMI (z-Score)	3.04 ± 0.89	3.17 ± 0.99	2.82 ± 0.65
Parent education university level, %			
Neither parent	66	72	58
One parent	17	16	18
Both parents	17	12	24
PA, mean (SD)			
SB (min/day)	561.07 ± 60.09	553.46 ± 60.32	572.7 ± 58.62
LPA (min/day)	275.85 ± 39.3	271.39 ± 38.42	282.65 ± 40.17
MVPA (min/day)	54.94 ± 20.8	62.01 ± 22.05	43.86 ± 12.36
Sleep behaviours, mean (SD)			
Wake time (hh:mm)	08:07 ± 00:34	08:05 ± 00:33	08:10 ± 00:36
Sleep onset time (hh:mm)	23:02 ± 00:40	23:01 ± 00:41	23:04 ± 00:38
Total time in bed (min/day)	527.24 ± 31.87	526.14 ± 33.55	528.93 ± 29.49
Total sleep time (min/day)	457.78 ± 34.91	455.39 ± 32.42	461.42 ± 38.58
Sleep efficiency (%)	84.53 ± 4.92	84.01 ± 4.39	85.34 ± 5.6
WASO time (min/day)	77.14 ± 23.8	79.92 ± 20.51	72.89 ± 27.85
Number of WASO (nr.)	23.4 ± 4.43	24.04 ± 4.08	22.42 ± 4.81
Valid days (nr.)	6.95 ± 0.4	6.97 ± 0.49	6.92 ± 0.35
Academic achievement, mean (SD)			
Reading (standard score)	108.17 ± 13.13	108.31 ± 11.17	107.95 ± 15.83
Mathematics (standard score)	101.8 ± 10.79	102.45 ± 11.42	100.82 ± 9.81
Writing (standard score)	113.6 ± 12.86	112.66 ± 12.02	115.05 ± 14.1
Academic skills (standard score)	118.66 ± 16.18	117.76 ± 14.77	120.03 ± 18.26
Academic fluency (standard score)	103.56 ± 11.92	104.03 ± 10.7	102.84 ± 13.69
Academic applications (standard score)	99.31 ± 9.12	99.81 ± 9.18	98.55 ± 9.1
Total achievement (standard score)	109.14 ± 11.98	109.03 ± 10.85	109.29 ± 13.67
Executive function, mean (SD)			
Cognitive flexibility (z-Score)	-0.03 ± 0.81	0.08 ± 0.82	-0.2 ± 0.77
Inhibition (s)	41.9 ± 17.31	38.97 ± 15.31	46.38 ± 19.34
Working memory (nr. correct answers)	65.54 ± 16.44	67 ± 16.46	63.31 ± 16.39
IQ, mean (SD)	98.45 ± 12.34	97.02 ± 12.14	100.63 ± 12.48

Data are presented as mean ± SD. Statistically significant values are shown in bold.

Whole-brain volumetric analyses were conducted using the Statistical Parametric Mapping software (SPM12; Wellcome Department of Cognitive Neurology, London, UK) implemented in Matlab (The MathWorks, Inc, Natick, MA). Imaging pre-processing included quality control, motion correction, spatial normalization to an MNI template, and spatial smoothing. Detailed information about pre-processing steps is described elsewhere[54]. Hippocampal volumetric analyses were conducted using FSL version 5.0.7. Specifically, we used FIRST in FSL. FIRST is a semi-automated model-based subcortical segmentation tool which uses the Bayesian framework from shape and appearance models obtained from manually segmented images from the Center for Morphometric Analysis, Massachusetts General Hospital, Boston, MA, USA [375]. Briefly, FIRST runs a two-stage affine registration to a standard space template (i.e., MNI space) using 12 degrees of freedom and uses a subcortical mask to exclude voxels outside subcortical regions. Second, the subcortical regions, including the hippocampus, are segmented for both hemispheres separately. The manual volumetric region labels are parameterized as surface meshes and modelled as a point distribution model. In addition, the hippocampus segmentation from FIRST was then split based on the centre of gravity of the region into anterior and posterior sub-regions for each hemisphere separately. This resulted in separate anterior and posterior hippocampal segmentations for each participant [376,377]. The final segmentations were visually inspected for quality. The volume of each region was obtained from FIRST in mm^3.

Abbreviations in this page:
DFT: design fluency test
FSL: FMRIB's Software Library
FIRST: FMRIB's Integrated Registration and Segmentation Tool
MNI: Montreal Neurological Institute
TMT: trail making test

Academic achievement, executive function and intelligence quotient

Academic achievement was assessed with the Spanish version of the Woodcock-Johnson III battery, which is a valid and reliable (internal consistency reliability coefficient > 0.9) measure of academic achievement in children [71]. Children completed a total of 12 tests from this battery including reading, language, mathematics and sciences during one session of 100-120 min. Tests were independently checked by two trained evaluators and then scores were processed in the Compuscore and profile software (v. 3.1., Riverside Publishing Company, Itasca, IL, USA). We used standardized scores of broad reading, mathematics and writing scores, as well as composite scores representing academic skills (answers accuracy), academic fluency (processing speed), academic applications (problem solving) and total academic achievement score.

Executive function domains included cognitive flexibility, inhibition and working memory as described elsewhere [266,337]. Cognitive flexibility was assessed with the second and fourth conditions of the DFT and the third and fourth conditions of the TMT [72,414]. Both the DFT and the TMT are valid and reliable for measuring

Abbreviations in this page:
BMI: body mass index
DNMS: Delayed non-match-to-sample
GMV: grey matter volume
IQ: intelligence quotient
K-BIT: Kaufman brief intelligence test

cognitive flexibility in children [72,73]. The score from these tests was standardized by sex as Z-score and then averaged to obtain a unique indicator of cognitive flexibility. The Stroop test [415] was used as a valid and reliable indicator of inhibition [72–75]. Performance time of condition 3 (i.e., inhibiting reading by naming color) minus condition 1 (i.e., color naming) was used as previously reported [266,337]. Finally, working memory was measured from a modified version of the DNMS computerized task which has been validated to assess working memory [76]. In brief, a total of 16 practice trials plus 140 experimental trials were presented in two separated conditions (i.e., low- and high-memory load). Response accuracy for the high-load condition was used [309].

The IQ quotient was assessed with the Spanish version of the K-BIT [77]. The K-BIT shows a coefficient α for the validity of 0.86 to 0.93 in its Spanish version[77]. Crystallized and fluid IQ components were assessed with vocabulary and matrices sub-tests from K-BIT, respectively. Both sub-test scores were summed to obtain the IQ score.

Confounders

Participants' weight, height, peak height velocity and parents' education level were obtained as part of the protocol of the Active-Brains project [63]. Weight and height were measured twice consecutively with an electronic scale (SECA 861, Hamburg, Germany) and a stadiometer (SECA 225, Hamburg, Germany), respectively, and average values were used in analyses. BMI was calculated as kg/m^2. Peak height velocity was derived from standing and sitting height as a continuous measure of maturational status [92]. Parents reported whether both of them, one of them or none of them reached university level education. Total brain volume was derived from FreeSurfer software version 5.3.0 (http://surfer.nmr.mgh.harvard.edu) as the sum of total white matter volume and total GMV.

Statistics

Participants' descriptive characteristics were summarized as mean and SD or percentages. All variables were checked for normality. Included and excluded participants did not significantly differ in sociodemographic and anthropometric variables (all $P>0.05$). Based on previous studies[400,401], we tested sex, peak height velocity, parent university education level and total brain volume as confounders in sensitivity analyses. As all models remained similar with and without adjustment for total brain volume, we excluded it from the covariates. Thus, the association between sleep behaviours (i.e., sleep onset, wake-up time, total bedtime, total sleep time, sleep efficiency, WASO time and WASO number) and GMV was analysed using whole-brain voxel-wise multiple regression models, adjusted for sex, peak height velocity and parent university education level (i.e., basic

confounders). Sensitivity analyses were performed adding BMI as confounder to the previous model and all significant associations presented in this study remained unchanged (data not shown). Additionally, we extracted the eigenvalues from the peak coordinates of each significant cluster. The associations of the extracted mean GMV from significant clusters and academic achievement, executive function and IQ were studied with linear regression models adjusted for basic confounders. The Benjamini-Hochberg procedure was applied to account for the random effect in multiple comparisons for every dependent domain (i.e., academic achievement, executive function and IQ) with q=0.1. Then, we performed additional independent models adding either SB, LPA or MVPA as confounders (they were not all included in the same model because of the high risk of multicollinearity among these variables).

The statistical threshold in the imaging analyses was calculated with AlphaSim, as implemented in Resting-State fMRI Data Analysis Toolkit toolbox (RESTplus) [416]. Parameters were defined as follows: cluster connection radius (rmm)=5mm and the actual smoothness of the data after model estimation, incorporating a grey mask volume of 128190 voxels. The voxel-level alpha significance (threshold, $p < 0.001$ uncorrected) along with the appropriate cluster size for controlling for multiple comparisons in each analysis were indicated in the results. The resulting cluster extents were further adjusted to account for the non-isotropic smoothness of structural images[417].

Multiple linear regression models were used to study the associations between sleep behaviours and ROI hippocampal GMV adjusted for basic confounders. Additional models adjusting for SB, LPA or MVPA were performed. All statistical analyses were performed in R (v. 3.4.4, https://cran.r-project.org/), except those involving imaging data which were performed using the GLM approach implemented in SPM12.

Abbreviations in this page:
BMI: body mass index
IQ: intelligence quotient
LPA: light physical activity
MVPA: moderate-to-vigorous physical activity
ROI: region-of-interest
SB: sedentary behaviour
WASO: wakening after sleep onset

Results

Sociodemographic and anthropometric characteristics, sleep behaviours, academic achievement, executive function and IQ scores of participants are reported in **Table 27.**

Whole-brain associations of sleep behaviours with grey matter volume

Table 28 presents the sleep behaviours inversely associated (no positive associations were found) with GMV in the whole-brain volumetric analyses adjusted for sex, peak height velocity, parent education level, and additionally for SB and PA. A later wake-up time was associated with lower GMV in 2 bilateral clusters in the inferior temporal gyrus (Left: peak t=3.58, *k*=186; Right: peak t=3.76, k=243), and 3 more clusters in the fusiform gyrus (peak t=3.76, *k*=138), the

Abbreviations in this page:
GMV: grey matter volume
LPA: light physical activity
MVPA: moderate-to-vigorous physical activity
SB: sedentary behaviour
WASO: wakening after sleep onset

supramarginal gyrus (peak t=4.00, *k*=412) and the superior parietal cortex (peak t=4.12, *k*=56). A longer total bedtime was associated with lower GMV in the postcentral gyrus (peak t=4.22, *k*=257), but no association with total sleep time was found. A longer WASO time was associated with lower GMV in the superior temporal pole (peak t=3.61, *k*=98), the precuneus (peak t=3.67, *k*=400) and the superior parietal cortex (peak t=3.61, *k*=150). The number of WASO was inversely associated with GMV in two bilateral clusters in the postcentral gyrus (Right: peak t=3.96, *k*=418; Left: peak t=3.67, *k*=117) and in two more clusters in the medial superior frontal gyrus (peak t=3.73, *k*=448) and the superior parietal cortex (peak t=3.63, *k*=125). All these clusters showed P < 0.001 and they remained significant after additional adjustments for SB, LPA or MVPA. Associated clusters are visually presented in **Figure 38**. **Table S1** (**ESM 9**) shows bivariate correlation coefficients and confidence intervals between sleep behaviours.

Table 28
Brain regions showing significant negative associations of sleep behaviours with GMV (n = 96)

						Basic confounders		Basic conf. + SB		Basic conf. + LPA		Basic conf. + MVPA	
Predictors	Brain regions	X	Y	Z	Hemisphere	Peak t	Cluster size	Peak t	Cluster size	Peak t	Cluster size	Peak t	Cluster size
Wake time (hh:mm)	Inferior temporal gyrus	-41	-12	-44	Left	3.58	186	3.54	164	3.56	164	3.55	164
		47	-21	-35	Right	3.76	243	3.73	218	3.71	210	3.75	222
	Fusiform gyrus	-30	-47	-18	Left	3.76	138	3.78	144	3.79	149	3.80	152
	Supramarginal gyrus	-48	-50	33	Left	4.00	412	3.96	375	3.94	404	3.96	402
	Superior parietal cortex	29	-75	57	Right	4.12	56	4.08	52	4.09	50	4.10	52
Total time in bed (min/day)	Postcentral gyrus	-33	-42	56	Left	4.22	257	4.60	420	4.41	335	4.18	267
WASO time (min/day)	Superior temporal pole	29	18	-29	Right	3.61	98	3.67	132	3.59	89	3.56	86
	Precuneus	-2	-39	48	Left	3.67	400	3.65	400	3.64	269	3.68	478
	Superior parietal cortex	18	-60	71	Right	3.61	150	3.57	145	3.51	79	3.59	122
Number of WASO (nr.)	Medial superior frontal gyrus	12	45	42	Right	3.73	448	3.69	402	3.76	287	3.69	422
	Postcentral gyrus	63	-12	23	Right	3.96	418	3.90	372	3.75	283	3.96	431
		-56	-39	53	Left	3.67	117	3.62	102	3.56	63	3.64	109
	Superior parietal cortex	-20	-75	54	Left	3.63	125	3.61	124	3.77	161	3.61	114

Whole-brain voxel-wise multiple regression models were used. Basic confounders are sex, peak height velocity (years) and parent education university level (neither/one/both). All contrasts were thresholded using AlphaSim at P<0.001 with k=47 for wake-up time, k=57 for bedtime, k=46 for WASO time, k=55 for number of WASO for the basic confounders model, and remained similar for the rest of models, and surpassed Hayasaka correction. Anatomical coordinates (X, Y, Z) are given in Montreal Neurological Institute (MNI) Atlas space.
No clusters were significatively associated with sleep onset time, sleep time or sleep efficiency

Sleep behaviours with grey matter volume in the hippocampus

Figure 39 depicts scatter plots for the associations between total sleep time, sleep efficiency and WASO time with GMV in the left and right hippocampus. Associations with the right hippocampus were positive for total sleep time (β=0.187, P=0.049) and sleep efficiency (β=0.220, P=0.019) and negative for WASO time (β=-0.202, P=0.033). Specific associations for the anterior and posterior hippocampal sub-regions can be found as **ESM 9** (**Table S3**).

ESM 9 (scan or click here)

Grey matter volume associations with academic achievement, executive function and intelligence quotient

Higher GMV in those clusters related to wake-up time were associated with higher academic achievement scores. Specifically, four

out of the five clusters were associated with one or more academic achievement indicators (β ranging from 0.217 to 0.333, all P<0.028); and one cluster (i.e., supramarginal gyrus) was also associated with IQ (**Table 29**). GMV in the cluster related to total bedtime (left post-central gyrus) was associated with various academic achievement indicators, i.e., reading, academic skills, academic fluency, academic applications and total achievement (β ranging from 0.225 to 0.355, P's<0.032) (**Table 29**). Additionally, clusters in the precuneus and the superior parietal cortex, which were previously associated with WASO time, were also associated with academic achievement indicators (β ranging from 0.232 to 0.309, P's<0.028) (**Table 30**). The remaining clusters were not associated with academic achievement, executive function or IQ (all P>0.05).

Abbreviations in this page:
GMV: grey matter volume
IQ: intelligence quotient
LPA: light physical activity
MNI: Montreal Neurological Institute
MVPA: moderate-to-vigorous physical activity
ROI: region-of-interest
SB: sedentary behaviour
WASO: wakening after sleep onset

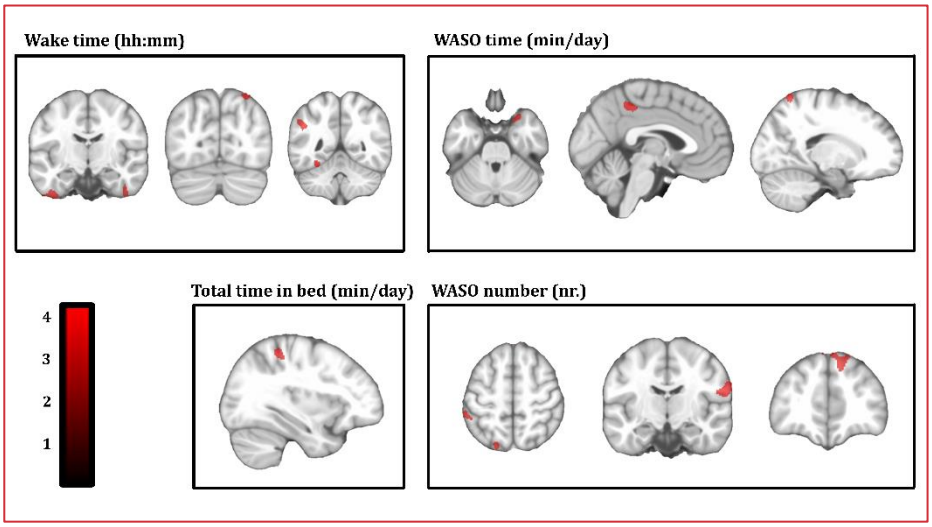


Figure 38
Brain regions showing negative associations of sleep behaviors with GMV in children with overweight/obesity.
Analyses were adjusted for sex, peak height velocity and parent education university level. All contrasts were thresholded using AlphaSim at P<0.001 with k=47 for wake-up time, k=57 for bedtime, k=46 for WASO time, k=55 for number of WASO, and surpassed Hayasaka correction. Anatomical coordinates (X, Y, Z) are given in MNI Atlas space.
The color bar represents t-values scale.

Finally, GMV in right hippocampus was not associated with academic achievement, executive function and IQ (**Table 31**). We only found a negative association with inhibition (β = -0.222 to 0.271, P = 0.02) surpassing the correction for multiple comparisons.

Discussion

Our findings support an association between some sleep behaviours (i.e., sleep timing, duration and pattern) and GMV in the cortical and subcortical brain structures, including the hippocampus, in children with overweight or obesity. Specifically, earlier wake-up time, shorter total bedtime, lower WASO time and the number of WASO were associated with greater GMV in one or several brain structures. Additionally, ROI analyses in the hippocampus depicted associations of a longer total sleep time, higher sleep efficiency and a lower WASO time with the GMV in the right hippocampus. All the associations mentioned are adjusted for sex, peak height velocity and parent education university level. Nearly every association remained following adjustment for SB, LPA or MVPA. The identified brain structures associated with sleep behaviours were also positively associated with academic achievement and, to a lesser extent, with IQ

Abbreviations in this page:
GMV: grey matter volume
ROI: region-of-interest
WASO: wakening after sleep onset

(but not with executive function). These findings should be interpreted with caution because of the multiple tests performed. Further studies should investigate the brain regions reported in this study.

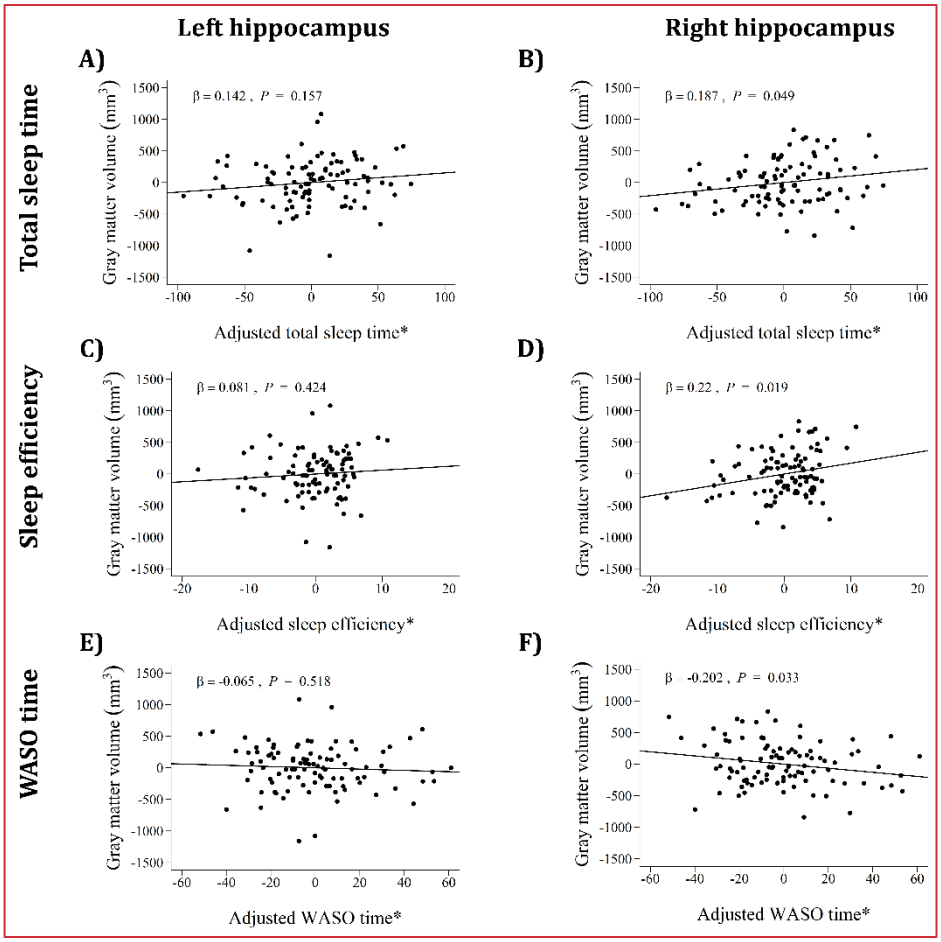


Figure 39
Linear regression analyses of the association of total sleep time (A, B), sleep efficiency (C, D) and WASO time (E, F) with GMV in the left and the right hippocampus.
*Analyses were adjusted for sex, peak height velocity and parent education university level.

Whole-brain associations of sleep behaviours with grey matter volume

Only two previous studies investigated the associations between sleep behaviours and GMV in children and/or adolescents using a whole-brain volumetric approach [400,401]. The whole-brain volumetric approach depicted a complete picture of sleep and GMV associations at a whole-brain level, rather than at a ROI level. Thus, the whole-brain approach affords the exploration of associations not described in previous studies. These two studies focused separately on weekdays and weekend days. This study was the first one investigating the week as a whole (using weighted averages to account for the correspondent weight of school days and weekend days in the daily life) and, therefore, considering the associations with brain as a result of the sleep behaviours during both weekdays and weekend days.

In regard to sleep timing, later wake-up times were associated with lower GMV in several cortical structures, such as the inferior temporal, the fusiform, the supramarginal gyri and the superior parietal cortex. Sleep onset was not associated with GMV in any brain region. The inferior temporal gyrus, the fusiform gyrus and the supramarginal gyrus are important for reading and language

processing, word recognition and posture and gesture identification [418,419]. GMV in these structures was positively associated with academic fluency, mathematics, total achievement and IQ score. The superior parietal cortex is important for processing spatiotemporal and visual information [420], which seems to be important for academic achievement as its GMV associated positively with mathematics, academic fluency, academic applications and total achievement. One previous study found that self-reported earlier wake-up time during weekends was associated with greater GMV in the medial frontal orbital and the anterior cingulate cortices in 14 year-olds [400]. Diversity in the associated brain regions can respond to a different stage in brain development (pre-adolescence vs. adolescence) and/or different methods of assessing sleep (i.e., self-reported vs. objective). Consistent with our findings, Urrila et al. found a positive association between GMV and school grades [400], which suggests that larger GMV was linked to better academic achievement. This study complements the previous one by adding a detailed study of the associations with different academic abilities and by including executive function and IQ.

Abbreviations in this page:
GMV: grey matter volume
IQ: intelligence quotient

Table 29
Associations of GMV from regions associated with wake-up time and total time in bed with academic achievement, executive function and IQ (N=96).

	Wake time					Total time in bed
	L Inferior temporal gyrus	R Inferior temporal gyrus	L Fusiform gyrus	L Supramarginal gyrus	R Superior parietal cortex	L Postcentral gyrus
Academic achievement						
Reading	0.121	-0.012	0.160	0.104	0.174	**0.355****
Mathematics	0.160	0.166	**0.224***	**0.241***	**0.333****	0.216*
Writing	0.149	0.027	0.200	0.170	0.207*	0.163
Academic skills	0.097	0.006	0.150	0.120	0.152	**0.225***
Academic fluency	**0.256***	0.126	**0.231***	**0.264***	**0.294****	**0.244***
Academic applications	0.136	0.070	0.181	0.155	**0.276****	**0.281****
Total achievement	0.171	0.058	**0.217***	0.195	**0.260****	**0.296****
Executive function						
Cognitive flexibility	0.184	0.020	0.191*	0.126	-0.001	0.075
Inhibition	-0.04	-0.057	0.051	-0.192	-0.146	0.185
Working memory	-0.03	0.069	0.065	0.113	0.152	0.210*
IQ	0.105	0.045	0.183	**0.269***	0.102	0.209*

Multiple linear regression models adjusted for sex, peak height velocity (years) and parent education university level (neither/one/both).
Bolded font indicates that the specific association surpassed the Benjamini-Hochberg correction for multiple comparison tests (performed for each domain, i.e., academic achievement, executive function and IQ).
* $P < 0.05$
** $P < 0.01$

Likewise, total bedtime associated with higher GMV in the postcentral gyrus. Considering the negative association with wake-up time and the lack of association with sleep onset, it seems logical that longer total bedtimes are due to later wake-up times, which are both negatively associated with GMV in several brain regions. None of the previous studies [400,401] have found associations with GMV in the postcentral gyrus, but differing sample characteristics and methodological inconsistencies may account for this discrepancy. The postcentral gyrus is located in the primary somatosensory cortex. Our findings may suggest that those children who stay longer in bed

Abbreviations in this page:
ESM: electronic supplementary material
GMV: grey matter volume
WASO: wakening after sleep onset

interact less with the environment and, in turn, do not stimulate this brain region adequately. This positively affects academic achievement, since GMV in this specific cluster was positively associated with reading, academic skills, academic fluency, academic applications and total achievement. However, this should not be interpreted as a negative consequence of longer sleep periods. Total sleep time was not associated with GMV in any region, but total bedtime was. It may indicate that those children who stayed in bed longer (especially when engaged in non-sleeping time activities) have smaller GMV in certain regions, which may lead to poorer academic achievement.

In regard to sleep patterns, longer WASO time was related to lower GMV in the superior temporal pole, the precuneus and the right superior parietal cortex; and a higher number of WASO with lower GMV in the superior medial frontal and the postcentral gyri, and the left superior parietal cortex. GMV in the precuneus and the right superior parietal cortex were positively associated with reading, mathematics, academic skills, fluency, applications and total achievement. Consistently, Urrila et al. also found sleep behaviours associated with GMV in the precuneus and, in turn, this GMV was associated with school grade average[400]. The precuneus has been related to visuospatial perception and, together with the hippocampus, to episodic memory[421,422]. These functions seem to be important for academic achievement. It is also noteworthy that the right superior parietal cortex was negatively associated with wake time and WASO time in our sample. Since these two sleep behaviours were not correlated with each other (see **ESM 9**, **Table S1**), we cannot assume that these findings are overlapping. Likewise, the left postcentral gyrus was associated with bedtime and the number of WASO, indicators which did not correlate with each other. Only the cluster related to bedtime was associated with academic achievement.

ESM 9 (scan or click here)

Briefly, these findings have complemented previous literature by confirming some specific areas which were found associated with sleep behaviours and by describing associations which have not been previously reported. Brain development during childhood and adolescence is heavily dependent upon the age of participants, with age having a differential relationship relative to which brain areas are more or less sensitive to sleep behaviours[372]. In this regard, our sample comprised 8-11-year-old children, while previous studies were focused on older adolescents. Furthermore, all of the participants in the current study presented overweight or obesity, which could alter the relationship between sleep and brain outcomes, including development[374], and may account for differences in the findings between the current study and previous research. Likewise, we found eight cortical regions associated with sleep, which resulted in a high number of statistical tests that were employed to study their

association with academic achievement, executive function and IQ. Although we applied a correction for multiple comparisons, we advise caution in the interpretations and suggest further investigation of these findings.

All of the above-mentioned associations were not affected by including SB; LPA or MVPA in the models. Therefore, SB, LPA and MVPA do not appear to influence the association between sleep behaviours and GMV in our sample of children with overweight or obesity.

Abbreviations in this page:
ESM: electronic supplementary material
GMV: grey matter volume
IQ: intelligence quotient
LPA: light physical activity
MVPA: moderate-to-vigorous physical activity
ROI: region-of-interest
SB: sedentary behaviour
WASO: wakening after sleep onset

Sleep behaviours with grey matter volume in the hippocampus

The hippocampus is a subcortical brain structure which, in constant communication with cortical structures, has been found to be crucial for memory consolidation during sleeping [402,403]. A major hypothesis on this link points out the memory consolidation process, which occurs predominantly during sleep [403]. Our ROI analyses of the hippocampus showed that total sleep time was positively associated with GMV in the right hippocampus. This finding agrees with Taki et al. [401], who found similar associations using self-reported sleep behaviours in 5-18-year-old children. Specifically, they found a longer total sleep time during weekdays to be associated with the hippocampus, but not weekend total sleep time. This study compliments their findings by objectively assessing a representative whole week, including both weekdays and weekends.

Table 30
Associations of GMV from regions associated with WASO time and number of WASO with academic achievement, executive function and intelligence (N=96).

	WASO time			Number of WASO			
	R Superior temporal pole	L Precuneus	R Superior parietal cortex	R Medial superior frontal gyrus	R Postcentral gyrus	L Postcentral gyrus	L Superior parietal cortex
Academic achievement							
Reading	0.122	**0.249***	**0.252***	0.138	0.216*	0.031	0.109
Mathematics	0.193	**0.249***	**0.309****	0.153	0.181	0.001	0.069
Writing	0.091	0.219*	0.079	-0.013	-0.031	-0.100	0.022
Academic skills	0.169	**0.241***	0.194	0.044	0.163	0.009	0.100
Academic fluency	0.050	**0.249***	0.193	0.146	0.137	-0.057	0.088
Academic applications	0.125	0.170	**0.232***	0.104	0.038	-0.065	-0.021
Total achievement	0.163	**0.274****	**0.248***	0.098	0.147	-0.028	0.074
Executive function							
Cognitive flexibility	0.009	-0.065	-0.065	-0.021	-0.104	-0.114	-0.123
Inhibition	0.037	-0.069	-0.064	0.030	-0.068	0.046	-0.094
Working memory	-0.054	0.040	0.088	0.095	0.028	0.094	0.012
IQ	0.105	0.233*	0.060	-0.084	-0.005	-0.076	-0.034

Multiple linear regression models adjusted for sex, peak height velocity (years) and parent education university level (neither/one/both).
Bolded font indicates that the specific association surpassed the Benjamini-Hochberg procedure for multiple comparison tests (performed for each domain, i.e., academic achievement, executive function and IQ).
* $P < 0.05$
** $P < 0.01$

Furthermore, higher sleep efficiency and shorter WASO time were associated with greater hippocampal GMV in the right hemisphere. This is the first study to investigate the associations between these variables and GMV in children. Of note, sleep efficiency and WASO time were highly correlated in this sample (see **ESM 9, Table S1**). Our conclusion is that WASO time should be as short as possible, meaning that sleep efficiency and GMV in the hippocampus would be

ESM 9 (scan or click here)

Abbreviations in this page:
GMV: grey matter volume
IQ: intelligence quotient
PA: physical activity
ROI: region-of-interest
SB: sedentary behaviour

higher. It is noteworthy that our whole-brain volumetric analyses failed in finding associations between sleep behaviours and GMV in the hippocampal regions as the ROI analysis did. The whole-brain analysis requires enough contiguous voxels associated with sleep behaviours to consider a significant association, which makes this analysis stricter than the ROI.

It seems that the right hippocampus is more sensitive to sleep behaviours than the left hippocampus in children with overweight or obesity. However, GMV in the right hippocampus was not clearly associated with academic achievement, executive function or IQ (i.e., only the anterior sub-section of the right hippocampus associated with inhibition).

Table 31
Standardized beta coefficients for the association of GMV in hippocampal regions and academic achievement, executive function, and IQ (N=96).

	Right hippocampus	Right anterior hippocampus	Right posterior hippocampus
Academic achievement			
Reading	0.106	0.063	0.161
Mathematics	0.06	0.043	0.087
Writing	0.128	0.114	0.139
Academic skills	0.059	0.015	0.122
Academic fluency	0.164	0.156	0.16
Academic applications	0.099	0.092	0.106
Total achievement	0.116	0.085	0.155
Executive function			
Cognitive flexibility	-0.03	-0.038	-0.018
Inhibition	-0.204*	**-0.222***	-0.159
Working memory	-0.03	-0.045	-0.009
IQ	0.059	0.046	0.08

Multiple linear regression models adjusted for sex, peak height velocity (years) and parent education university level (neither/one/both).
Bolded font indicates that the specific association surpassed the Benjamini-Hochberg procedure for multiple comparison tests (performed for each domain, i.e., academic achievement, executive function and IQ).
*Indicates p<0.05
**Indicates p<0.01
† Indicates that the score is multiplied by -1 (i.e., a positive association is interpreted as higher inhibition)

Limitations and Strengths

Several limitations of this study should be acknowledged. First, the cross-sectional design of the study does not allow a causal interpretation of the findings. Next, accelerometer-based estimates of sleep do not represent sleep itself, but an estimation based on movement patterns, so our findings should be interpreted with caution. However, accelerometers are the less-invasive objective method to assess sleep behaviours in free-living settings, while also providing good validity [38,40]. Likewise, nap time cannot be accurately identified via accelerometers, and we did not collect self-report information on naps; thus, it could be that part of the daily sleep is missing in our estimates. Nevertheless, strengths of this study include: its relatively large sample size (96 children with valid MRI); the consideration of SB and PA as a potential confounding factor for sleep behaviours and GMV; the objective assessment of sleep behaviours during a whole week; the standardized tests for the measurement of academic achievement rather than school grades; and the focus on

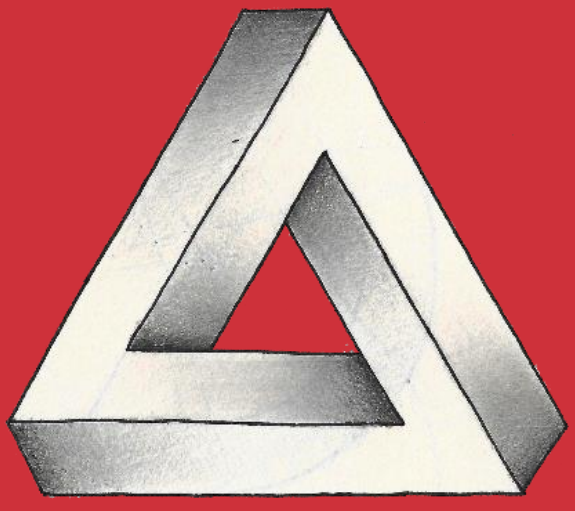
children with overweight or obesity, given the bidirectional associations between obesity and sleep behaviours, PA and brain development.

Conclusion

In conclusion, our findings indicate that sleep behaviours, including timing, duration and patterns, are associated with GMV and, subsequently, GMV is associated with academic achievement and IQ in children with overweight or obesity. It seems that the superior parietal and postcentral cortices are the most consistent regions associated with sleep, being also associated with academic achievement indicators. We should also highlight that WASO time was associated with GMV in both cortical structures and, subsequently, related strongly to academic achievement. Total sleep time, sleep efficiency and WASO time seem to be specifically associated with the right hippocampus, but this subcortical region did not associate with academic achievement, executive function or IQ. Sleep behaviours seem important for GMV and academic achievement and, to a lesser extent, for IQ, but they were not associated with executive function. All these associations remained significant after considering the potential effect of SB, LPA or MVPA.

Abbreviations in this page:
GMV: grey matter volume
IQ: intelligence quotient
LPA: light physical activity
MVPA: moderate-to-vigorous physical activity
PA: physical activity
WASO: wakening after sleep onset

Study X



Activity-rest circadian pattern and academic achievement, executive function and intelligence in children with overweight or obesity

Migueles JH, Martinez-Nicolas A, Cadenas-Sanchez C, Esteban-Cornejo I, Muntaner-Mas A, Mora-Gonzalez J, Rodriguez-Ayllon M, Madrid JA, Rol MA, Hillman CH, Catena A, Ortega FB.
Scand J Med Sci Sports 2020
DOI: 10.1111/sms.13862

Contents

Abstract

Background | Lifestyle behaviours such as PA, SB and sleep are associated with academic achievement, executive function and crystallized intelligence in children. Less attention has been paid to the circadian rhythm of PA and resting (namely the activity-rest pattern) in relation to these outcomes in children with overweight or obesity.

Aims | This study aimed to analyse the associations of activity-rest pattern indicators with academic achievement, executive function and IQ; and to explore whether these associations are mediated by the total GMV among children with overweight or obesity.

Methods | Ninety-five children (10±1 year, 37 girls) with overweight or obesity (based on the WOF BMI cut-off points) were included in this cross-sectional study. Hip- and wrist-worn ActiGraph GT3X+ accelerometers were used to assess the activity-rest pattern. IS, IV, the mean value of the lowest 5h (L5) and the mean value of the maximum 10h (M10) of activity and their respective timing (TL5, TM10) were used as indicators of the activity-rest pattern throughout the day. Chronotype and social jetlag were used as indicators of circadian preference. Academic achievement, executive function and intelligence were assessed with standardized tests. GMV was acquired by MRI.

Results | IS was positively associated with executive function (β=0.244, P=0.014). IV was negatively associated with mathematics and academic applications (β: -0.211 to -0.238, P's ≤ 0.026). Later TM10 in the day was related to lower writing, academic skills and intelligence (β: -0.229 to -0.271, P's ≤ 0.025). None of the associations found were mediated by GMV.

Conclusion | A non-fragmented and stable activity-rest pattern and earlier physical activity in the day were associated with better academic achievement, executive function and IQ in children with overweight or obesity. Further studies are required to corroborate or contrast our findings.

Abbreviations in this page:
BMI: body mass index
CPM: counts per minute
GMV: grey matter volume
IS: interdaily stability
IV: intradaily variability
L5: mean CPM of the 5 h with the lowest activity
M10: mean CPM of the 10 h with the highest activity
MRI: magnetic resonance imaging
PA: physical activity
SB: sedentary behaviour
TL5: timing of the L5
TM10: timing of the M10
VMCounts: activity counts in the vector magnitude
WOF: World Obesity Federation

Key Points

Question
Is the activity-rest pattern associated with academic achievement, executive function, and/or IQ in children with overweight or obesity?

Findings
Higher IS, lower IV, and earlier TM10 were associated with academic achievement and IQ. We did not find evidence of mediation via GMV in this study

Meaning
Brain health may be benefitted from a stable and non-fragmented activity-rest pattern in children with overweight or obesity.

Introduction

Enhancing brain development to reach optimal cognitive functioning and academic success in youth is a universal public health goal [51]. Lifestyle behaviours such as PA, SB and sleep are associated with academic achievement, executive function and crystallized intelligence in children [423–425]. Less attention has been paid to the circadian rhythm of PA and resting in relation to these outcomes, namely the activity-rest pattern. The activity-rest pattern defines the timing and stability of these behaviours throughout the day and across days. Unstable and fragmented patterns are associated with obesity, lower CRF, and mood disorders in youth [60,426].

To date, the understanding of whether the activity-rest pattern is related to academic achievement, executive function or IQ in children is inconclusive. A previous study found a moderating effect of sleep type on attention and IQ among children. Specifically, they observed higher scores in attention and IQ in early (i.e., usually classified as sleep midpoint before 3 am) versus late types of sleep timing among children [427]. A prior meta-analysis reported poorer academic achievement for those adolescents and young adults with a late sleep type (i.e., late bedtime and rise up, usually derived from sleep midpoint after 4 am) [395]. Furthermore, social jetlag (i.e., time difference between sleep midpoint in weekdays and weekends) is negatively associated with cognitive and academic performance in adolescents [428]; however, information in children is lacking.

PA, SB, and sleep have been associated with GMV in specific brain areas, which may explain, in part, their associations with academic achievement, executive function and IQ.[54,332,400,429,430] In children with overweight or obesity, we found a moderator effect of weight status on the association of PA and GMV in the hippocampus [332]. Additionally, Esteban-Cornejo et al. found that cardiorespiratory fitness (which can be effectively modified through PA) is associated with GMV in seven different brain regions important for academic achievement (e.g., premotor cortex, supplementary motor cortex and hippocampus) [54]. Relative to sleep, Urrila et al. found that shorter time in bed and later bedtime hours are associated with lower GMV and poorer school grade average [400]. Therefore, if the activity-rest pattern is associated with academic achievement, executive function and/or IQ, GMV could mediate this relationship.

Pediatric obesity is associated with a fragmented and unstable activity-rest pattern [60] and impaired brain health in children [54,332,374] (**Figure 40** shows an unstable/fragmented [participant 1] and a stable/non-fragmented [participant 2] pattern examples). Thus, the study of children with overweight or obesity should consider appropriate activity-rest patterns in the design of interventions

Abbreviations in this page:
CRF: cardiorespiratory fitness
GMV: grey matter volume
IQ: intelligence quotient
PA: physical activity
SB: sedentary behaviour

Abbreviations in this page:
BMI: body mass index
GMV: grey matter volume
IQ: intelligence quotient
IS: interdaily stability
IV: intradaily variability
PA: physical activity
RCT: randomized controlled trial
WOF: World Obesity Federation

to benefit their cardiometabolic and brain health. However, no previous studies have focused on the associations of activity-rest patterns with brain health outcomes (including academic, cognitive, and brain structure outcomes) in children with overweight or obesity, nor in children in general. This study aimed to: study the associations of the activity-rest pattern indicators with academic achievement, executive function and IQ; and explore whether these associations are mediated by GMV in children with overweight/obesity.

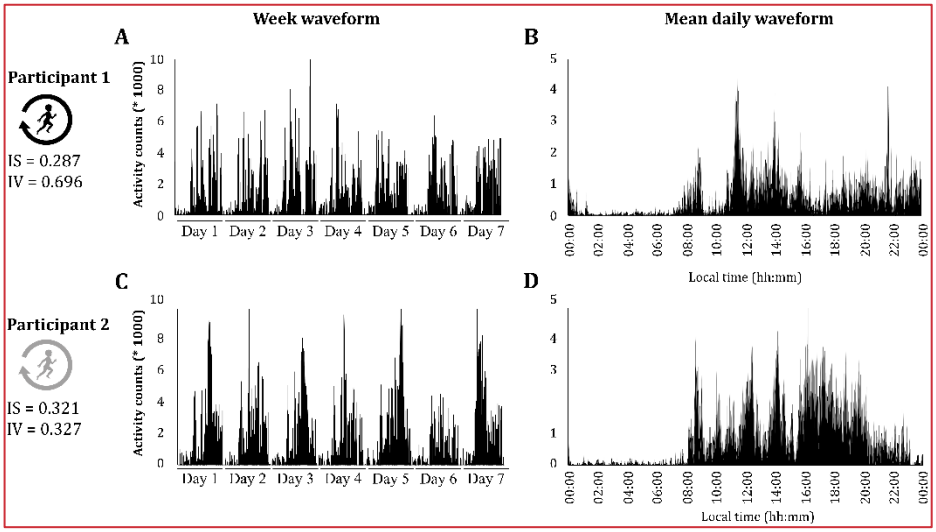


Figure 40
Examples of differing interdaily stability (IS) and intradaily variability (IV) among two participants of the ActiveBrains project. Participant 1 presents a more fragmented and less stable activity-rest pattern than participant 2.

Methods

Participants and study design

This study included baseline data from the ActiveBrains project [63], which is a RCT intended to examine the effect of a 20-week PA intervention on brain structure, brain function, cognitive performance, academic achievement, and physical and mental health outcomes in children with overweight or obesity. Children were recruited in three waves from hospitals and health care centers from Granada, public and private schools, and local media advertising. Out of the 110 children enrolled, those with valid data at baseline were included in this cross-sectional analysis (N=95, 10±1 year, 37 girls). Detailed rationale and methods have been described elsewhere [63]. Briefly, participants were 8-11 years old with overweight or obesity (based on the WOF BMI cut-off points). Participants' weight and height were obtained and their BMI was calculated as weight (kg) divided by squared height (m^2). Weight and height were measured twice consecutively with an electronic scale (SECA 861, Hamburg, Germany) and a stadiometer (SECA 225, Hamburg, Germany), respectively and averaged values were used. Baseline data were collected from November 2014 to February 2016. Parents were informed of the purpose of the study and parental written informed consents were obtained. The ActiveBrains project was approved by the Ethics Committee on Human Research of the University of Granada.

Activity-rest pattern

Participants wore accelerometers (ActiGraph GT3X+, Pensacola, FL, USA) on their right hip and non-dominant wrist for seven days (24 hours/day) as previously reported [266]. In brief, accelerometers collected data at 100 Hz with a dynamic range of ±6 *G*. Raw accelerations were processed with the default filter in the ActiLife software (ActiGraph, Pensacola, FL, USA) and the VMCounts were aggregated over 30 s epochs. Periods of 30 consecutive minutes recording 0 CPM were considered non-wear time and excluded from analyses. Those children wearing devices more than 16 hours/day for at least 4 days (including 1 weekend day) were included.

Abbreviations in this page:
BMI: body mass index
CPM: counts per minute
IS: interdaily stability
IV: intradaily variability
L5: mean CPM of the 5 h with the lowest activity
M10: mean CPM of the 10 h with the highest activity
TL5: timing of the L5
TM10: timing of the M10
VMCounts: activity counts in the vector magnitude

Table 32
Description and interpretation of the activity-rest parameters.

Parameter	Derived from	Description and interpretation
Interdaily stability (IS)	Accelerometer	Ratio of activity level variance within each 24-hour period to the overall variance. The higher the IS, the higher the stability of the activity-rest pattern.
Intradaily variability (IV)	Accelerometer	Ratio of the mean squares of the difference between consecutive hours and the mean squares around the overall mean. The larger the IV, the larger the fragmentation of the activity-rest pattern.
Lowest 5-h (L5)	Accelerometer	Average CPM over 10-min epochs during the consecutive 5 h with the lowest CPM values. The higher the L5, the higher the intensity of the activity in this part of the day (usually sleep period).
Lowest 5-h timing (TL5)	Accelerometer	Central time of the L5 occurrence. Indicated as hh:mm.
Maximum 10-h (M10)	Accelerometer	Average CPM over 10-min epochs during the consecutive 10 h with the highest CPM values. The higher the M10, the higher the intensity of the activity in this part of the day (usually wake period).
Maximum 10-h timing (TM10)	Accelerometer	Central time of the M10 occurrence. Indicated as hh:mm.
Social jetlag	Sleep diaries	Absolute time difference in hours of the sleep midpoint calculated in weekend days (free days) minus weekdays (school days). The larger the social jetlag, the larger the difference in the timing of sleep of weekdays versus weekend days.
Early sleep type	Sleep diaries	Participants whose sleep midpoint was before 4 am.
Neither sleep type	Sleep diaries	Participants whose sleep midpoint was between 4 am and 5 am.
Late sleep type	Sleep diaries	Participants whose sleep midpoint was after 5 am.

Since the accelerometer data did not adjust to the cosine function (i.e., sinusoid wave), we used non-parametric analysis was performed to characterize the activity-rest pattern [60]. We calculated the following parameters: interdaily stability (IS), ratio of activity level variance within each 24-hour pattern to the overall variance, that indicates the degree of consistency from day to day; intradaily variability (IV), ratio of the mean squares of the difference between consecutive hours and the mean squares around the overall mean, that is a measure of rhythm fragmentation within a day; average CPM over 10-min epochs during the 5 consecutive hours with the lowest CPM in the day (L5) and its respective timing (TL5); average CPM

Abbreviations in this page:
CPM: counts per minute
DFT: design fluency test
IS: interdaily stability
IV: intradaily variability
L5: mean CPM of the 5 h with the lowest activity
M10: mean CPM of the 10 h with the highest activity
TL5: timing of the L5
TM10: timing of the M10
TMT: trail making test

ESM 10 (scan or click here)

over 10-min epochs during the 10 consecutive hours with the highest CPM in the day (M10) and its respective timing (TM10). A stable and non-fragmented activity-rest pattern would be characterized by high IS and low IV, respectively. Likewise, lower L5, earlier TL5, higher M10 and earlier TM10 have been reported to be beneficial for health [431,432]. Brief descriptions and interpretation of activity-rest indicators are presented in **Table 32**. All these variables were calculated twice, using the acceleration signal from the hip-worn and the wrist-worn accelerometers. Analyses using the wrist data are presented in the **ESM 10**.

Additionally, participants were instructed to complete a sleep log with bedtimes (i.e., onset, cessation) every day. Sleep onset and offset were used to classify participants as early, neither or late sleep types as described by Roenneberg et al [433]. Midpoint of sleep corrected for sleep deficit was calculated from the times recorded in the sleep diaries and using the formula proposed by Roenneberg et al [433]. Participants were classified as early sleep type if the midpoint of their reported sleep time was before 4 am or late sleep type if it occurred after 5 am. Those children with a sleep midpoint between 4 am and 5 am were classified as neither sleep type. We used 4 am and 5 am cut points instead of the standard 3 am and 4 am because Spanish time is one hour ahead of its geographical time zone [432]. In addition, social jetlag was quantified as the absolute difference between midsleep on school days and weekend days.

Academic achievement, executive function and intelligence quotient

Academic achievement was assessed with the Spanish version of the Woodcock-Johnson III battery, which provides a standardized, valid, and reliable measure of academic achievement.[71] Children completed a total of 12 tests from this battery including reading, language, mathematics and sciences during one session of 100-120 min. All tests were performed individually by a trained evaluator in the laboratory. Tests were independently checked by two trained evaluators and scores were then processed in the Compuscore software (v. 3.1., Riverside Publishing Company, Itasca, IL, USA). We used the broad reading, mathematics and writing scores, as well as composite scores representing academic skills (answer accuracy), academic fluency (processing speed), academic applications (problem-solving) and total academic achievement score.

Executive function domains included cognitive flexibility, inhibition and working memory as described elsewhere [337]. Cognitive flexibility was assessed with two different tests: the third condition of the DFT and the fourth condition of the TMT [72]. The DFT condition consisted in connecting dots, switching bland and empty dots, using four straight lines to design as many novel shapes as possible in one minute. The TMT condition required participants to draw lines

to connect numbers and letters in numeric and alphabetic order as fast as possible (e.g., 1-A-2-B-3-C). The score from these tests (i.e., number of correct designs in the DFT and completion time in the TMT) was standardized by sex as Z-score and then averaged to obtain a unique indicator of cognitive flexibility. A modified version of the Stroop test [415] was used as an indicator of inhibition. Performance time of condition 3 (i.e., inhibiting reading by naming color) minus condition 1 (i.e., color naming) was used as previously reported [337]. As higher values indicate lower inhibition, we inverted this score for analyses. Finally, working memory was measured from a modified version of the DNMS computerized task [76]. A total of 16 practice trials plus 140 experimental trials were presented in two separated conditions (i.e., low- and high-memory load). Response accuracy for the high-load condition was used as an indicator of working memory. We calculated a composite executive function z-score as an overall measure of executive function, representing the sum of the three domains (i.e., cognitive flexibility, inhibition, and working memory).

IQ was assessed with the Spanish version of the K-BIT [77]. Crystallized and fluid IQ components were assessed with the vocabulary and matrices sub-tests, respectively. Both sub-test scores were summed to obtain a total IQ score.

Magnetic resonance imaging data acquisition and processing

MRIs were collected on a 3.0 Tesla Siemens Magnetom Tim Trio scanner (Siemens Medical Solutions, Erlangen, Germany) with a 32-channel head coil. High-resolution, T1-weighted images were acquired using a 3D MPRAGE protocol. The acquisition parameters were the following: repetition time=2,300 ms; echo time=3.1 ms; inversion time=900 ms; flip angle=9º; field of view=256 x 256; acquisition matrix=320 x 320, 208 slices; resolution=0.8 x 0.8 x 0.8 mm; and scan duration=6 min and 34 s [54].

We used the FreeSurfer software version 5.3.0 (http://surfer.nmr.mgh.harvard.edu) and the standard processing pipeline known as 'recon-all' to assess the total GMV [434]. Imaging pre-processing included quality control, motion correction, spatial normalization to an MNI template and spatial smoothing. Detailed information about pre-processing steps is described elsewhere [54].

Confounders

Participants' weight, height, peak height velocity, parents' education level, and CRF were obtained as part of the protocol of the ActiveBrains project [63]. Weight and height were measured twice consecutively with an electronic scale (SECA 861, Hamburg, Germany) and a stadiometer (SECA 225, Hamburg, Germany), respectively and averaged values were used. BMI was calculated as weight (kg) divided by squared height (m^2). Peak height velocity was derived from

Abbreviations in this page:
BMI: body mass index
CRF: cardiorespiratory fitness
DFT: design fluency test
DNMS: delayed non-match-to-sample
GMV: grey matter volume
IQ: intelligence quotient
K-BIT: Kaufman brief intelligence test
MNI: Montreal Neurological Institute
MPRAGE: magnetization-prepared rapid gradient-echo
MRI: magnetic resonance image
TMT: trail making test

Abbreviations in this page:
ANCOVA: analysis of covariance
CI: confidence interval
GMV: grey matter volume
IQ: intelligence quotient
SD: standard deviation
VO_2max: maximal oxygen consumption

standing and sitting height as a continuous measure of maturational status.[92] Peak height velocity represents the difference (in years) between chronological age and the age at peak height velocity. Cardiorespiratory fitness was assessed with a maximal incremental treadmill tests recommended by the American College of Sports Medicine for poorly fit children and with the 20-m shuttle run test. The treadmill test consisted of walking on a treadmill at a constant speed (4.8 km/h) starting at a 6% slope with grade increments of 1% per min until volitional exhaustion. VO_2max was obtained if participants reached volitional fatigue (>8 points in the OMNI scale), plateau in VO_2max during the last two stages of the test, achieved >85% of their age-predicted maximum heart rate, and/or respiratory exchange ration ≥1.0. We used the Léger equation to derive VO_2max from the 20-m shuttle run test. Parents reported their highest level of education finished and they were categorized as: both of them, one of them or neither of them reaching university level.

Statistics

Participants' descriptive characteristics were summarized as mean and SD or percentages (%). All variables were checked for normal distribution prior to analysis. Hierarchical stepwise regressions were performed with potential confounders as explanatory variables and academic achievement, executive function, and IQ indicators as explained variables. Confounders were selected based on the change in the explained variance and the significance level (sex, peak height velocity offset, and parental education university level). Multiple linear regression models were used to study the associations between activity-rest pattern indicators and academic achievement, executive function and IQ. Analysis of covariance (ANCOVA) models were used to investigate differences in academic achievement, executive function and IQ across early, neither and late sleep types. Simple mediation analyses were performed to investigate whether the associations of activity-rest pattern indicators with academic achievement, executive function and IQ were mediated by total GMV. The PROCESS macro version 2.16.3, model 4, with 5000 bias-corrected (BC) bootstrap samples and 95% CIs was used. Mediation is assessed by the indirect effect of the pertinent activity-rest pattern indicator (independent variable) on academic achievement, executive function, or IQ (dependent variable) through total GMV (mediator). The total (c path), direct (c′ path) and indirect (a*b paths) effects are graphically presented in **Figure 41**. The Sobel test was used to determine whether the indirect effect was equal to zero. The statistical significance level was set at $P<0.05$ and the Benjamini-Hochberg correction for multiple comparison testing was applied [435]. Analyses were performed using the IBM SPSS Statistics for Windows version 22.0 (Armonk, NY: IBM).

Table 33. Descriptive characteristics of participants

	All (N = 95)	Early sleep type† (N = 31)	Neither sleep type† (N = 41)	Late sleep type† (N = 23)	Unadjusted *P*
Age (years)	10.00 (1.13)	9.82 (1.16)	10.04 (1.10)	10.18 (1.15)	0.492
Peak height velocity (years)	-2.32 (0.95)	-2.49 (0.99)	-2.3 (0.84)	-2.14 (1.07)	0.412
BMI (kg/m²)	26.73 (3.59)	26.94 (4.31)	26.51 (3.30)	26.83 (3.10)	0.870
Parent education university level, %					
Neither parent	67	61	66	78	
One parent	16	23	15	9	0.579
Both parents	17	16	19	13	
Activity-rest pattern (derived from the hip accelerometer)					
IS	0.32 (0.05)	0.33 (0.05)	0.33 (0.04)	0.30 (0.05)	0.060
IV	0.49 (0.10)	0.50 (0.11)	0.49 (0.09)	0.47 (0.09)	0.482
L5 (CPM)	28.91 (14.75)	29.97 (11.74)	26.06 (11.05)	32.57 (22.12)	0.213
TL5 (hh:mm)	03:36 (01:12)	03:07 (01:26)	03:36 (00:58)	04:19 (00:43)	**<0.001**
M10 (CPM)	1032.87 (247.05)	1086.8 (293.22)	985.26 (222.71)	1045.07 (211.84)	0.219
TM10 (hh:mm)	15:50 (1:26)	15:36 (1:26)	16:05 (1:12)	15:50 (1:12)	0.388
Social Jetlag (h)	1.39 (0.73)	0.72 (0.39)	1.40 (0.44)	2.27 (0.44)	**<0.001**
Academic achievement‡					
Academic skills	118.63 (16.27)	119.00 (16.02)	120.51 (15.39)	114.78 (18.11)	0.401
Academic fluency	103.52 (11.97)	103.19 (10.87)	105.17 (12.57)	101 (12.34)	0.407
Academic applications	99.22 (9.13)	100.16 (8.44)	99.95 (9.98)	96.65 (8.28)	0.302
Reading	108.17 (13.2)	107.81 (12.67)	108.66 (12.91)	107.78 (14.91)	0.952
Mathematics	101.71 (10.8)	102.71 (10.85)	102.76 (10.85)	98.48 (10.48)	0.261
Writing	113.51 (12.89)	114.45 (12.02)	116.49 (11.46)	106.91 (14.53)	**0.014**
Total achievement	109.06 (12.02)	109.58 (12.11)	110.66 (12.03)	105.52 (11.66)	0.252
Executive function					
Cognitive flexibility (z-score)§	-0.03 (0.81)	-0.01 (0.77)	-0.02 (0.90)	-0.09 (0.71)	0.925
Inhibition (s*-1)	-41.98 (17.38)	-44.31 (15.23)	-40.34 (19.65)	-41.77 (16.12)	0.633
Working memory (%)	65.52 (16.53)	64.11 (14.4)	68.16 (15.87)	62.74 (20.07)	0.473
Executive function score (z-score)	-0.03 (0.69)	-0.10 (0.60)	0.06 (0.74)	-0.10 (0.73)	0.536
IQ‡					
Cristal IQ	102.94 (12.92)	104.87 (12.74)	102.00 (13.43)	102.00 (12.51)	0.602
Fluid IQ	97.34 (12.86)	96.03 (11.84)	97.49 (14.29)	98.83 (11.79)	0.733
Total IQ	98.40 (12.39)	99.13 (12.32)	98.1 (13.03)	97.96 (11.82)	0.924
Gray matter					
Volume (mm³)	730380 (65537)	731569 (67453)	731068 (59507)	727553 (75483)	0.972

Data are presented as mean (SD) except for parent education university level (%).
† Early sleep type includes children whose sleep midpoint is before 4 am, Neither sleep type represents sleep midpoints between 4 am and 5 am and Late sleep type includes children with sleep midpoint after 5 am.
‡ Academic achievement and IQ indicators were calculated based on standardized scores centered at 100.
§ Cognitive flexibility was assessed with two tests: the design fluency test and the trail making test. Standardized composite scores are presented.

Results

Sociodemographic and anthropometric characteristics, activity-rest pattern indicators, academic achievement, executive function and IQ scores of participants are reported in the (**Table 33**). Descriptive values of the activity-rest indicators derived from the wrist-worn accelerometer are presented in **ESM 10** (**Table S1**)

Activity-rest pattern and academic achievement, executive function and intelligence quotient

As shown in **Table 34**, IS was positively associated with working memory (β=0.244, P=0.014); IV was negatively associated with mathematics (β=-0.238, P=0.026) and academic applications (β=-0.262, P=0.013); later TM10 was associated with lower writing (β=-0.229, P=0.025), academic skills (β=-0.256, P=0.008) and total IQ (β=-0.271, P=0.006). Otherwise, social jetlag was not associated with any of the outcomes (all P's>0.05). No associations were found between activity-rest pattern indicators derived from the wrist-worn accelerometer and academic achievement, executive function, or IQ (**ESM 10**, **Table S2**).

Abbreviations in this page:
BMI: body mass index
CPM: counts per minute
IQ: intelligence quotient
IS: interdaily stability
IV: intradaily variability
L5: mean CPM of the 5 h with the lowest activity
M10: mean CPM of the 10 h with the highest activity
SD: standard deviation
TL5: timing of the L5
TM10: timing of the M10

ESM 10 (scan or click here)

Abbreviations in this page:
CPM: counts per minute
GMV: grey matter volume
IQ: intelligence quotient
IS: interdaily stability
IV: intradaily variability
M10: mean CPM of the 10 h with the highest activity
TM10: timing of the M10

No differences in academic achievement, executive function and IQ across early, neither and late sleep types were found, except for the writing (Neither vs Late: 8.4 points, P=0.029) (**Table 35**).

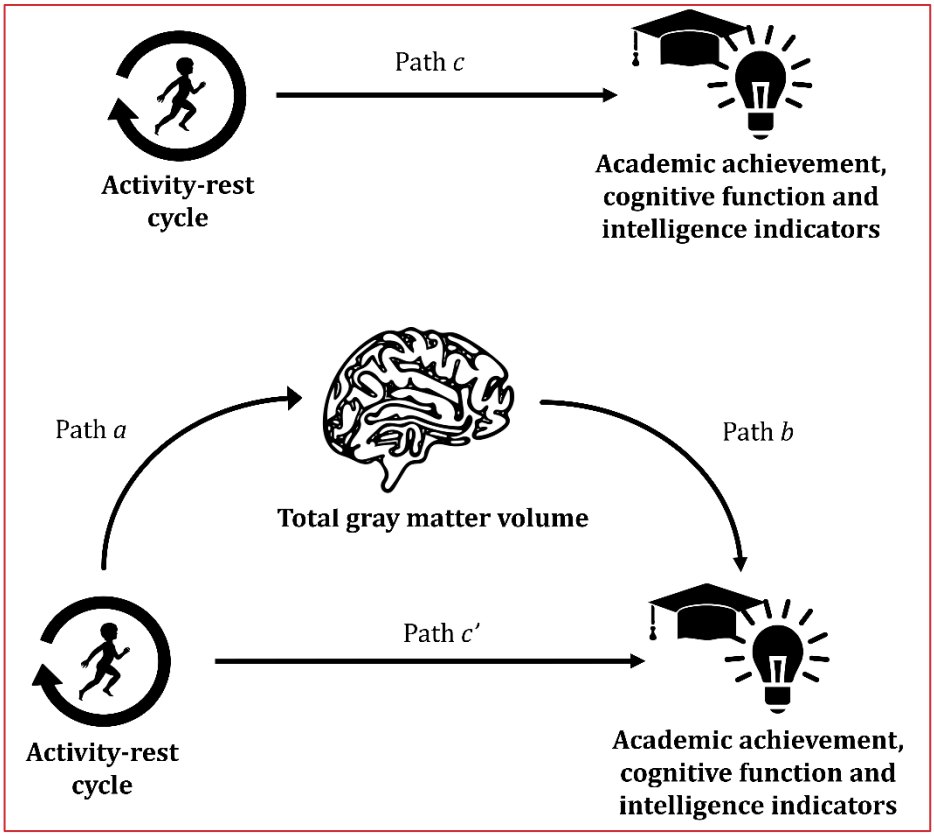


Figure 41
Diagram representing the study simple mediation analyses. Path c shows the association between independent and dependent variables. Paths a x b show the natural indirect effect pathway, and c' shows the natural direct effect pathway.

Mediation role of total grey matter volume

Prior to performing mediation models, we investigated the bivariate correlations of total GMV (i.e., mediator) with academic achievement, executive function, and IQ indicators (i.e., outcomes). We observed positive associations with mathematics (β=0.233, P=0.023), academic fluency (β=0.242, P=0.018), academic applications (β=0.229, P=0.025), and total academic achievement (β=0.215, P=0.037). Associations with other outcomes only demonstrated trends, such as: reading (β=0.187, P=0.070), cognitive flexibility (β=0.186, P = 0.071), executive function score (β=0.192, P=0.062), and crystallized IQ (β=0.192, P=0.063). Mediation analyses were only performed in those associations which were found significant (**Table 34**). **Table 36** shows that none of the associations were significantly mediated by the total GMV.

Discussion

Our findings support the association of activity-rest pattern with academic achievement, executive function and IQ in children with overweight or obesity. Specifically: lower IV and earlier TM10 are associated with higher academic achievement; higher IS and lower IV are associated with better executive function; earlier TM10 is associated with higher IQ scores; and none of these associations were mediated by total GMV. In brief, stable and less fragmented

activity-rest patterns, and earlier PA performance in the day are related to better academic achievement, executive function, and IQ.

Table 34
Standardized β coefficients from linear regression models investigating the association of activity-rest pattern parameters derived from the hip accelerometer with academic achievement, executive function and IQ in children with overweight or obesity (n = 95).

	IS	IV	L5	TL5	M10	TM10	Social Jetlag
Academic achievement							
Reading	0.139	-0.219*	0.074	-0.046	-0.048	-0.120	0.091
Mathematics	0.178	**-0.238***	0.181	-0.177	-0.113	-0.099	0.007
Writing	0.158	-0.082	-0.108	-0.195	-0.076	**-0.229***	-0.038
Academic skills	0.165	-0.159	-0.076	-0.122	-0.109	**-0.256****	0.031
Academic fluency	0.155	-0.176	0.200*	-0.156	-0.077	-0.073	0.058
Academic applications	0.109	**-0.262***	0.149	-0.106	-0.018	-0.004	0.014
Total achievement	0.183	-0.225*	0.045	-0.155	-0.088	-0.181	0.029
Executive function							
Cognitive flexibility	0.184*	-0.171	0.04	-0.085	0.040	0.017	-0.055
Inhibition	0.013	-0.120	0.107	-0.046	-0.166	-0.129	0.050
Working memory	**0.244***	-0.197	0.09	-0.117	-0.153	-0.086	-0.068
Executive function score	0.209*	-0.211*	0.091	-0.104	-0.105	-0.054	-0.047
IQ							
Cristal IQ	0.046	-0.097	0.117	-0.089	-0.007	-0.176	0.017
Fluid IQ	0.104	-0.108	0.215*	-0.057	-0.105	-0.219*	0.060
Total IQ	0.099	-0.119	0.206*	-0.087	-0.071	**-0.271****	0.014

Linear regression models were adjusted for sex, peak height velocity offset and parent education university level.
Bolded font indicates that the significant association (i.e., $p < 0.05$) surpassed the Benjamini-Hochberg correction for multiple comparison tests.
* $P < 0.05$
** $P < 0.01$

Interestingly, associations were only significant using the activity-rest pattern determined with the hip-worn accelerometers. A potential explanation is that wrist-worn accelerometers capture higher variability in the accelerations (see **Table 33** and **Table S1, ESM 10**). Thus, we may be underpowered to find associations using the wrist-worn accelerometer data in this study. Likewise, we recently found that accelerometer-derived PA metrics from hip-worn and wrist-worn accelerometers are hardly comparable [337]. Thus far, it is not clear which body attachment site is better in the characterization of the activity-rest pattern [233]. In this regard, we found that the activity-rest pattern measured at hip is more related to brain health than those measured at wrist.

Higher IS was associated with higher executive function (mainly with working memory) in this sample, but it was not associated with either academic achievement or IQ scores. No previous studies have investigated IS or similar metrics in relation to academic achievement, executive function or IQ in children. Yet, previous studies have found IS to be inversely associated with cognitive function in adults [436] and with emotional face processing in 1-year-old infants [437]. As such, higher IS appears beneficial for cognitive processes in different populations.

Fragmented activity-rest patterns (higher IV) were associated with lower academic achievement and executive function in children with overweight/obesity. This suggests that organized PA is preferable to sporadic and unorganized bouts. No previous studies have directly focused on the IV in relation to academic achievement, executive function and IQ. However, systematic reviews and meta-analyses found that intervention studies using PA breaks during school did

Abbreviations in this page:
CPM: counts per minute
IQ: intelligence quotient
IS: interdaily stability
IV: intradaily variability
L5: mean CPM of the 5 h with the lowest activity
M10: mean CPM of the 10 h with the highest activity
PA: physical activity
TL5: timing of the L5
TM10: timing of the M10

ESM 10 (scan or click here)

not improve academic achievement, but organized PA interventions did (e.g., after-school programs), mainly in mathematics [55,438,439]. These findings corroborate ours, since we found that less fragmented activity-rest patterns (lower IV) were associated with higher scores in mathematics and academic applications (i.e., problem solving).

Table 35
Means and standard error (SE) of academic achievement, executive function and IQ indicators according to early, neither and late types in children with overweight or obesity (n = 95).

	Early sleep type† (N = 31)		Neither sleep type† (N = 41)		Late sleep type† (N = 23)		
	Mean	SE	Mean	SE	Mean	SE	*P*
Academic achievement							
Reading	107.28	2.25	108.42	1.94	108.92	2.62	0.881
Mathematics	101.81	1.78	102.63	1.54	99.91	2.07	0.576
Writing	114.03	2.18	116.27	1.88	107.88	2.54	**0.033**‡
Academic skills	118.49	2.70	120.13	2.33	116.15	3.15	0.598
Academic fluency	102.75	2.12	105.06	1083	101.81	2.47	0.518
Academic applications	99.34	1.50	99.88	1.29	97.89	1.75	0.659
Total achievement	108.86	1.99	110.42	1.72	106.92	2.32	0.476
Executive function							
Cognitive flexibility	0.04	0.13	-0.01	0.11	-0.17	0.15	0.557
Inhibition	43.72	2.81	40.26	2.43	42.71	3.28	0.468
Working memory	64.91	2.89	67.83	2.50	61.52	3.37	0.322
Executive function score	-0.06	0.10	0.06	0.09	-0.17	0.12	0.264
IQ							
Cristal IQ	103.86	2.16	101.84	1.87	103.65	2.52	0.738
Fluid IQ	95.68	2.26	97.22	1.95	99.78	2.63	0.504
Total IQ	98.34	2.09	97.84	1.81	99.47	2.44	0.866

Analyses were adjusted for sex, peak height velocity offset and parent education university level.
† Early sleep type includes children whose sleep midpoint is before 4 am, Neither sleep type represents sleep midpoints between 4 am and 5 am and Late sleep type includes children with sleep midpoint after 5 am.
‡ Bonferroni correction for pairwise comparisons: Writing scores were significantly lower for the late group compared with neither group (-8.4 points; P = 0.029)
Bolded font indicates statistical significance (P < 0.05)

Abbreviations in this page:
CPM: counts per minute
IQ: intelligence quotient
IV: intradaily variability
M10: mean CPM of the 10 h with the highest activity
TM10: timing of the M10
SE: standard error

Furthermore, we found that earlier TM10 is associated with better writing, academic skills and fluid IQ. Concomitantly, those children with a late sleep type (defined as sleep midpoint after 5 am) showed lower scores in writing. In agreement, a previous meta-analysis found lower academic achievement in adolescents with a late compared to early sleep type [395]. Likewise, lower IQ scores were observed in those children with late sleep type [427]. Altogether, it can be speculated that those children with later PA performance and late sleep type had attenuated associations (or no association) with academic achievement and IQ compared to children with earlier PA performance or early sleep type. More studies should investigate the appropriate timing for PA performance with larger cohorts and well-designed studies.

By contrast, social jetlag was not associated with academic achievement, executive function or IQ. Although Social jetlag is negatively related to academic achievement and cognitive function in adolescents [428], its relevance at younger ages is under-reported. A possible explanation is that children still present an overall low social jetlag since their activity during weekends is not substantially displaced to night-time, which would support our findings. More studies with larger sample sizes are needed to confirm our findings.

Previous findings provide support for total GMV in various brain regions as a potential underlying factor that could explain the observed positive associations of PA and sleep with academic achievement, executive function and IQ [400]. Since we were underpowered to test the various mediation models in this study, we decided to consider total GMV as a potential mediator. However, we did not find that total GMV mediated the associations between activity-rest pattern and academic achievement, executive function and IQ in children with overweight/obesity. This may indicate that a detailed study of certain brain regions would be needed to find the mediation role of GMV. In this regard, the main brain region underlying circadian rhythms is the suprachiasmatic nucleus in the hypothalamus. Given the small size of this brain region (i.e., around 2 mm^3) further complications may be added when trying to isolate this region. Other mechanistic pathways including variability in functional connectivity should be also investigated.

Abbreviations in this page:
CPM: counts per minute
GMV: grey matter volume
IQ: intelligence quotient
IS: interdaily stability
IV: intradaily variability
L5: mean CPM of the 5 h with the lowest activity
M10: mean CPM of the 10 h with the highest activity
PA: physical activity
SE: standard error
TM10: timing of the M10

Table 36. Total, direct and indirect effects of the simple mediation analyses investigating total GMV as a mediator in the association of activity-rest circadian rhythm with academic achievement, executive function and IQ (n = 95).

Predictor	Outcome	Total effect (*c*)	Direct effect (*c'*)	Path *a*	Path *b*	Indirect effect (*ab*)	BC 95% CI (lower, upper)	Sobel test
	Academic achievement							
IV	Reading	**-0.219 (0.109)***	-0.207 (0.111)	-0.171 (0.101)	0.069 (0.114)	-0.012 (0.024)	-0.100, 0.013	-0.570 (0.021)
IV	Mathematics	**-0.238 (0.105)***	**-0.215 (0.106)***	-0.171 (0.101)	0.131 (0.109)	-0.022 (0.031)	-0.130, 0.010	-0.980 (0.023)
IV	Academic applications	**-0.262 (0.104)***	**-0.234 (0.105)***	-0.171 (0.101)	0.163 (0.108)	-0.028 (0.031)	-0.133, 0.006	-1.127 (0.025)
IV	Total achievement	**-0.225 (0.106)***	-0.204 (0.108)	-0.171 (0.101)	0.123 (0.110)	-0.021 (0.028)	-0.120, 0.008	-0.933 (0.023)
L5	Academic fluency	**0.200 (0.100)***	0.189 (0.099)	0.061 (0.090)	0.192 (0.116)	0.012 (0.018)	-0.008, 0.074	0.627 (0.019)
TM10	Writing	**-0.229 (0.101)***	**-0.223 (0.100)***	-0.041 (0.093)	0.145 (0.114)	-0.006 (0.016)	-0.059, 0.014	-0.417 (0.014)
TM10	Academic skills	**-0.256 (0.095)****	**-0.255 (0.095)****	-0.041 (0.093)	0.041 (0.108)	-0.002 (0.010)	-0.036, 0.010	-0.288 (0.006)
	Executive function							
IS	Cognitive flexibility	**0.149 (0.074)***	0.150 (0.075)	-0.012 (0.093)	0.055 (0.084)	-0.001 (0.013)	-0.038, 0.018	-0.127 (0.005)
IS	Working memory	**0.233 (0.099)***	**0.234 (0.099)***	-0.012 (0.093)	0.082 (0.112)	-0.001 (0.017)	-0.050, 0.027	-0.127 (0.008)
IS	Executive function score	**0.145 (0.058)***	**0.145 (0.058)***	-0.012 (0.093)	0.017 (0.066)	0.0001 (0.008)	-0.019, 0.014	-0.115 (0.002)
IV	Executive function score	**-0.146 (0.064)***	**-0.148 (0.065)***	-0.171 (0.101)	-0.012 (0.067)	0.002 (0.013)	-0.016, 0.040	0.178 (0.012)
	IQ							
L5	Fluid IQ	**0.215 (0.099)***	**0.206 (0.099)***	0.061 (0.090)	0.160 (0.115)	0.010 (0.015)	-0.007, 0.061	0.609 (0.016)
L5	Total IQ	**0.206 (0.095)***	**0.192 (0.093)***	0.061 (0.090)	**0.227 (0.109)***	0.014 (0.019)	-0.013, 0.071	0.645 (0.021)
TM10	Fluid IQ	**-0.219 (0.102)***	**-0.213 (0.101)***	-0.041 (0.093)	0.166 (0.115)	-0.007 (0.018)	-0.065, 0.015	-0.422 (0.016)
TM10	Total IQ	**-0.271 (0.096)****	**-0.262 (0.094)****	-0.041 (0.093)	**0.229 (0.107)***	-0.009 (0.023)	-0.076, 0.023	-0.432 (0.022)

Data presented as standardized coefficients (SE) and BC 95%CI based on 5000 bootstraps. All analyses were adjusted for sex, peak height velocity offset and parent education university level.

Thus far, there is little information on whether PA and sleep benefits also appear specifically in children with overweight/obesity. For example, Crova et al. found higher benefits of PA on cognitive performance in children with overweight compared to normal-weight children [378], although most previous research did not find the weight status to be a modifier of this association [55]. In this study, we found some promising associations that could guide the appropriate designs of lifestyle interventions to amplify the effects of PA on brain health in children with overweight/obesity. The major limitations of this study are the cross-sectional design and the

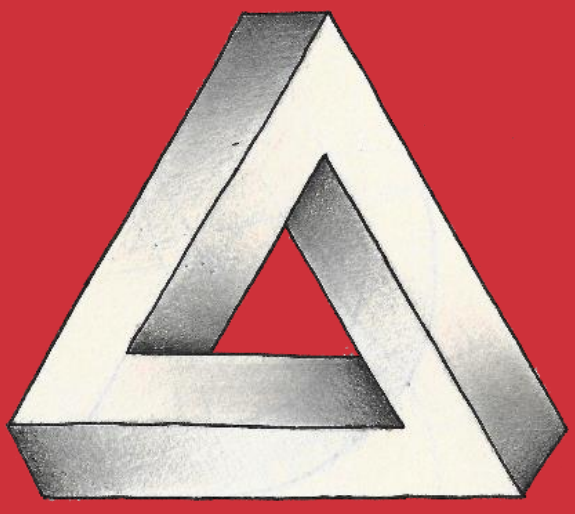
Abbreviations in this page:
GMV: grey matter volume
IQ: intelligence quotient
MRI: magnetic resonance image
PA: physical activity

limited sample size for the mediation analyses. The cross-sectional design hinders the investigation of the mechanisms underlying the associations found. Likewise, there might be uncontrolled confounders, e.g., screen time before bed, which could affect the activity-rest patterns. Such influences on the association with brain health outcomes should be further investigated. Strengths of this study are: the use of objective assessment of the activity-rest pattern from hip- and wrist-worn accelerometers; the use of standardized tests to assess academic achievement, cognition and IQ; the use of MRI (gold-standard) to study the potential mechanistical pathway of the GMV for the observed associations; and the focus on children with overweight or obesity, which are a public health concern.

Perspectives

A more stable and less fragmented activity-rest pattern is associated with better academic achievement, executive function and IQ in children with overweight/obesity. Likewise, earlier PA occurrence is preferable for academic achievement, executive function and IQ in this population. Altogether, these findings provide insights into factors that contribute to the appropriate design of effective PA programs for children with overweight or obesity. Future studies should confirm these findings with larger sample sizes that investigate the mechanistic pathways responsible for the associations depicted.

Study XI



Effects of exercise on cardiometabolic and mental health in children with overweight or obesity: The ActiveBrains randomized controlled trial

Migueles JH*, Cadenas-Sanchez C*, [...], Ortega FB.
***Co-first authors**
In process to be submitted

Contents

Abstract

Background | Childhood obesity is a risk factor for later T2D and CVDs, as well as mental disorders.

Aims | The aim of our study was to investigate the effects of an exercise program on cardiometabolic and mental health in children with overweight or obesity.

Methods | A total of 98 (N = 47 exercise group, mean age 10.4±1.1 years) children with overweight or obesity participated in the two-arm ActiveBrains RCT. The ActiveBrains exercise program included a combination of aerobic and resistance training from Monday to Friday during 20 weeks. The outcomes were cardiometabolic health and mental health.

Results | The ActiveBrains exercise program reduced metabolic syndrome (MetS) risk score around 30% (95% confidence interval, CI: -0.65 to 0.03) compared to control group, which was confirmed with the use of two different validated MetS risk scores. The exercise program had a positive effect on cardiometabolic health by reducing the BMI (-0.60 kg/m^2, 95% CI: -1.07 to -0.13), the fat mass index (-0.70 kg/m^2, 95% CI: -1.03 to -0.36), and visceral adipose tissue (-34.05 g, 95% CI: -61.38 to -6.73); and increasing cardiorespiratory fitness (+3.07 ml/kg/min, 95% CI: 0.68 to 5.45) in the exercise group compared to the control group. More participants in the exercise group reduced their MetS risk score compared to the control group (MetS 1: 30 vs. 45%; MetS 2: 23 vs. 59%). No effects were observed on mental health outcomes.

Conclusion | The ActiveBrains exercise program improved cardiometabolic health in children with overweight or obesity, but it had no effect on mental health. These findings support public health initiatives promoting exercise programs in children with obesity to prevent future comorbidities.

Clinical Trial registration no. NCT02295072

Abbreviations in this page:
BMI: body mass index
CI: confidence interval
CVD: cardiovascular disease
MetS: metabolic syndrome
MVPA: moderate-to-vigorous physical activity
RCT: randomized controlled trial
SB: sedentary behaviour
T2D: type 2 diabetes

Key Points

Question
Can a 20-week exercise program improve the cardiometabolic and mental health in children with overweight or obesity?

Findings
The ActiveBrains exercise program improved the cardiometabolic health and more children experienced meaningful changes in the exercise group compared with the control group. Mental health was not affected

Meaning
These findings support public health initiatives promoting exercise in children with overweight or obesity and should be considered in future exercise programs

Introduction

Around 380 million people live with T2D worldwide [440], which might double their risk of CVD [441]. Obesity is a risk factor for T2D and CVD [48], and this condition may onset in childhood. Of concern, T2D is increasing in youth with obesity across the globe [50,442]. Thus, best practice prevention of T2D and CVD should initiate fighting against obesity in childhood. Children with obesity are often characterized by poor cardiometabolic [45,46,443] and mental health [47]. Exercise is considered an essential component of obesity treatment programs in children due to its physical, psychological, and cognitive benefits [51]. However, few studies have examined the parallel effects of exercise on cardiometabolic and mental health in children with obesity, being therefore unable to compare whether the effect sizes are larger in the cardiometabolic or mental health dimension.

The HEARTY study demonstrated exercise benefits on cardiometabolic health [444], physical fitness [445], and mental health [446] in adolescents with obesity. In children, previous trials found benefits of exercise on visceral fat [447,448], HDL, LDL [447,449], insulin resistance [448], blood pressure [450], body composition [448,449], and CRF [448–450]. However, none of these studies in children included mental health indicators to elucidate the parallel effects of exercise in cardiometabolic and mental health in children with obesity. Indeed, a recent systematic review and meta-analysis conducted by our group found insufficient evidence for the effect of exercise on the children's mental health [451].

Although previous studies found effects on at least one dimension of cardiometabolic health, the effects were inconsistent across studies. Moreover, none of the abovementioned trials analyzed the effects of the exercise program on a composite metabolic syndrome (MetS) risk score, which is relevant to quantify the overall risk for future T2D, CVD, and other cardiometabolic diseases [452,453]. Likewise, only one previous trial reported the percentage of children with obesity experiencing meaningful changes in their cardiometabolic health and body composition [454]. Given the well-known multi-organ effects of exercise, it is relevant to shed light on the parallel effects of exercise on cardiometabolic and mental health in children with obesity. This holistic approach will guide policy makers in the implementation of strategies to target an array of health outcomes in children with obesity. Therefore, the primary aim of our study was to investigate the effects of a 20-week exercise program on cardiometabolic and mental health in children with overweight or obesity. Our secondary aim was to examine the within-individual variability in the effects observed.

Abbreviations in this page:
CRF: cardiorespiratory fitness
CVD: cardiovascular disease
HDL: high-density lipoprotein cholesterol
HEARTY: Healthy Eating Aerobic and Resistance Training in Youth
LDL: low-density lipoprotein cholesterol
MetS: metabolic syndrome
T2D: type 2 diabetes

Abbreviations in this page:
ADHD: attention-deficit hyper-activity disorder
ESM: electronic supplementary material
MVPA: moderate-to-vigorous physical activity
PA: physical activity
RCT: randomized controlled trial

Methods

Study design and participants

This study investigated the effects on secondary outcomes of the ActiveBrains RCT (http://profith.ugr.es/activebrains, NCT02295072). The primary aim of the ActiveBrains RCT was to investigate the effect of an exercise program on brain and cognitive function in children with overweight and obesity [63]. The rationale, protocol, and inclusion criteria are described elsewhere [63]. In brief, a total of 109 pre-pubertal children (8-11 years) with overweight or obesity [248], and not presenting any neurological or physical problem (including ADHD) took part in the ActiveBrains RCT [63]. Participants were randomly assigned to either the exercise program (N = 57) or the wait-list control group (N = 52). A person not involved in the assessments performed a computer-based simple randomization procedure for group allocation after baseline data collection. Participants entered in the ActiveBrains RCT in three different waves for feasibility reasons. **Figure 41** depicts the participants flowchart of the project. The ActiveBrains RCT [63] was approved by the Human Research Ethics Committee of the University of Granada.

Exercise program

The ActiveBrains exercise program had a duration of 20 weeks and was based on the global recommendations on PA for children (i.e., 60 min/day of MVPA, including aerobic and muscle-bone strengthening activities) [8]. Five sessions per week were offered (from Monday to Friday), of which participants were recommended to attend at least to three sessions per week. Each session lasted 90 min and included warm-up (5-10 min), aerobic exercise (60 min), resistance training (20 min), and cool-down (5-10 min). The ActiveBrains exercise program was mainly based on aerobically-demanding games and resistance exercises using bodyweight, elastic resistance band (i.e., therabands, TM resistance bands), and stability balls (i.e., fitballs), involving all major muscle groups [63]. Detailed information about the exercise program can be found in **ESM 11**. Heart rate monitors (POLAR RS300X, Polar Electro Oy Inc., Kempele, Finland) were programmed for each child based on their baseline age, sex, weight, height, and maximum heart rate obtained in a previous incremental maximal test. Heart rate monitoring was used to track participants' exercise intensity in every session and adapt the games or instructions if needed. Children spent an average of 38 min per session above 80% of their maximum heart rate. The SAAFE principles (i.e., Supportive, Active, Autonomous, Fair and Enjoyable) proposed by Lubans et al. to maximize the effects of exercise on physical and mental health were met [455]. All sessions were delivered by trainers with university studies in sport sciences. Weekly meetings

ESM 11 (scan or click here)

were conducted with the trainers to determine if the sessions were running as they were previously programmed.

Table 37
Descriptive characteristics of the ActiveBrains participants meeting the per-protocol criteria at baseline.

	All		Control group		Exercise group	
	N	Mean (SD) or %	N	Mean (SD) or %	N	Mean (SD) or %
Sex						
Girls (n %)	41	42%	25	49%	16	34%
Boys (n %)	57	58%	26	51%	31	66%
At risk of dyslipidemia	42	43%	26	51%	16	34%
Pre-diabetes	3	3%	0	0%	3	6%
Pre-hypertension	9	9%	4	8%	5	11%
Age (y)	98	10.1 (1.1)	51	10.1 (1.1)	47	10.0 (1.1)
Peak height velocity (y)	98	-2.2 (1.0)	51	-2.1 (1.1)	47	-2.4 (0.9)
Cardiometabolic health						
LDL (mg/dL)	85	101.1 (24.6)	44	103.1 (24.3)	41	98.9 (24.9)
HDL (mg/dL)	94	50.9 (11.2)	49	50.0 (10.5)	45	52.1 (11.9)
Triglycerides (mg/dL)	94	96.3 (46.0)	46	100.1 (48.8)	45	92.2 (43.1)
Triglycerides-to-HDL (mg/dL)	94	2.2 (1.9)	49	2.2 (1.5)	45	2.2 (2.2)
HOMA	88	2.9 (2.0)	45	2.8 (2-0)	43	3.0 (2.1)
Mean arterial pressure (mmHG)	96	71.3 (12.3)	49	70.4 (12.5)	47	72.3 (12.2)
Body mass index (kg/m^2)	98	26.9 (3.6)	51	26.4 (3.0)	47	27.4 (4.1)
Fat mass index (kg/m^2)	97	11.8 (2.8)	50	11.4 (2.3)	47	12.4 (3.2)
Lean mass index (kg/m^2)	97	13.9 (1.5)	50	13.9 (1.4)	47	14.0 (1.6)
Waist circumference (cm)	96	90.5 (9.8)	51	89.7 (8.6)	47	91.3 (11.0)
Visceral adipose tissue (g)	79	402.8 (114.7)	41	391.4 (105.7)	38	415.1 (123.9)
CRF performance (laps)	96	15.4 (6.9)	49	15.7 (6.9)	47	15.0 (6.9)
Speed-agility fitness (s)	96	15.2 (1.6)	49	15.0 (1.6)	47	15.3 (1.5)
Upper-limb muscular strength (kg)	97	16.8 (4.2)	50	16.9 (4.0)	47	16.7 (4.5)
Lower-limb muscular strength (cm)	96	104.0 (17.8)	49	105.6 (18.6)	47	102.3 (17.0)
MetS score 1*	92	-0.011 (0.551)	47	-0.064 (0.525)	45	0.045 (0.577)
MetS score 2†	86	-0.039 (0.636)	44	-0.093 (0.597)	42	0.017 (0.677)
Mental health						
Stress (0 – 30)‡	95	5.9 (3.3)	49	6.3 (3.1)	46	5.4 (3.4)
Anxiety (20 – 60)‡	94	33.7 (7.3)	49	34.3 (7.0)	45	33.0 (7.5)
Depression (0 – 54)‡	98	8.6 (5.1)	51	9.3 (5.2)	47	7.7 (4.9)
Negative affect (10 – 30)‡	94	16.1 (3.5)	49	16.4 (3.6)	45	15.7 (3.5)
Positive affect (10 – 30)‡	96	24.3 (2.9)	50	24.4 (2.9)	46	24.3 (3.0)
Happiness (4 – 28)‡	98	22.7 (3.9)	51	22.3 (3.7)	47	23.1 (4.1)
Optimism (6 – 30)‡	97	21.9 (4.0)	50	21.9 (4.1)	47	22.0 (3.9)
Self-efficacy (10 – 40)‡	97	30.8 (4.9)	50	30.3 (5.2)	47	31.3 (4.4)
Self-concept (30 – 300)‡	96	227.0 (29.0)	49	225.2 (30.1)	47	228.8 (28.0)
Self-esteem (10 – 40)‡	97	32.9 (4.7)	50	32.5 (4.4)	47	33.3 (4.9)
Psychological ill-being**	87	0.025 (0.739)	46	0.133 (0.705)	41	-0.097 (0.766)
Psychological well-being††	94	-0.028 (0.641)	48	-0.091 (0.636)	46	0.037 (0.647)
Total mental health‡‡	85	-0.016 (0.559)	44	-0.094 (0.533)	41	0.067 (0.582)

Data analyses were primarily conducted under the per-protocol principle, i.e., attending to 70% of the sessions or keep the usual lifestyle for exercise and control groups, respectively.
*MetS score 1 was calculated as the normalized mean of the z-scores for HDL cholesterol, waist circumference, triglycerides, glucose, and the average of systolic and diastolic blood pressure.
†MetS score 2 was calculated as the normalized mean of the z-scores for waist circumference, triglycerides to high-density lipoprotein ratio, mean arterial pressure and fasting insulin.
‡Score range for the questionnaire.
**Psychological ill-being was calculated as the normalized mean of the z-score for stress, anxiety, depression, and negative affect.
††Psychological well-being was calculated as the normalized mean of the z-score for positive affect, happiness, optimism, self-efficacy, self-concept, and self-esteem.
‡‡ Total mental health was calculated as the normalized mean of the z-score for all mental health indicators.
***Standard T-scores (i.e., 50 ± 10) from the Behavioral Assessment System for Children (BASC).

Abbreviations in this page:
CRF: cardiorespiratory fitness
HDL: high-density lipoprotein cholesterol
HOMA: homeostatic model assessment
LDL: low-density lipoprotein cholesterol
MetS: metabolic syndrome
PA: physical activity
SD: standard deviation

Measurements

Measurements were conducted at baseline and repeated after the exercise program finished in both the exercise and control groups. PA was monitored at baseline and during the delivery of the exercise program (10th week) in both groups. Assessors were not blinded to the participants' group allocation due to budget restrictions in the staff employed for the trial. Detailed information on the measurements is provided elsewhere [63]. Sociodemographic data were reported by children and their parents. At baseline, biological maturation was assessed with the peak height velocity from

Abbreviations in this page:
ADHD: attention-deficit hyperactivity disorder
ESM: electronic supplementary material
HDL: high-density lipoprotein cholesterol
HOMA: homeostatic model assessment
ITT: intention-to-treat
LDL: low-density lipoprotein cholesterol
MetS: metabolic syndrome

height and sitting height measurements using the Moore's equations [92].

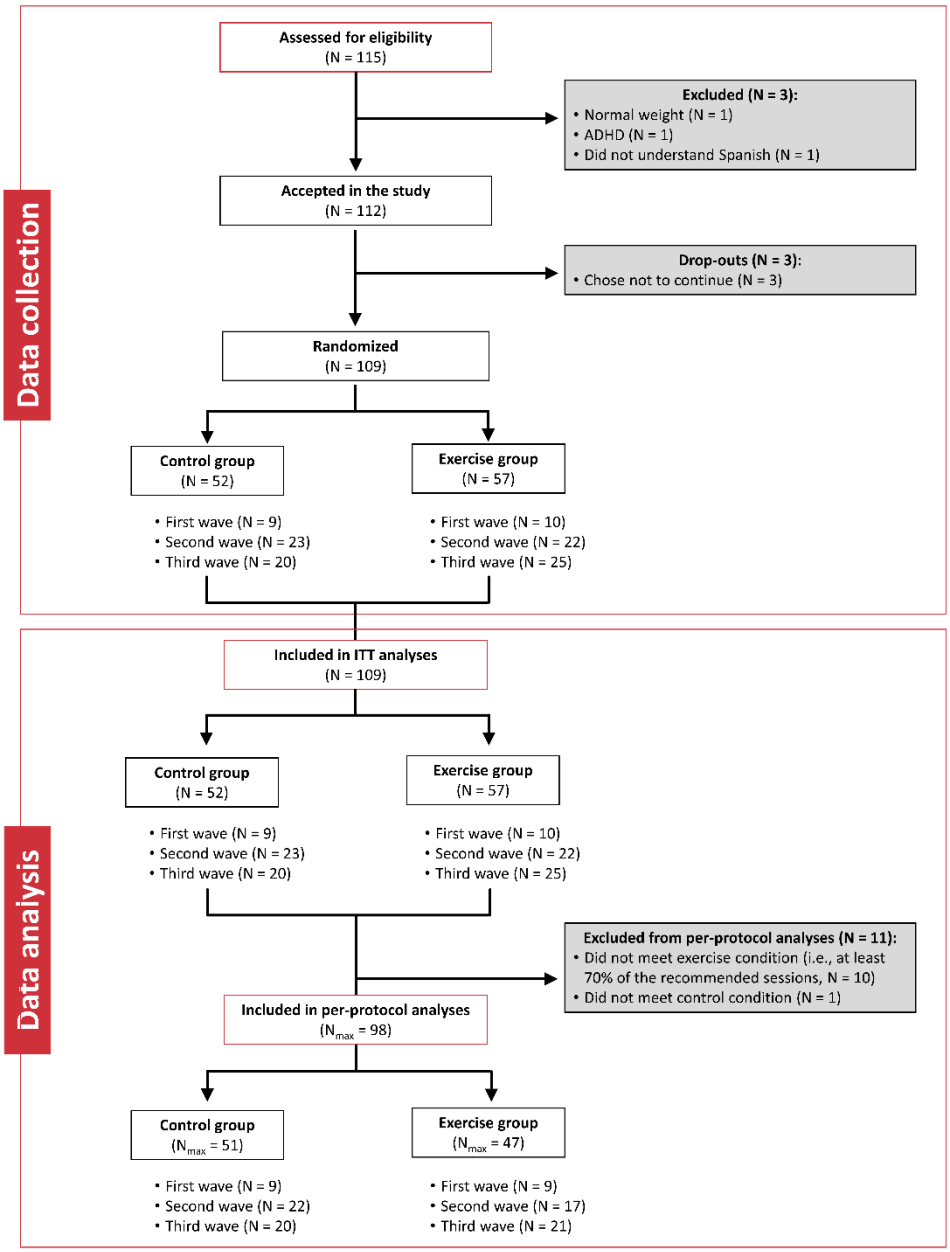


Figure 42
Flowchart of the study.
For final ITT analyses, those participants that left the study during the exercise program or did not complete the post-exercise program assessments were imputed (see Statistical section).
Nmax: Maximum N for analyses, it changes depending on the variable, see Table 1 and **ESM 12** Tables S1 and S3 for the main study outcomes.

ESM 12 (scan or click here)

Cardiometabolic health

Cardiometabolic health outcomes included the traditionally-considered risk factors for MetS (i.e., hyperglycemia, hypertension, and dyslipidemia) [63], as well as body composition, inflammatory biomarkers, and physical fitness, which are closely related to cardiometabolic health [143,456]. Blood lipids biomarkers included fasting LDL, HDL cholesterol, and triglycerides. The triglycerides-to-HDL ratio was calculated. Fasting insulin and glucose were obtained from blood samples and the HOMA (homeostatic model assessment) index was calculated as insulin (μU/mL) multiplied by glucose (mg/dL) and divided by 405. All blood samples were collected at the hospital between 8:00 am and 10:30am after a minimum of 8h overnight fasting, and were analyzed by an accredited laboratory. Systolic and diastolic blood pressure were assessed twice in a sitting position from

the left arm with an automatic sphygmomanometer (Omron M6, Hoofddorp, The Netherlands), and the lowest values out of the two measures was retained for analyses. The systolic and diastolic average, and the mean arterial pressure were calculated. Additionally, the risk of dyslipidemia (based on an alteration of the blood lipids, either triglycerides and/or HDL), pre-diabetes (glucose) and pre-hypertension (systolic and diastolic blood pressure) were classified based on the age- and sex-specific cut-offs that are linked to the Adult Treatment Panel III and International Diabetes Federation criteria [457].

Body weight and height were measured twice with an electronic scale and a stadiometer (SECA, Hamburg, Germany) with participants barefoot and wearing light clothes, and the averages were recorded. BMI was calculated using the standard equation: kg/m^2. Whole-body fat mass and lean mass, and visceral adipose tissue were measured via DXA (Discovery Horizon® DXA system, Hologic, Canada ULC). Fat mass index and lean mass index were calculated as fat or lean mass in kilograms divided by height in meters squared (kg/m^2). Abdominal obesity was represented by the average waist circumference from two measurements following standard protocols [458]. Physical fitness components (i.e., CRF, speed-agility, and muscular fitness) were assessed using feasible, reliable, and valid tests for children [88–90]. Specifically, CRF was assessed using the 20 m shuttle-run test [459]. The number of completed laps was recorded and the Léger's equation was used to estimate the relative VO_2max in ml/kg/min [91]. Speed-agility fitness was assessed through the 4x10 m shuttle run test, where lower values indicate higher performance (i.e., less time to complete the circuit). Muscular fitness was assessed by the handgrip strength (i.e., upper-limb muscular strength) and the standing long jump tests (i.e., lower-limb muscular strength). To account for the body mass, relative upper-limbs strength was calculated as the ratio of the handgrip test to the body weight. We multiplied the distance reached jumping by the bodyweight to calculate the lower-limbs absolute strength (taking body weight into account). Both upper- and lower-limb strength indicators were merged in a composite score indicator of overall muscular fitness. Detailed information on the physical fitness testing can be found elsewhere [63].

Finally, two previously validated MetS risk scores were calculated [452,453]. The MetS score 1 averaged the specific z-scores for the variables included in the most-used definition of MetS (i.e., triglycerides, inverted HDL, glucose, and the average of systolic, and diastolic blood pressure, and waist circumference) [453]. The MetS score 2, which has shown good structural validity in children at cross-sectional and longitudinal level [460], was calculated as the

Abbreviations in this page:
BMI: body mass index
CRF: cardiorespiratory fitness
DXA: dual-energy X-ray absorptiometry
HDL: high-density lipoprotein cholesterol
MetS: metabolic syndrome
VO_2max: Maximal oxygen consumption

Abbreviations in this page:
HDL: high-density lipoprotein cholesterol
LPA: light physical activity
MVPA: moderate-to-vigorous physical activity
PA: physical activity
SB: sedentary behaviour

normalized mean of the z-scores of waist circumference, triglycerides-to-HDL, mean arterial pressure, and fasting insulin [452].

Mental health

Children filled in the mental health questionnaires in three separate days. Both psychological ill-being and well-being components of mental health were assessed using valid self-reported questionnaires. Psychological ill-being included measures of stress (Children's Daily Stress Inventory, scored from 0 to 30), anxiety (State-Trait Anxiety Inventory for Children, scored from 20 to 60), depression (Children's Depression Inventory, scored from 0 to 54) and negative affect (Positive and Negative Affect Schedule for Children, scored from 10 to 30). Otherwise, psychological well-being included positive affect (Positive and Negative Affect Schedule for Children, scored from 10 to 30), happiness (Subjective Happiness Scale, scored from 4 to 28), optimism (Life Orientation Test-Revised, scored from 6 to 30), self-efficacy (General Self-Efficacy, scored from 10 to 40), self-concept (Five-Factor Self-concept questionnaire, scored from 30 to 300) and self-esteem (Rosenberg Self-Esteem Scale, scored from 10 to 40). A detailed description of the mental health indicators assessment and the tests psychometric information can be found elsewhere [63]. Composite standardized scores were calculated for psychological ill-being (i.e., stress, anxiety, depression and negative affect), psychological well-being (i.e., positive affect, happiness, optimism, self-efficacy, self-concept, and self-esteem) and total mental health (i.e., psychological ill-being multiplied by -1 and psychological well-being).

Physical activity assessment

Accelerometer-determined daily time spent in PA, SB, and sleep during the intervention were used to assess the ActiveBrains exercise program fidelity (and possible contamination/compensation effects). Accelerometers (GT3X+, ActiGraph, Pensacola, Florida, USA) were placed on the right hip and the non-dominant wrist to monitor PA for seven days at baseline and during the intervention delivered period. The accelerometers raw data were processed as described elsewhere [266], following the practical recommendations previously done by our group [233]. In brief, a minimum of four valid days (i.e., ≥16 hours/day), including at least one weekend day, was required to be included in the analyses. We used the GGIR software [200] to identify the night sleep periods using an automated algorithm guided by the self-reported sleep times [40,202]. Then, waking time was classified into MVPA, LPA, and SB using children-specific cut-points [61,62,68].

Statistical analysis

Power analyses showed that a sample size of 98 children is enough to detect low-to-medium effect sizes assuming an α error of 0.05 and 80% statistical power. Characteristics of the study participants are presented as mean and SD, or frequency and percentage. Prior to analyses, each outcome was winsorized when needed by replacing extreme values for the closest valid value to avoid the outliers' influence [461]. Then, baseline z-scores of the outcomes were calculated by subtracting their mean and dividing by their SD. Post-exercise z-scores were calculated relative to the baseline mean and SD as a standardized measure of the effect size. Therefore, these post-exercise z-scores are indicative of change (i.e., the number of SDs that the outcomes deviated from their mean baseline value) [461].

ANCOVA was used to test the effects of the ActiveBrains exercise program on each outcome. For this, post-exercise outcome values were the dependent variables, group (i.e., exercise vs. control) as fixed factor, and baseline outcome data as covariates [461]. Analyses were primarily conducted under the per-protocol principle, i.e., attending to 70% of the sessions. Additionally, as sensitivity analysis, we analyzed the data under the ITT principle including all participants and imputing the missing data using predictive mean matching multiple imputations [462]. We checked that missing data was missing at random prior to performing the multiple imputation.

The within-individual change distribution was studied and the changes exceeding 0.2 Cohen's D were considered meaningful, as this is accepted as the minimum relevant standardized effect size of an intervention [454]. Chi-square tests were used to compare the rate of meaningful changes observed in the exercise and the control group. Finally, changes in the daily time spent in MVPA, LPA, SB, and sleep during the exercise program (or usual lifestyle for the control group) were analyzed. We calculated the isometric log-ratios between each group's compositional mean and the overall compositional mean after centering the data at baseline and during exercise [326]. Then, we subtracted the baseline composition from the during-exercise (or usual lifestyle) composition. Values in the change composition are represented as proportional changes (%) from the baseline overall composition. The Hotelling's T-squared test for multivariate pair-wise comparisons was used to compare the change in the time-use composition between the control and the exercise group. All the statistical procedures were performed using the R software (v. 4.0.0.). A significant difference level of $P < 0.05$ was set.

Abbreviations in this page:
ANCOVA: analysis of covariance
ITT: intention-to-treat
LPA: light physical activity
MVPA: moderate-to-vigorous physical activity
SB: sedentary behaviour
SD: standard deviation

Abbreviations in this page:
BMI: body mass index
CRF: cardiorespiratory fitness
HDL: high-density lipoprotein cholesterol
HOMA: homeostatic model assessment
LDL: low-density lipoprotein cholesterol
MetS: metabolic syndrome

Results

A total of 98 participants (N = 47 exercise group) were included in the per-protocol analysis (90% of the participants initially enrolled in the trial adhered to the recommended study protocol, i.e., 3 sessions/week) (**Figure 42**). Participants in the exercise and control groups presented similar baseline characteristics (**Table 37**). At baseline, 43% of the children were at risk of dyslipidemia, 3% presented pre-diabetes, and 9% pre-hypertension. No significant differences were found regarding basic characteristics between the participants included in the per-protocol analysis (N=98), and the rest of the participants (N=11, all $P > 0.05$).

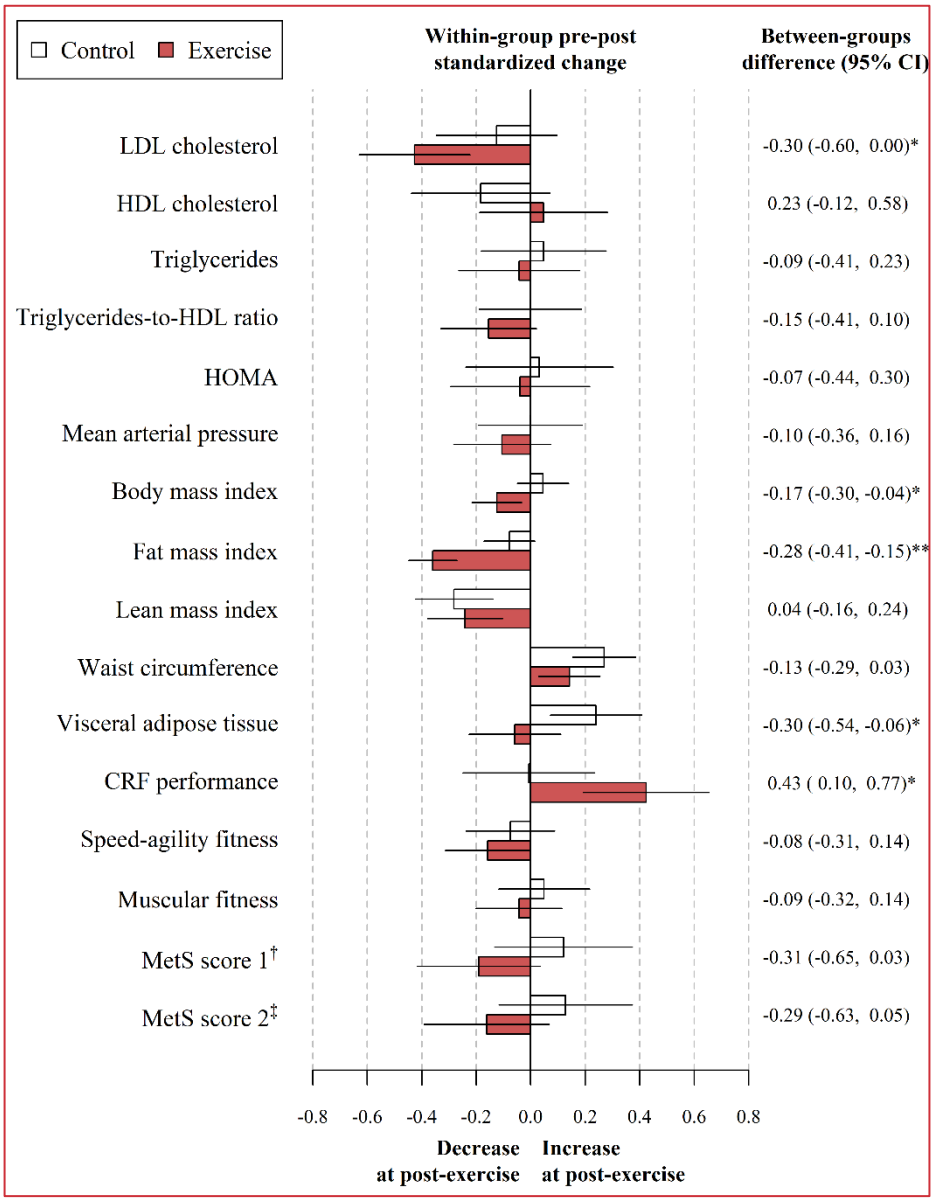


Figure 43
Effects of the ActiveBrains exercise program z-score pre-post change between groups in cardiometabolic risk outcomes (per-protocol analyses).
Data analyses were primarily conducted under the per-protocol principle, i.e., attending to 70% of the sessions. Baseline z-score of the outcomes were calculated by subtracting the mean value and dividing by the SD of each outcome. Post-exercise z-scores were calculated relative to the mean and SD of the baseline values, being a z-score of the change in each outcome, i.e., (post-exercisei – baseline mean) / baseline SD.
†MetS score 1 was calculated as the normalized mean of the z-scores for HDL cholesterol, waist circumference, triglycerides, glucose, and the average of systolic and diastolic blood pressure.
‡MetS score 2 was calculated as the normalized mean of the z-scores for waist circumference, triglycerides to high-density lipoprotein ratio, mean arterial pressure and fasting insulin.
*$P < 0.05$
** $P < 0.01$

ESM 12 (scan or click here)

Cardiometabolic health

Figure 43 and **Table S1 (ESM 12)** show the within- and between-groups pre-post differences in cardiometabolic health outcomes in a graphical and tabulated format, respectively. We found a reduction in the LDL of 7.40 mg/dL (-14.82 to 0.016), the BMI (-0.60 kg/m^2, -1.07 to -0.13), the fat mass index (-0.70kg/m^2, -1.03 to -0.36),

and the visceral adipose tissue of -34.05 g (–61.38 to –6.73) in the exercise group compared to the control group. CRF performance was higher after the exercise program in the exercise group compared to the control group (+3.07 laps, 0.68 to 5.45). The absolute muscular fitness was significantly reduced in the exercise group compared to controls (effect size = -0.16, -0.319 to -0.008), yet this difference was not present in the relative muscular fitness score (**Table S1, ESM 12**). The ActiveBrains exercise program produced a meaningful (yet not statistically significant) reduction on the MetS scores in exercisers compared to controls (effect sizes = -0.30 [-0.652 to 0.029] and -0.29 [-0.628 to 0.048], in MetS score 1 and 2 respectively). Overall, the intention-to-treat analyses showed similar effects than observed in the per-protocol analysis (**ESM 12**, **Table S2**).

More participants in the exercise group showed meaningful changes (i.e., reduction of ≥ 0.2 SD) compared to the control group in fat mass index (79% vs. 36%, $P < 0.001$) and CRF performance (65% vs. 38%, $P = 0.020$) **(Figure 44)**. A marginal difference, yet non-significant, was found in favor to exercise in BMI (34% vs. 16%, $P = 0.081$). Significantly, more participants experienced meaningful changes in the exercise group compared to the control group in the MetS risk scores (MetS score 1: 45% vs. 30%, $P=0.047$; MetS score 2: 59% vs. 23%, $P = 0.007$). Furthermore, since a 5% reduction in body weight is considered clinically relevant [463], we observed that the proportion of participants experiencing a clinically relevant change in the fat mass index was significantly higher in the exercise group compared to controls (34% vs. 79%, $P<0.001$, **ESM 12, Figure S1**).

Mental health

Figure 45 and **Table S3 (ESM 12)** show that the ActiveBrains exercise program did not affect any mental health outcome. Similarly, intention-to-treat analyses showed no effect of the ActiveBrains exercise program on mental health (**ESM 12**, **Table S4**).

Evidence of the exercise program fidelity

As a measure of exercise program fidelity, **Figure 46** shows the exercise-induced changes in physical behaviors derived from the hip- and the wrist-worn accelerometers (Panel A and Panel B, respectively). A no significant but marginal difference in the time-use composition was found using the estimates based on the hip-worn accelerometer ($P = 0.079$, Panel A). Otherwise, the time-use composition was found significantly different between groups using the wrist-worn accelerometer estimates ($P = 0.002$, Panel B). Both the hip- and the wrist-based estimations agree on that the exercise group increased more the time devoted to MVPA than the control group (hip: +15% vs. +7% from baseline; wrist: +21% vs. +7% from baseline). Likewise, the control group did not substantially alter their time in LPA, SB, and sleep from baseline, while the exercise group

Abbreviations in this page:
BMI: body mass index
CRF: cardiorespiratory fitness
ESM: electronic supplementary material
LPA: light physical activity
MetS: metabolic syndrome
MVPA: moderate-to-vigorous physical activity
SB: sedentary behaviour
SD: standard deviation

ESM 12 (scan or click here)

substantially reduced their SB (hip: -6%; wrist: -14%) and sleep time (hip: -8%; wrist: -9%).

Abbreviations in this page:
CRF: cardiorespiratory fitness
HDL: high-density lipoprotein cholesterol
LDL: low-density lipoprotein cholesterol
MetS: metabolic syndrome
RCT: randomized controlled trial
SB: sedentary behaviour

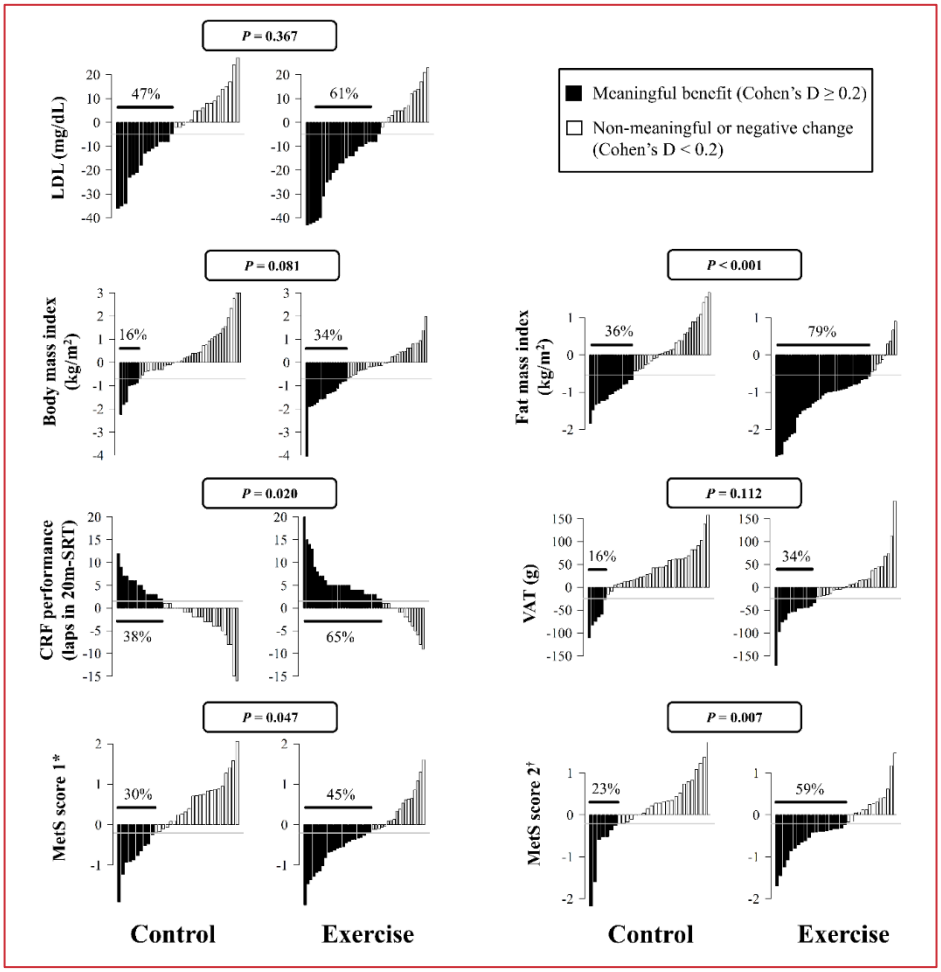


Figure 44
Pre-post change distribution in the outcomes significantly affected by the exercise program.
Data analyses were primarily conducted under the per-protocol principle, i.e., attending to 70% of the sessions.
*MetS score 1 was calculated as the re-normalized mean of the z-scores for HDL cholesterol. waist circumference. triglycerides. glucose. and the average of systolic and diastolic blood pressure.
†MetS score 2 was calculated as the re-normalized mean of the z-scores for waist circumference, triglycerides to high-density lipoprotein ratio, mean arterial pressure and fasting insulin.
P values from chi-squared test.

Discussion

The primary aim of our study was to investigate the effects of the ActiveBrains exercise program on cardiometabolic and mental health in children with overweight or obesity. Secondarily, we examined the within-individual variability in the effects observed. The ActiveBrains RCT demonstrated that a 20-week exercise program reduced the MetS risk around 30% (i.e., ~one-third of one SD of normalized MetS scores) compared to usual lifestyle in children with overweight or obesity. It seems that changes in visceral adipose tissue and LDL could drive this risk reduction. The participant rate presenting meaningful changes in the MetS risk scores was markedly and significantly higher among exercisers than controls. These findings were confirmed by two different and valid MetS risk scores [452,453]. In addition, the ActiveBrains exercise program substantially improved the children's body composition and the CRF, compared to those randomized to the control group. Similar to cardiometabolic health, the participant rate experiencing meaningful changes in body composition and CRF was higher in the exercise group than in the control group. No effects were observed on mental health. Finally, we observed that the exercise group incremented

notably the daily time devoted to MVPA and reduced SB compared to controls during the exercise program, which confirms the exercise program fidelity.

Cardiometabolic health

In our opinion, our models depicted a sizeable effect, yet non-significant, in the MetS risk (i.e., 0.3 SDs reduction in exercise compared to control group). To note that the sample size for the MetS scores was substantially reduced due to data missingness (i.e., 74 and 64 cases for MetS score 1 and 2, versus 91 cases for fat mass index), which may explain the non-significant *P* values for the effect sizes observed. Post hoc power analyses showed with the current sample size for the two MetS risk scores, i.e., 74 and 64 participants, we have power to detect changes ≥ 0.32 SDs with and ≥ 0.34 SDs respectively for a given α error of 5% and statistical power of 95%. However, if the sample size for the MetS risk scores would have been 91 participants (instead of 74 and 64 as it currently is due to missingness), as it is for FMI and other outcomes in our project, we would have power to detect changes ≥ 0.28 SDs with an α error of 5% and statistical power of 95%. These calculations clearly support the notion that the effect size observed in our project for the two MetS risk scores, i.e., 0.3 SDs, is meaningful and would have been a significant ($p<0.05$) if these outcomes would have had the same sample size as the other study outcomes (e.g., fat mass index, 91 participants). The fact that out intervention improved MetS risk scores is further supported by the higher proportion of participants meaningfully reducing the risk in the exercise group compared to controls. The risk reduction was probably driven by improvements in blood lipids, total and visceral adiposity, and CRF. In this regard, our exercise group reduced their fasting LDL cholesterol by 8 mg/dL and their visceral adipose tissue by 34 g compared to the control group. Other blood lipid and adiposity markers showed a positive trend in the exercise compared to the control group, but it was not statistically significant (e.g., waist circumference and HDL cholesterol).

Our results agree with recent meta-analyses in children overweight or obesity on that concurrent aerobic and resistance training is effective to improve blood lipids, mainly LDL and triglycerides [464,465]. Likewise, we observed consistent effects with previous trials in children with overweight or obesity, such as reductions in visceral fat [447,448], LDL [447], and increments in HDL cholesterol [449,450]. Two of the previous RCTs in children with obesity additionally found effects on insulin resistance [448,450] while we did not. We believe that differences in the participants' baseline characteristics may account for our lack of effects on glucose metabolism biomarkers. In this regard, Farpour-Lambert et al. only included children with obesity (excluding children with overweight) [450], while

Abbreviations in this page:
CRF: cardiorespiratory fitness
HDL: high-density lipoprotein cholesterol
LDL: low-density lipoprotein cholesterol
MetS: metabolic syndrome
RCT: randomized controlled trial
SD: standard deviation

Abbreviations in this page:
BMI: body mass index
RCT: randomized controlled trial
SD: standard deviation

the PLAY RCT analyzed 222 participants, of whom 85% had obesity and 28% were pre-diabetic [448]. Our sample included 98 children with overweight or obesity, of whom 77% had obesity, and only 3% were pre-diabetic. Otherwise, almost half of our participants were at risk of dyslipidemia [457]. Thus, there was more room for improvements in blood lipids and adiposity than it was in glycemic metabolism. Our participants were at healthy glycemic levels at baseline, which could produce a floor effect and the so observed lack of effects.

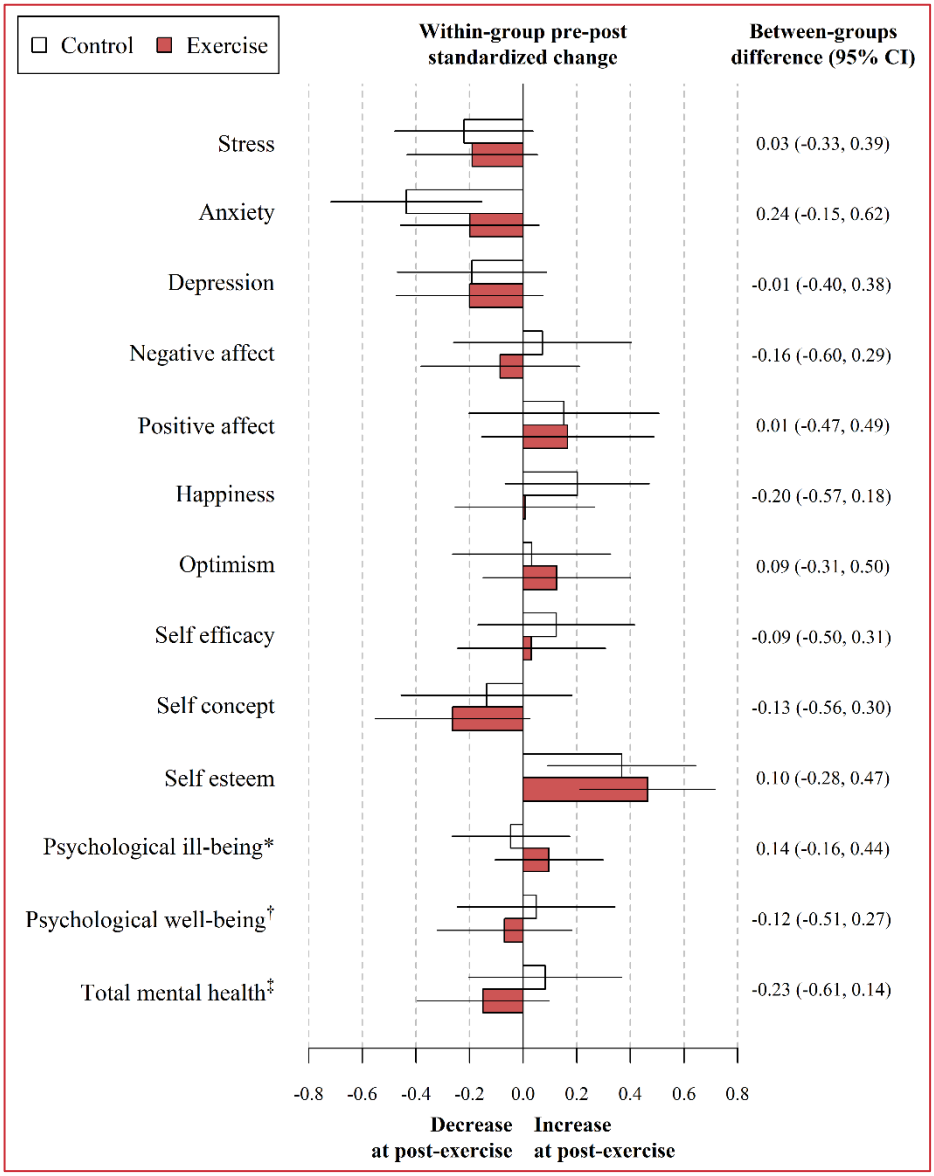


Figure 45
Effects of the ActiveBrains exercise program z-score pre-post change between groups in mental health (per-protocol analyses).
Data analyses were primarily conducted under the per-protocol principle, i.e., attending to 70% of the sessions. Baseline z-score of the outcomes were calculated by subtracting the mean value and dividing by the SD of each outcome. Post-exercise z-scores were calculated relative to the mean and SD of the baseline values, being a z-score of the change in each outcome, i.e., (post-exercise$_i$ – baseline mean) / baseline SD.
*Psychological ill-being was calculated as the normalized mean of the z-score for stress, anxiety, depression, and negative affect.
†Psychological well-being was calculated as the normalized mean of the z-score for positive affect, happiness, optimism, self-efficacy, self-concept, and self-esteem.
‡ Total mental health was calculated as the normalized mean of the z-score for all mental health indicators.

Children in the ActiveBrains exercise program improved the body composition by reducing their total and visceral fat mass. These results are in line with the previous literature in children with obesity regarding the reductions in BMI and fat mass [447–450]. Likewise, a recent network meta-analysis concluded that aerobic or the combined aerobic and resistance training effectively reduced adiposity outcomes with similar magnitude as we observed in our study (BMI ~0.7 kg/m^2 vs. our findings: 0.6 kg/m^2) in children and adolescents with overweight or obesity [464]. No less important, we found that a higher rate of participants experienced meaningful reductions

in their fat mass (i.e., 79% in the exercise vs. 36% in the control group), which is in line with the EFIGRO trial findings [454]. Otherwise, the lean mass index was not affected by the ActiveBrains exercise program. This is in line with a previous study using a similar indicator of lean mass [447]. However, another RCT in children with obesity with a similar dose of resistance training described improvements in FFM (+1.2 kg compared to controls) [450].

Regarding physical fitness, the ActiveBrains exercise program improved CRF both the performance in the test (laps) and estimated VO_2max. These results agree with previous trials in children with overweight or obesity, independently of the exercise protocol followed [447–450]. The ActiveBrains exercise program did not improve the children's speed-agility or muscular fitness, which agrees with the EFIGRO trial findings, which used a similar exercise protocol in a sample of similar characteristics to ours [447]. The specificity of the resistance exercises performed might explain this finding, i.e., body-weight exercises instead of weightlifting may have produced benefits in muscular endurance instead of maximal strength or power (as measured by the handgrip and the standing long jump tests). Although the absolute muscular fitness was reduced in the exercise group compared to the control, this difference disappeared after using a score relative to the body weight. This is in line with the lack of effects of the ActiveBrains exercise program on the lean mass index. Thereby, given that not only CRF but also other fitness components such as speed-agility or muscular fitness are relevant for health later in life [23], further trials targeting speed-agility and muscular fitness in children with overweight or obesity are needed. As previously found [454], the participant rate meaningfully improving CRF was higher in the exercise group compared to controls.

None of the previous studies analyzed the effects of the exercise program on the composite MetS risk scores, which hampers comparisons in this regard. We believe it is a strength of our study to quantify the effect on the composite MetS risk scores, which are valid measures of risk for T2D, CVD, and other cardiometabolic diseases [452,453]. Our findings are further strengthened by the investigation of the proportion of children experiencing meaningful changes in the control and the exercise groups. Using a similar approach, the EFIGRO trial found higher rates of meaningful changes in hepatic fat [454].

Mental health

Furthermore, the ActiveBrains exercise program did not affect mental health (i.e., psychological ill-being and well-being) in children with overweight or obesity. These results agree with a recent meta-analysis in which no overall effect of exercise on mental health in children was described (effect size = 0.209, P = 0.141) [451].

Abbreviations in this page:
CRF: cardiorespiratory fitness
CVD: cardiovascular disease
FFM: fat-free mass
MetS: metabolic syndrome
RCT: randomized controlled trial
T2D: type 2 diabetes
VO_2max: maximum oxygen consumption

Abbreviations in this page:
CRF: cardiorespiratory fitness
CVD: cardiovascular disease
MetS: metabolic syndrome
RCT: randomized controlled trial
T2D: type 2 diabetes

However, the studies included examining the effect of exercise in children with overweight or obesity showed inconsistent findings [466,467]. Seabra et al. concluded that a 20-week football program improved self-esteem in boys with overweight [466]. Alternatively, Romero-Perez et al., found no significant changes after 20-week aerobic exercise training in anxiety and a small reduction in depression in children with obesity [467]. Differences in our findings and the previous studies could be explained by the heterogeneity of the exercise program (type, only aerobic vs. aerobic and resistance training; and frequency, 2 vs. 3 to 5 sessions per week), characteristics of the study sample (sex, weight status), the mental health outcomes examined (individual dimensions vs. a complete set of psychological ill-being and well-being outcomes); and the study design (non-RCT vs. RCT).

Although our intervention is in line with the principles defined by Lubans et al. [455] to maximize the effects of exercise on mental health, we believe that the lack of a protocol to control whether the sessions were conducted under these principles (to adapt the sessions accordingly if needed) could explain the lack of effect observed. Likewise, it is possible that the large number of questionnaires implied a high burden to participants. Since the mental health of children with obesity is likely impaired [47,50], further studies proposing effective lifestyle interventions to improve their mental health are urgently needed. We recommend that future trials monitor the acute effects on mental health in every session to find the best type, frequency, intensity, and duration of exercises to target mental health. Over time, high-intensity interval training might be a good strategy for improving mental health [468].

Practical implications

First, our study is the first quantifying the risk reduction for MetS produced by an exercise program in childhood obesity. Children with overweight or obesity are at high risk for future T2D and CVD [48]. Here, we demonstrate that a combined aerobic and resistance training exercise program reduced risk of MetS by around 30% in these children, probably driven by improvements in the blood lipids biomarkers, total and visceral adiposity, and CRF [469,470]. Another important implication of our study relates to the mental health benefits of exercise. Lubans et al. already described that exercise on its own may not be enough to improve mental health, and they proposed a series of characteristics of the exercise that should be followed to maximize the effects [455]. Based on the study findings, i.e., no effect on mental health, and the experience after carrying out the ActiveBrains RCT, we would recommend the monitoring of the exercise sessions (i.e., intensity and methodology) and examination of mental health outcomes after each session to adapt the

exercise program as needed. Last, our study sheds light on the exercise-induced changes in the time devoted to MVPA, LPA, SB, and sleep in children. We found that the exercise program promoted a healthier daily time-use composition, which is important for public health strategies by supporting that children with overweight or obesity might benefit from exercise programs by increasing the time spent in MVPA and decreasing the health consequences of SB.

Abbreviations in this page:
LPA: light physical activity
MetS: metabolic syndrome
MVPA: moderate-to-vigorous physical activity
SB: sedentary behaviour
T2D: type 2 diabetes

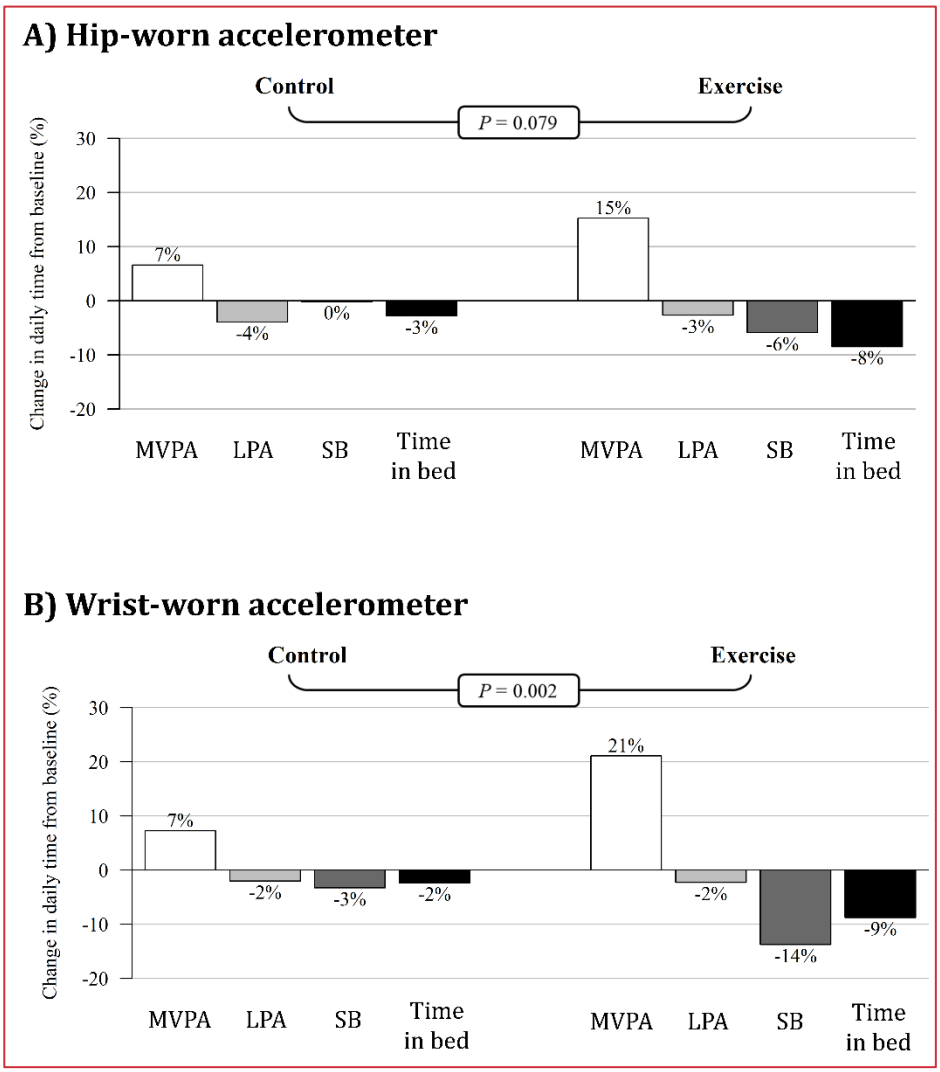


Figure 46
Evidence of exercise program fidelity based on behavioral changes assessed from hip- (panel A) and wrist-worn (panel B) accelerometers observed from baseline to during the implementation of the ActiveBrains exercise program.
Data analyses were primarily conducted under the per-protocol principle, i.e., attending to 70% of the sessions. Isometric log-ratios between each group's compositional mean and the overall compositional mean after centering the data at baseline and during exercise were calculated.
P value from Hotelling's T-squared test for pair-wise comparison of multivariate means.

Limitations and strengths

The strengths of our study include: the holistic view of both physical and mental health with a complete array of outcomes in children with obesity; the quantification of two composite MetS risk scores; the description of the distribution of the meaningful change at individual level; the analysis of the exercise-induced changes in the daily time devoted to MVPA, LPA, SB and sleep using objective measures; the use of gold-standard measures of cardiometabolic health and body composition, reliable and valid physical fitness tests and mental health questionnaires; and the heart rate monitoring using individualized and 'a priori' programmed heart rate monitors. However, there are some limitations that should be noted. These study findings might be limited by the relatively low sample size for

Abbreviations in this page:
CRF: cardiorespiratory fitness
DXA: dual-energy X-ray absorptiometry
MetS: metabolic syndrome

some outcomes, which could make some of the statistical analyses underpowered to detect significant differences and by the fact that some of the evaluators were not blinded to the group allocation. We believe that most of the outcomes included in our study are objective or contained a large number of objective elements which are unlikely to be influenced by assessor blinding (i.e., cardiometabolic health, blood markers assessed in external laboratory, and body composition by DXA). However, it is possible that the lack of findings in mental health is explained by the large number of questionnaires (although they were collected in three different days to lower the participants' burden).

Conclusion

The ActiveBrains exercise program improved cardiometabolic health in children with overweight or obesity. The MetS risk score might be reduced by around 30%, which could be due to improvements in blood lipids, total and visceral adiposity, and CRF. The lack of the effect on mental health need to focus more attention on exercise environment and delivery. These findings support public health initiatives promoting exercise programs in children with obesity to reduce their risk for later comorbidities.

LIMITATIONS AND STRENGTHS

"There is no such thing as simple. Simple is hard"

Martin Scorsese

CONTENTS

Limitations

Abbreviations in this page:
PA: physical activity
PAEE: physical activity-related energy expenditure
SB: sedentary behaviour

The findings of the present Doctoral Thesis should be interpreted with caution due to a number of limitations. Specific limitations of each study are presented in the Results and Discussion section of each study and an overall view of the main limitations is presented here.

Section I includes a systematic review, a software description article, four observational studies, and an expert consensus statement. The systematic review (**Study I**) set basis for the rest of the thesis describing the state-of-the-art in the measurement of physical behaviours (i.e., PA, SB, and sleep) with accelerometers, and identifying some research gaps that were approached in this Thesis. Notably, this review was limited by the only inclusion of one accelerometer model (ActiGraph GT3X/+), other brands and models were omitted by inclusion criteria. This could limit our understanding of the research field, yet this is unlikely given that ActiGraph was the brand by far most used in the field at that time and roughly 200 studies were included in the review [41]. An update including not only ActiGraph, but also other brands and open-access algorithms would provide useful information and make the field closer to a harmonization of accelerometer data collection and processing protocols. The **Study II** describes the software used for the accelerometer data processing in this Thesis (i.e., GGIR). The GGIR software could be limited in some functionalities, such as the lack of a validated algorithm to estimate sleep behaviours for children or the activity-rest indicators. However, GGIR continuously grows upon the research field demands with open-source contribution and usability policies.

Studies III-VI are a set of cross-sectional analyses investigating the capacity of acceleration metrics to estimate PAEE and comparing accelerometer data derived from different data collection and processing protocols. The fact that they are cross-sectional studies is less of a concern given the research objectives of these studies. However, longitudinal designs could have opened the venue to the investigation of accelerometer data variability along time. Likewise, the participants differed across studies, limiting the generalization of findings from some studies to others. In this regard, the estimation of PAEE from acceleration metrics (Study III) was performed in 5-year olds, which limited the usability of the PAEE equations in the other studies included in this Thesis. The rest of studies in the thesis focused on the participants of the ActiveBrains project (8-11-year olds with overweight or obesity). Although the sample is younger in Study III, this study provides valuable information and addressed a research gap in the field related to PAEE estimation from wrist accelerometers. The fact that the Study IV was conducted in young adults

Abbreviations in this page:
ENMO: Euclidean norm minus 1*G*
PA: physical activity
RCT: randomized controlled trial
SB: sedentary behaviour
VMCounts: activity counts in the vector magnitude

is less concerning since our aim here related to the comparability of the acceleration metrics, being less relevant the sample characteristics as long as they provide a wide range of accelerations in which the metrics were tested.

Another limitation that was repeated in Studies IV-VI is the lack of a criterion measure of the physical behaviours (i.e., PA, SB, and sleep) or related outcomes (e.g., steps). Under that constrain, our studies were limited to the comparability of the different protocols, metrics, and descriptors. However, we could not make inferences on what specific protocol, metric, or descriptor was the most accurate to quantify physical behaviours. Likewise, study V only included published ENMO and VMCounts cut-points, which could be complemented now by recently published cut-points for other acceleration metrics. Given that we used cross-sectional designs, we could not test the reliability and consistency of these protocols, metrics, and descriptors. This limitation may be also considered in the studies in Section II. Our decisions on the specific protocols followed in each study of Section II were made based on the specific requirements of the research question to answer in each study. Additionally, sensitivity analyses on how different data collection and processing protocols could have affected the specific findings of each study were carried out.

Similarly, **Section II** has several limitations to be considered. The cross-sectional design of **Studies VIII-X** limits the causation interpretation. Thus, the associations observed between physical behaviours and brain health in this Thesis should be considered with caution and further confirmed in well-designed RCTs. Additionally, some of the analyses performed may lack enough statistical power given the limited sample size. In this regard, the ActiveBrains RCT was powered to find low-to-medium effect sizes on its primary research question (i.e., effects of exercise on brain outcomes, academic achievement, and cognition). However, some analyses cross-sectional analyses (especially the stratification of sample as in the **Study VIII**) on secondary outcomes (or some) might be underpowered. Thus, we may be ignoring some associations that future studies should investigate. Likewise, the use of accelerometers to quantify sleep-related behaviours might be criticized by the fact that we are not actually measuring sleep, but we estimate it from movement patterns. However, accelerometers are the less-invasive objective method to assess sleep behaviours in free-living settings, while also providing good validity [38,40]. Sleep periods during the day (naps) are ignored in this Thesis upon the unavailability of an algorithm providing valid and objective measures of this behaviour. Finally, the ActiveBrains data collection implied a high burden to participants, primarily given the number of questionnaires they had to fill in. This

might affect those variables derived from questionnaires, where the subjectivity might produce inaccuracies. However, this is not a concern for many of the variables in this Thesis (cardiometabolic variables, body composition, physical activity) as they were objectively collected; to note that we separated the assessments in five different days to reduce the participants' burden.

Strengths

Notwithstanding, an overall view of the strengths of this Thesis is worth mentioning. First, both **Section I** and **Section II** benefit from a variety of study designs (i.e., systematic review, cross-sectional studies, consensus statement, and RCT) well complemented to reach meaningful conclusions for the PA promotion in children with overweight or obesity. Also relative to both sections, participants in all the studies reached high wear times with the accelerometers, which allowed for a better representation of their movement along the day. Likewise, the use of accelerometers to quantify physical behaviours, as the less invasive tool to objectively quantify physical behaviours in free living is another important strength of the Thesis.

Specifically, in Section I, a major strength arises from the systematic review (**Study I**) as it is a comprehensive study that summarizes information from the methodologies of a large number of studies in separate sections for age groups. This allows researchers to gain insights in the comparability of different methods and access meaningful information for assistance in the selection of specific data collection and processing decisions for their sample and research question characteristics. The GGIR software (described in **Study II**) is another strength of this thesis as it allows for transparent (open-access) data processing of the raw accelerometer data. The use of doubly-labelled water, which is a gold standard for the measurement of PAEE in **Study II** is another important strength to consider. Besides, the in-depth investigation of different data processing protocols and their comparability is a step forward to reach the harmonization of findings in the field (**Studies III-VI**). In this sense, we included from a variety of acceleration metrics (e.g., ENMO, MAD, VMCounts) to inferences of specific behaviours, such as MVPA, sleep, or steps among others.

In **Section II**, gold-standard measurement techniques were used for GMV (MRIs, **Studies VIII-IX**), cardiometabolic health and body composition (blood samples, DXA, **Study XI**). This provide accurate and reliable measures of the cardiometabolic and brain health of our participants. Similarly, the use of standard measures of academic achievement, executive function, and IQ is preferable over other measures that are more likely to be affected by subjectivity (such as academic grades). Another strength is the use of advanced

Abbreviations in this page:
DXA: dual-energy x-ray absorptiometry
ENMO: Euclidean norm minus 1G
GMV: grey matter volume
MAD: mean amplitude deviation
MRI: magnetic resonance image
MVPA: moderate-to-vigorous physical activity
PA: physical activity
PAEE: physical activity-related energy expenditure
RCT: randomized controlled trial
SB: sedentary behaviour
VMCounts: activity counts in the vector magnitude

analytical approaches (e.g., compositional data analysis, multivariate pattern analysis, mediation models), which allows appropriate conclusions by handling the closure and multicollinearity often observed among accelerometer data descriptors. Finally, the focus on children with overweight or obesity, which can be greatly benefited from the conclusions obtained in this Thesis including a complete array of cardiometabolic and brain health outcomes.

FUTURE RESEARCH DIRECTIONS

“I don't know where I'm going from here, but I promise it won't be boring”

David Bowie

CONTENTS

Looking back while moving forward

Abbreviations in this page:
PA: physical activity

As in sculling, the 'boat' containing the research on PA measurement and on associations of PA with health should move forward while the rowers (researchers) keep an eye to backwards. I borrow this metaphor, which I completely agree with, from Prof. I-Min Lee, one of the most influential epidemiologists in PA and health of the last decades.

The **Section I** of this Thesis focuses on PA measurement with accelerometers, the method of preference to objectively monitor physical behaviours in free-living settings at the moment. May the **Study I** serve as the look to the past with a comprehensive systematic review on the data collection and processing decisions of choice during the last years in the field. Main advantages and limitations of the chosen decisions were described, which allowed us to define some future directions:

- There was a need for further research on the comparability of data collected from different body attachment sites, being the hip and wrists the preferred by the field at that time.
- The field primarily relied on a single manufacturer of accelerometers (ActiGraph), which comes with proprietary data processing methods conducted. Thus, another important conclusion for future research was the need of open-source algorithms to process raw accelerometer data and obtain consistent physical behaviour estimates.
- We also observed a clear under-reporting of the data processing protocols followed in many of the studies included in the systematic review. We recommend future studies in the field to provide a clear description on how the data were processed to enhance data comparability and reproducibility of findings.
- At the moment, the field would benefit from an update of this systematic review, which should be expanded to other accelerometer brands and open-access algorithms.

Some of these directions were followed by the rest of the studies included in this Thesis. Specifically, **Section I** of this Thesis contributes to the field by making comparing backward (proprietary information) algorithms with newer (open-access) algorithms to process raw accelerometer data and gain insight into physical behaviours quantification. The following research directions are recommended attending at the **Section I** studies' findings:

Abbreviations in this page:
PA: physical activity
PAEE: physical activity-related energy expenditure
RCT: randomized controlled trial

- The usefulness of open-access metrics is supported in future studies given that we observed higher correlations between such metrics and PAEE (measured with a gold-standard) than using proprietary activity counts (**Study III**). However, future studies should also confirm this information in other cohorts of varying characteristics.
- Likewise, **Study IV** demonstrated that the movement pattern identified by different acceleration metrics is highly comparable, unless aggressive filtering techniques are used. This also supports the use of acceleration metrics and encourage more studies on whether the movement pattern associations with health vary upon the use of different acceleration metrics.
- Dominant-wrist cut-points for PA intensity classification in adults were missing in the field, and **Study IV** provided a comparable set of cut-points to others previously published in non-dominant wrist data. However, future studies should cross-validate these cut-points against indirect calorimetry to really understand their usefulness.
- **Study V** shows that PA intensity and SB estimated from different cut-points are hardly comparable. Further data harmonization efforts are needed, as well as meta-analyses using data from cut-points validation studies to propose a consensual set of cut-points to be used in different settings/projects.
- One *step* further in the investigation of different acceleration metrics that provide meaningful information on PA is done in **Study VI**. Step-based metrics are a good proxy to PA levels in children with overweight or obesity, and walking seems to be a feasible intervention alternative to meaningfully increase PA in children with overweight or obesity. However, further RCTs using on walking are needed to confirm this finding.

The **Study VII** serves as link between physical behaviour measurements and the investigation of their associations with health. The sculling metaphor is still present in this study, in which we summarize the analytical approaches more frequently-used in the field (look backward) while proposing the research field needs and recommend future steps (moving forward). In this regard, the authors of this consensus article agree that investigations determining associations between physical behaviours and health should be extended to understand how is the physical behaviours interplay in their relationship

with health outcomes. We defined a short- and long-term agenda for the field, which is summarized as follows:

- Clear communication on the rationale and limitations of the analytical approaches used in studies.
- Investigation of new analytical approaches to handle the accelerometer data singularities (primarily closure and collinearity between variables).
- Triangulation of results from different analytical approaches might be the best alternative at the moment to quantify the associations of physical behaviours with health outcomes. Following the decision tree designed in this consensus article (**Figure 32**) may assist researcher on the selection of analytical approaches for a given research question.
- Investigation of machine learning for diagnostic/prognostic purposes is encouraged in the field.

Finally, **Section II** is built upon the information obtained from the studies in **Section I**. All the data collection and processing decisions, as well as the analytical approaches used, were made based on the findings obtained in the previous studies. Future directions can be also derived from the Section II studies in relation to the associations of physical behaviours with cardiometabolic and brain health in children with overweight or obesity:

- First, we discussed previous findings on the lack of associations between PA and hippocampal GMV in children (**Study VIII**). We observed than using appropriate analytical approaches, and considering the weight status as moderator, some associations were observed. This finding is not even close to be conclusive, yet it opens a conversation worth exploring in future studies.
- We also observed that some sleep-related behaviours were associated with GMV in several cortical brain regions and the hippocampus, and this seemed to impact the children's academic achievement and IQ (**Study IX**). This may open the venue for well-designed RCTs to investigate strategies to improve sleep quality in children and their impact on their brain health.
- We also investigated the activity-rest pattern, its stability and fragmentation, in relation to the brain health of children with overweight or obesity (**Study XI**). Future RCTs should consider stable and timely-appropriate exercise interventions when targeting brain health in children with overweight or obesity.

Likewise, we conducted a RCT on the effects of a 20-week exercise program on the cardiometabolic and mental health of children

Abbreviations in this page:
GMV: grey matter volume
IQ: intelligence quotient
PA: physical activity
RCT: randomized controlled trial

Abbreviations in this page:
PA: physical activity
RCT: randomized controlled trial

with overweight or obesity (**Study XI**). From this study, we derive the following recommendations for future research:

- We used accelerometers in investigate the exercise program fidelity in **Study XI**. Monitoring physical behaviours patterns at baseline and during the exercise program implementation allows to quantify the participants compliance with the exercise program, as well as the contamination (likelihood of controls to engage on PA during the experiment) and compensation (likelihood of exercises of quitting from other daily living PA during the experiment) effects. Future studies should consider this to measure the intervention fidelity.
- We also believe that including a post-session acute evaluation of some mental health outcomes would have benefitted our exercise program by allowing to adapt the exercise program to the participants' needs. Future RCTs targeting mental health through exercise interventions should consider this.

Last but not least, there is a need of well-designed RCTs in children with overweight or obesity to confirm the findings presented in this Thesis.

CONCLUSIONS

"If you can't explain it simply, you don't understand it well enough"

Albert Einstein

CONTENTS

General conclusion

The conclusions from this International Doctoral Thesis are of value for the PA measurement and the PA epidemiology research fields. **Section I** provides valuable recommendations on best-practice accelerometer data collection and processing techniques to measure physical behaviours in free-living settings, as well as consensus recommendations on analytical approaches for the field of PA epidemiology. Moreover, this Doctoral Thesis highlights the value of open-source data processing algorithms for the field, being more comparable and reproducible than proprietary algorithms. Likewise, The **Section II** demonstrates the important role of PA, SB, sleep, and the activity-rest pattern in relation with brain health outcomes in children with overweight or obesity. Finally, this Doctoral Thesis has demonstrated that meaningful and positive changes in cardiometabolic health in children with overweight or obesity can be obtained with a 20-week exercise program, which should inform future health programs.

Abbreviations in this page:
PA: physical activity
PAEE: physical activity-related energy expenditure
TEE: total energy expenditure
VMCounts: activity counts in the vector magnitude

Specific conclusions

The specific conclusions reached in the studies included in this International Doctoral Thesis are detailed as follows:

1. The systematic review (**Study I**) on accelerometer data collection and processing decisions provide practical considerations for the decision making based on existing literature. Importantly, researchers in the field should make their decisions depending on the population's age. Likewise, when selecting a specific cut-point or algorithm, it is important to apply the same criteria as in the original validation/calibration study.
2. GGIR (described in **Study II**) provides a tool for researchers to derive variables that characterize PA and sleep assessed in an objective manner. As a PhD student, I have actively helped to check and improve GGIR, in particular the time-use analysis functionality.
3. A higher performance to predict PAEE and TEE was observed for the open-source acceleration metrics versus VMCounts (**Study III**). Open-source acceleration metrics from the wrist explained up to 84% of the variance in TEE and 67% in PAEE after adjustments for weight and height. Overall, open-source metrics explained around 10-20% more of the variation in TEE and PAEE than VMCounts.
4. Higher acceleration metric values were observed in the dominant wrist versus the non-dominant wrist in **Study IV**). ENMO and LFENMO were the metrics that

Abbreviations in this page:
CPM: counts per minute
ENMO: Euclidean norm minus 1 *G*
GMV: grey matter volume
LFENMO: ENMO of the low-pass filtered raw accelerations
LMVPA: light-moderate-vigorous physical activity
MAD: mean amplitude deviation
MVPA: moderate-to-vigorous physical activity
PA: physical activity
PAEE: physical activity-related energy expenditure
SB: sedentary behaviour
TEE: total energy expenditure
VMCounts: activity counts in the

compared the best, and to some extent, they also showed good comparability with MAD for daily average values and for the movement pattern identified throughout the day. However, VMCounts were demonstrated to be less comparable to the previously-mentioned metrics.

5. Large discrepancies were observed in the time spent in SB and PA intensities across cut-points relative to different body attachment sites and acceleration metrics in overweight or obese children (**Study V**). Furthermore, we provide a comprehensive comparison between available cut-points in order to better understand which cut-points provide comparable results and which ones not. Also, our data showed that it is not currently possible (and probably will never be) to know the prevalence of a population meeting the PA guidelines based on accelerometer data, with differences from nearly none to nearly everyone meeting the guidelines.
6. Step-based metrics including steps/day and various cadence-based intensity indicators seem to capture the majority of PA (as recorded by daily CPM, LMVPA, and MVPA with accelerometers) in children with overweight or obesity (**Study VI**).
7. The expert consensus article provided a comprehensive description of the analytical approaches most-frequently used in the field to investigate the associations of physical behaviours (i.e., PA, SB, and sleep) with health (**Study VII**). Advantages and limitations of each approach are exposed and practical recommendations on the best-suited approaches for a given research question are recommended. The authors also agreed on a set of consensus points and research needs which are relevant for the physical behaviour epidemiology.
8. Findings from the Study VIII indicate that PA and SB were not associated with GMV in the hippocampus in children with overweight or obesity. However, we found some evidence of moderation by weight status in the associations, so that reducing SB and engaging in more MVPA were associated with greater GMV in the right hippocampus. Specifically, reallocating 20 min/day from SB to MVPA would be associated with 100 mm^3 more GMV in the right hippocampus in children with obesity type I.
9. Sleep behaviours, including timing, duration, and patterns, were associated with GMV and, subsequently,

GMV was associated with academic achievement and IQ in children with overweight or obesity (**Study IX**). Total sleep time, sleep efficiency and WASO time seem to be specifically associated with the right hippocampus, but this subcortical region did not associate with academic achievement, executive function or IQ. Sleep behaviours seem important for GMV and academic achievement and, to a lesser extent, for IQ, but they were not associated with executive function. These associations were independent of SB and PA intensity.

10. A more stable and less fragmented activity-rest pattern is associated with better academic achievement, executive function, and IQ in children with overweight or obesity (**Study X**). Likewise, earlier PA occurrence is preferable for academic achievement, executive function and IQ in this population. These associations seemed not to be mediated by overall brain GMV.
11. The ActiveBrains exercise program improved cardiometabolic health in children with overweight or obesity. The MetS risk score might be reduced by around 30%, which could be due to improvements in blood lipids, total and visceral adiposity, and CRF. The lack of effect on mental health need to focus more attention on exercise environment and delivery.

All these findings support public health initiatives promoting exercise programs in children with overweight or obesity, as well as they provide meaningful information for exercise program planification.

Abbreviations in this page:
CRF: cardiorespiratory fitness
GMV: grey matter volume
IQ: intelligence quotient
MetS: metabolic syndrome
PA: physical activity
PAEE: physical activity-related energy expenditure
SB: sedentary behaviour
TEE: total energy expenditure
WASO: wakening after sleep onset

REFERENCES

"If I have seen further it is by standing on the shoulders of Giants"

Isaac Newton

1 Juvenal. *The satires of Juvenal translated: with explanatory and classical notes, relating to the laws and customs of the Greeks and Romans*. Dublin : printed by George Faulkner, 1769. 1769. https://search.library.wisc.edu/catalog/9966326063602122

2 Morris JN, Heady JA, Raffle PA, *et al.* Coronary heart-disease and physical activity of work. *Lancet (London, England)* 1953;**262**:1111–20; concl. doi:10.1016/s0140-6736(53)91495-0

3 Warburton DER, Nicol CW, Bredin SSD. Health benefits of physical activity: the evidence. *CMAJ* 2006;**174**:801–9. doi:10.1503/cmaj.051351

4 Janssen I, LeBlanc AG. Systematic review of the health benefits of physical activity and fitness in school-aged children and youth. *Int J Behav Nutr Phys Act* 2010;**7**. doi:10.1186/1479-5868-7-40

5 Fletcher GF, Balady G, Blair SN, *et al.* Statement on exercise: benefits and recommendations for physical activity programs for all Americans. A statement for health professionals by the Committee on Exercise and Cardiac Rehabilitation of the Council on Clinical Cardiology, American Heart Associ. *Circulation* 1996;**94**:857–62. doi:10.1161/01.cir.94.4.857

6 American College of Sports Medicine A. *Guidelines for graded exercise testing and exercise prescription*. Philadelphia: 1975.

7 HHS USD of H and HS. *2008 Physical activity guidelines for Americans*. 2008. www.health.gov/paguidelines

8 WHO. *Global recommendations on physical activity for health*. 2010. doi:10.1080/11026480410034349

9 Hallal PC, Andersen LB, Bull FC, *et al.* Global physical activity levels: Surveillance progress, pitfalls, and prospects. *Lancet* 2012;**380**:247–57. doi:10.1016/S0140-6736(12)60646-1

10 Lee I-M, Shiroma EJ, Lobelo F, *et al.* Effect of physical inactivity on major non-communicable diseases worldwide: an analysis of burden of disease and life expectancy. *Lancet (London, England)* 2012;**380**:219–29. doi:10.1016/S0140-6736(12)61031-9

11 Pandey A, Salahuddin U, Garg S, *et al.* Continuous dose-response association between sedentary time and risk for cardiovascular disease a meta-analysis. *JAMA Cardiol* 2016;**1**:575–83. doi:10.1001/jamacardio.2016.1567

12 Rezende LFM, Sá TH, Mielke GI, *et al.* All-Cause Mortality Attributable to Sitting Time: Analysis of 54 Countries Worldwide. *Am J Prev Med* 2016;**51**:253–63. doi:10.1016/j.amepre.2016.01.022

13 Covassin N, Singh P. Sleep Duration and Cardiovascular Disease Risk: Epidemiologic and Experimental Evidence. *Sleep Med Clin* 2016;**11**:81–9. doi:10.1016/j.jsmc.2015.10.007

14 Bussmann JBJ, van den Berg-Emons RJG. To total amount of activity….. and beyond: perspectives on measuring physical behavior. *Front Psychol* 2013;**4**:463. doi:10.3389/fpsyg.2013.00463

15 Ortega FB, Konstabel K, Pasquali E, *et al.* Objectively Measured Physical Activity and Sedentary Time during Childhood, Adolescence and Young Adulthood: A Cohort Study. *PLoS One* 2013;**8**. doi:10.1371/journal.pone.0060871

16 Riddoch CJ, Andersen LB, Wedderkopp N, *et al.* Physical Activity Levels and Patterns of 9- and 15-yr-Old European Children. *Med Sci Sports Exerc* 2004;**36**:86–92. doi:10.1249/01.MSS.0000106174.43932.92

17 Pedišić Ž. Measurement issues and poor adjustments for physical activity and sleep undermine sedentary behaviour research - The focus should shift to the balance between sleep, sedentary behaviour, standing and activity. *Kinesiology* 2014;**46**:135–46.

18 Caspersen CJ, Powell KE, Christenson GM. Physical activity, exercise, and physical fitness: definitions and distinctions for health-related research. *Public Health Rep* 1985;**100**:126–31.

19 Tremblay MS, Aubert S, Barnes JD, *et al.* Sedentary Behavior Research Network (SBRN) - Terminology Consensus Project process and outcome. *Int J Behav Nutr Phys Act* 2017;**14**:1–17. doi:10.1186/s12966-017-0525-8

20 Vähä-Ypyä H, Husu P, Suni J, *et al.* Reliable recognition of lying, sitting, and standing with a hip-worn accelerometer. *Scand J Med Sci Sport* 2018;**28**:1092–102. doi:10.1111/sms.13017

21 Krueger JM, Frank MG, Wisor JP, *et al.* Sleep function : Toward elucidating an enigma. *Sleep Med Rev* 2016;**28**:46–54. doi:10.1016/j.smrv.2015.08.005

22 Ortega FB, Sui X, Lavie CJ, *et al.* Body mass index, the most widely used but also widely criticized index: would a criterion standard measure of total body fat be a better predictor of cardiovascular disease mortality? *Mayo Clin Proc* 2016;**91**:443–55. doi:10.1016/j.mayocp.2016.01.008

23 Ortega FB, Ruiz JR, Castillo MJ, *et al.* Physical fitness in childhood and adolescence: A powerful marker of health. *Int J Obes* 2008;**32**:1–11. doi:10.1038/sj.ijo.0803774

24 Erickson KI, Hillman C, Stillman CM, *et al.* Physical Activity, Cognition, and Brain Outcomes: A Review of the 2018 Physical Activity Guidelines. *Med Sci Sports Exerc* 2019;**51**:1242–51. doi:10.1249/MSS.0000000000001936

25 Duijvestijn M, van den Berg SW, Wendel-Vos GCW. Adhering to the 2017 Dutch Physical Activity Guidelines: A Trend over Time 2001-2018. *Int J Environ Res Public Health* 2020;**17**. doi:10.3390/ijerph17030681

26 Van Domelen DR, Koster A, Caserotti P, *et al.* Employment and physical activity in the U.S. *Am J Prev Med* 2011;**41**:136–45. doi:10.1016/j.amepre.2011.03.019

27 Kwak L, Berrigan D, Van Domelen D, *et al.* Examining differences in physical activity levels by employment status and/or job activity level: Gender-specific comparisons between the United States and Sweden. *J Sci Med Sport* 2016;**19**:482–

7. doi:10.1016/j.jsams.2015.05.008
28 Epstein L, Miller GJ, Stitt FW, *et al.* Vigorous exercise in leisure time, coronary risk-factors, and resting electrocardiogram in middle-aged male civil servants. *Br Heart J* 1976;**38**:403–9. doi:10.1136/hrt.38.4.403
29 Strath SJ, Kaminsky LA, Ainsworth BE, *et al.* Guide to the assessment of physical activity: Clinical and research applications: A scientific statement from the American Heart association. *Circulation* 2013;**128**:2259–79. doi:10.1161/01.cir.0000435708.67487.da
30 Baranowski T, Dworkin RJ, Cieslik CJ, *et al.* Reliability and Validity of Self Report of Aerobic Activity: Family Health Project. *Res Q Exerc Sport* 1984;**55**:309–17. doi:10.1080/02701367.1984.10608408
31 Sirard JR, Pate RR. Physical activity assessment in children and adolescents. *Sport Med* 2001;**31**:439–54. doi:10.2165/00007256-200131060-00004
32 Sallis JF. Self-report measures of children's physical activity. *J Sch Health* 1991;**61**:215–9. doi:10.1111/j.1746-1561.1991.tb06017.x
33 Ainsworth B, Cahalin L, Buman M, *et al.* The Current State of Physical Activity Assessment Tools. *Prog Cardiovasc Dis* 2015;**57**:387–95. doi:10.1016/j.pcad.2014.10.005
34 Welk GJ, Kim Y. Context of Physical Activity in a Representative Sample of Adults. *Med Sci Sports Exerc* 2015;**47**:2102–10. doi:10.1249/MSS.0000000000000641
35 Tudor-Locke C, Ham SA. Walking Behaviors Reported in the American Time Use Survey 2003–2005. *J Phys Act Heal* 2008;**5**:633–47. doi:10.1123/jpah.5.5.633
36 Crowley P, Skotte J, Stamatakis E, *et al.* 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* 2019;**16**:65. doi:10.1186/s12966-019-0835-0
37 Freedson PS, Melanson E, Sirard JR. Calibration of the computer science and applications, Inc. accelerometer. *Med Sci Sports Exerc* 1998;**30**:777–81. doi:10.1097/00005768-199805000-00021
38 Sadeh A, Sharkey KM, Carskadon MA. Activity-based sleep-wake identification: an empirical test of methodological issues. *Sleep* 1994;**17**:201–7. doi:7939118
39 Cole RJ, Kripke DF, Gruen W, *et al.* Automatic sleep wake identification from wrist activity - Cole et al 1992.pdf. Sleep. 1992;**15**:461–9.
40 Van Hees VT, Sabia S, Anderson KN, *et al.* A novel, open access method to assess sleep duration using a wrist-worn accelerometer. *PLoS One* 2015;**10**:1–13. doi:10.1371/journal.pone.0142533
41 Wijndaele K, Westgate K, Stephens SK, *et al.* Utilization and Harmonization of Adult Accelerometry Data: Review and Expert Consensus. *Med Sci Sports Exerc* 2015;**47**:2129–39. doi:10.1249/MSS.0000000000000661
42 MacMahon S, Baigent C, Duffy S, *et al.* Body-mass index and cause-specific mortality in 900 000 adults: Collaborative analyses of 57 prospective studies. *Lancet* 2009;**373**:1083–96. doi:10.1016/S0140-6736(09)60318-4
43 Bray GA, Frühbeck G, Ryan DH, *et al.* Management of obesity. *Lancet (London, England)* 2016;**387**:1947–56. doi:10.1016/S0140-6736(16)00271-3
44 World Health Organization. *Global Health Estimates 2016: Disease burden by Cause, Age, Sex, by Country and by Region, 2000-2016.* Geneva: : World Health Organization 2018.
45 Weihrauch-Blüher S, Schwarz P, Klusmann JH. Childhood obesity: increased risk for cardiometabolic disease and cancer in adulthood. *Metabolism* 2019;**92**:147–52. doi:10.1016/j.metabol.2018.12.001
46 Ortega FB, Lavie CJ, Blair SN. Obesity and cardiovascular disease. *Circ Res* 2016;**118**:1752–70. doi:10.1161/CIRCRESAHA.115.306883
47 Luppino FS, de Wit LM, Bouvy PF, *et al.* Overweight, obesity, and depression: a systematic review and meta-analysis of longitudinal studies. *Arch Gen Psychiatry* 2010;**67**:220–9. doi:10.1001/archgenpsychiatry.2010.2
48 Ayer J, Charakida M, Deanfield JE, *et al.* Lifetime risk: Childhood obesity and cardiovascular risk. *Eur Heart J* 2015;**36**:1371–6. doi:10.1093/eurheartj/ehv089
49 Bentham J, Di Cesare M, Bilano V, *et al.* Worldwide trends in body-mass index, underweight, overweight, and obesity from 1975 to 2016: a pooled analysis of 2416 population-based measurement studies in 128·9 million children, adolescents, and adults. *Lancet* 2017;**390**:2627–42. doi:10.1016/S0140-6736(17)32129-3
50 Nadeau KJ, Anderson BJ, Berg EG, *et al.* Youth-onset type 2 diabetes consensus report: Current status, challenges, and priorities. *Diabetes Care* 2016;**39**:1635–42. doi:10.2337/dc16-1066
51 2018 Physical Activity Guidelines Advisory Committee. *2018 Physical Activity Guidelines Advisory Committee Scientific Report*. Washington, DC: : U.S. Department of Health and Human Services, 2018 2018. https://health.gov/paguidelines/second-edition/report/pdf/pag_advisory_committee_report.pdf%0Ahttps://health.gov/paguidelines/second-edition/report/pdf/06_E_Systematic_Review_Literature_Search_Methodology.pdf
52 Escalante Y, Saavedra JM, García-Hermoso A, *et al.* Improvement of the lipid profile with exercise in obese children: A systematic review. *Prev Med (Baltim)* 2012;**54**:293–301. doi:10.1016/j.ypmed.2012.02.006
53 García-Hermoso A, Saavedra JM, Escalante Y. Effects of exercise on resting blood pressure in

obese children: A meta-analysis of randomized controlled trials. *Obes Rev* 2013;**14**:919–28. doi:10.1111/obr.12054

54 Esteban-Cornejo I, Cadenas-Sanchez C, Contreras-Rodriguez O, *et al.* A whole brain volumetric approach in overweight/obese children: Examining the association with different physical fitness components and academic performance. The ActiveBrains project. *Neuroimage* 2017;**159**:346–54. doi:10.1016/j.neuroimage.2017.08.011

55 Donnelly JE, Hillman CH, Castelli D, *et al.* Physical Activity, Fitness, Cognitive Function, and Academic Achievement in Children: A Systematic Review. *Med Sci Sports Exerc* 2016;**48**:1197–222. doi:10.1249/MSS.0000000000000901

56 Cohen N, Eichenbaum H. *Memory, Amnesia and the Hippocampal System*. Cambridge: 1993.

57 Davachi L. Item, context and relational episodic encoding in humans. *Curr Opin Neurobiol* 2006;**16**:693–700. doi:10.1016/j.conb.2006.10.012

58 Jaroudi W, Garami J, Garrido S, *et al.* Factors underlying cognitive decline in old age and Alzheimer's disease: The role of the hippocampus. *Rev Neurosci* 2017;**28**:705–14. doi:10.1515/revneuro-2016-0086

59 Luik AI, Zuurbier LA, Hofman A, *et al.* Stability and fragmentation of the activity rhythm across the sleep-wake cycle: The importance of age, lifestyle, and mental health. *Chronobiol Int* 2013;**30**:1223–30. doi:10.3109/07420528.2013.813528

60 Garaulet M, Martinez-Nicolas A, Ruiz JR, *et al.* Fragmentation of daily rhythms associates with obesity and cardiorespiratory fitness in adolescents: The HELENA study. *Clin Nutr* 2017;**36**:1558–66. doi:10.1016/j.clnu.2016.09.026

61 Hildebrand M, Van Hees VT, Hansen BH, *et al.* Age group comparability of raw accelerometer output from wrist- and hip-worn monitors. *Med Sci Sports Exerc* 2014;**46**:1816–24. doi:10.1249/MSS.0000000000000289

62 Hildebrand M, Hansen BH, van Hees VT, *et al.* Evaluation of raw acceleration sedentary thresholds in children and adults. *Scand J Med Sci Sports* 2017;**27**:1814–23. doi:10.1111/sms.12795

63 Cadenas-Sánchez C, Mora-González J, Migueles JH, *et al.* An exercise-based randomized controlled trial on brain, cognition, physical health and mental health in overweight/obese children (ActiveBrains project): Rationale, design and methods. *Contemp Clin Trials* 2016;**47**:315–24. doi:10.1016/j.cct.2016.02.007

64 Cole TJ, Lobstein T. Extended international (IOTF) body mass index cut-offs for thinness, overweight and obesity. *Pediatr Obes* 2012;**7**:284–94. doi:10.1111/j.2047-6310.2012.00064.x

65 Bervoets L, Massa G. Defining morbid obesity in children based on BMI 40 at age 18 using the extended international (IOTF) cut-offs. *Pediatr Obes* 2014;**9**:e94–8. doi:10.1111/j.2047-6310.2014.00217.x

66 Van Hees VT, Fang Z, Langford J, *et al.* Autocalibration of accelerometer data for free-living physical activity assessment using local gravity and temperature: An evaluation on four continents. *J Appl Physiol* 2014;**117**:738–44. doi:10.1152/japplphysiol.00421.2014

67 van Hees VT, Gorzelniak L, Dean León EC, *et al.* Separating Movement and Gravity Components in an Acceleration Signal and Implications for the Assessment of Human Daily Physical Activity. *PLoS One* 2013;**8**:1–10. doi:10.1371/journal.pone.0061691

68 Evenson KR, Catellier DJ, Gill K, *et al.* Calibration of two objective measures of physical activity for children. *J Sports Sci* 2008;**26**:1557–65. doi:10.1080/02640410802334196

69 Romanzini M, Petroski EL, Ohara D, *et al.* Calibration of ActiGraph GT3X, Actical and RT3 accelerometers in adolescents. *Eur J Sport Sci* 2014;**14**:91–9. doi:10.1080/17461391.2012.732614

70 Chandler JL, Brazendale K, Beets MW, *et al.* Classification of physical activity intensities using a wrist-worn accelerometer in 8-12-year-old children. *Pediatr Obes* 2016;**11**:120–7. doi:10.1111/ijpo.12033

71 Schrank FA, Mcgrew KS, Woodcock RW. Woodcock-Johnson ® III Assessment Service Bulletin Number 2 WJ III ® Technical Abstract. In: *WJ III Technical Manual (McGrew & Woodcock*. 2001.

72 Delis D, Kaplan E, Kramer J. *Delis-Kaplan executive function system (D-KEFS)*. San Antonio, TX: : The Psychological Corporation 2001.

73 Delis DC, Kramer JH, Kaplan E, *et al.* Reliability and validity of the Delis-Kaplan Executive Function System: An update. *J Int Neuropsychol Soc* 2004;**10**:301–3. doi:10.1017/S1355617704102191

74 Penner IK, Kobel M, Stcklin M, *et al.* The stroop task: Comparison between the original paradigm and computerized versions in children and adults. *Clin Neuropsychol* 2012;**26**:1142–53. doi:10.1080/13854046.2012.713513

75 Nyongesa MK, Ssewanyana D, Mutua AM, *et al.* Assessing executive function in adolescence: A scoping review of existing measures and their psychometric robustness. *Front Psychol* 2019;**10**. doi:10.3389/fpsyg.2019.00311

76 Murray, S T, Shapiro, D E. Fronto-temporal dysregulation in remitted bipolar patients: an fMRI delayed-non-match-to-sample (DNMS) study. *Bipolar Disord* 2009;**11**:351–60. doi:10.1016/j.cll.2010.01.003.Lyme

77 Kaufman AS, Kaufman NL. *Kaufman brief intelligence test.* Bloomington, MN: : Pearson 1990.

78 Trianes Torres MV, Blanca Mena MJ, Fernández Baena FJ, *et al.* Assessment of stress in childhood: Children's Daily Stress Inventory (Inventario

Infantil de Estresores Cotidiano, IIEC). *Psicothema* 2009;**21**:598–603.http://www.ncbi.nlm.nih.gov/pubmed/19861105

79 Rodrigo G, Lusiardo M. Spanish version of the Revised Children's Manifest Anxiety Scale in Uruguay: reliability and concurrent validity. *Psychol Rep* 1989;**65**:94. doi:10.2466/pr0.1989.65.1.94

80 Kovacs M. *The children's depression inventory (CDI)*. Toronto: 1992.

81 Watson D, Clark LA, Tellegen A. Development and validation of brief measures of positive and negative affect: The PANAS scales. J. Pers. Soc. Psychol. 1988;**54**:1063–70. doi:10.1037/0022-3514.54.6.1063

82 Lyubomirsky S, Lepper HS. A measure of subjective happiness: Preliminary reliability and construct validation. *Soc Indic Res* 1999;**46**:137–55. doi:10.1023/A:1006824100041

83 Scheier MF, Carver CS, Bridges MW. Distinguishing optimism from neuroticism (and trait anxiety, self-mastery, and self-esteem): a reevaluation of the Life Orientation Test. *J Pers Soc Psychol* 1994;**67**:1063–78. doi:10.1037//0022-3514.67.6.1063

84 Schwarzer R, Jerusalem M. Generalized self-efficacy scale. In: *Measures in health psychology: A user's portfolio*. Windsor, UK: : NFER-NELSON 1995. 35–7.

85 Martinez I, Musitu G, García J., *et al.* A cross-cultural analysis of the effects of family socialization on self-concept: Spain and Brazil. *Psicol Educ y Cult* 2003;**7**:239–59.

86 Rosenberg M. *Society and the adolescent self-image*. Princeton, NJ, USA: 1965.

87 Shah S, Novak S, Stapleton LM. Evaluation and comparison of models of metabolic syndrome using confirmatory factor analysis. *Eur J Epidemiol* 2006;**21**:343–9. doi:10.1007/s10654-006-9004-2

88 Artero EG, Espaa-Romero V, Castro-Piero J, *et al.* Reliability of field-based fitness tests in youth. *Int J Sports Med* 2011;**32**:159–69. doi:10.1055/s-0030-1268488

89 Castro-Piñero J, Artero EG, España-Romero V, *et al.* Criterion-related validity of field-based fitness tests in youth: A systematic review. *Br J Sports Med* 2010;**44**:934–43. doi:10.1136/bjsm.2009.058321

90 Ruiz JR, Castro-Piñero J, Artero EG, *et al.* Predictive validity of health-related fitness in youth: A systematic review. *Br J Sports Med* 2009;**43**:909–23. doi:10.1136/bjsm.2008.056499

91 Leger LA, Mercier D, Gadoury C, *et al.* The multistage 20 metre shuttle run test for aerobic fitness. *J Sports Sci* 1988;**6**:93–101. doi:10.1080/02640418808729800

92 Moore SA, McKay HA, Macdonald H, *et al.* Enhancing a somatic maturity prediction model. *Med Sci Sports Exerc* 2015;**47**:1755–64. doi:10.1249/MSS.0000000000000588

93 Delisle C, Sandin S, Forsum E, *et al.* A web- And mobile phone-based intervention to prevent obesity in 4-year-olds (MINISTOP): A population-based randomized controlled trial. *BMC Public Health* 2015;**15**:1–8. doi:10.1186/s12889-015-1444-8

94 Nyström CD, Forsum E, Henriksson H, *et al.* A mobile phone based method to assess energy and food intake in young children: A validation study against the doubly labelled water method and 24 h dietary recalls. *Nutrients* 2016;**8**:1–11. doi:10.3390/nu8010050

95 Delisle Nyström C, Pomeroy J, Henriksson P, *et al.* Evaluation of the wrist-worn ActiGraph wGT3x-BT for estimating activity energy expenditure in preschool children. *Eur J Clin Nutr* 2017;**71**:1212–7. doi:10.1038/ejcn.2017.114

96 Van Someren EJW, Lazeron RHC, Vonk BFM, *et al.* Gravitational artefact in frequency spectra of movement acceleration: Implications for actigraphy in young and elderly subjects. *J Neurosci Methods* 1996;**65**:55–62. doi:10.1016/0165-0270(95)00146-8

97 Bouten CVC, Koekkoek KTM, Verduin M, *et al.* A triaxial accelerometer and portable data processing unit for the assessment of daily physical activity. *IEEE Trans Biomed Eng* 1997;**44**:136–47. doi:10.1109/10.554760

98 Davies PSW, Coward WA, Gregory J, *et al.* Total energy expenditure and energy intake in the pre-school child: a comparison. *Br J Nutr* 1994;**72**:13–20. doi:10.1079/bjn19940005

99 Weir JB d. V. New methods for calculating metabolic rate with special reference to protein metabolism. *J Physiol* 1949;**109**:1–9. doi:10.1113/jphysiol.1949.sp004363

100 Black AE, Prentice AM, Coward WA. Use of food quotients to predict respiratory quotients for the doubly-labelled water method of measuring energy expenditure. Hum. Nutr. Clin. Nutr. 1986;**v. 40c**.

101 Henry C. Basal metabolic rate studies in humans: measurement and development of new equations. *Public Health Nutr* 2005;**8**:1133–52. doi:10.1079/phn2005801

102 Wells JCK, Williams JE, Chomtho S, *et al.* Pediatric reference data for lean tissue properties: Density and hydration from age 5 to 20 y. *Am J Clin Nutr* 2010;**91**:610–8. doi:10.3945/ajcn.2009.28428

103 Fletcher G, Balady G, Blair S, *et al.* Statement on Exercise : Benefits and Recommendations for Physical Activity Programs for All Americans. *Circulation* 1996;**94**:857–62. doi:10.1161/01.CIR.94.4.857

104 Warren JM, Ekelund U, Besson H, *et al.* Assessment of physical activity - a review of methodologies with reference to epidemiological research: a report of the exercise physiology section of the European Association of Cardiovascular Prevention and Rehabilitation. *Eur J Cardiovasc Prev Rehabil* 2010;**17**:127–39.

doi:10.1097/HJR.0b013e32832ed875

105 Bassett DR, Rowlands A, Trost SG. Calibration and validation of wearable monitors. *Med Sci Sports Exerc* 2012;**44**:32–8. doi:10.1249/MSS.0b013e3182399cf7

106 Rothney MP, Brychta RJ, Meade NN, *et al.* Validation of the actigraph two-regression model for predicting energy expenditure. *Med Sci Sports Exerc* 2010;**42**:1785–92. doi:10.1249/MSS.0b013e3181d5a984

107 Sasaki JE, John D, Freedson PS. Validation and comparison of ActiGraph activity monitors. *J Sci Med Sport* 2011;**14**:411–6. doi:10.1016/j.jsams.2011.04.003

108 Hänggi JM, Phillips LRS, Rowlands A V. Validation of the GT3X ActiGraph in children and comparison with the GT1M ActiGraph. *J Sci Med Sport* 2013;**16**:40–4. doi:10.1016/j.jsams.2012.05.012

109 Copeland JL, Esliger DW. Accelerometer assessment of physical activity in active, healthy older adults. *J Aging Phys Act* 2009;**17**:17–30. doi:10.1123/japa.17.1.17

110 Tudor-Locke C, Barreira T V., Schuna JM, *et al.* Fully automated waist-worn accelerometer algorithm for detecting children's sleep-period time separate from 24-h physical activity or sedentary behaviors. *Appl Physiol Nutr Metab* 2014;**39**:53–7. doi:10.1139/apnm-2013-0173

111 Barreira T V., Schuna JM, Mire EF, *et al.* Identifying Children's Nocturnal Sleep Using 24-h Waist Accelerometry. *Med Sci Sport Exerc* 2015;:1. doi:10.1249/MSS.0000000000000486

112 Crouter SE, Horton M, Bassett DR. Use of a two-regression model for estimating energy expenditure in children. *Med Sci Sports Exerc* 2012;**44**:1177–85. doi:10.1249/MSS.0b013e3182447825

113 Trost SG, Ward DS, Moorehead SM, *et al.* Validity of the computer science and applications (CSA) activity monitor in children. *Med Sci Sports Exerc* 1998;**30**:629–33. doi:10.1097/00005768-199804000-00023

114 Fairclough SJ, Noonan R, Rowlands A V., *et al.* Wear compliance and activity in children wearing wrist- and hip-mounted accelerometers. *Med Sci Sports Exerc* 2016;**48**:245–53. doi:10.1249/MSS.0000000000000771

115 Staudenmayer J, He S, Hickey A, *et al.* Methods to estimate aspects of physical activity and sedentary behavior from high-frequency wrist accelerometer measurements. *J Appl Physiol* 2015;**119**:396–403. doi:10.1152/japplphysiol.00026.2015

116 Aguilar-Farias N, Brown WJ, Peeters GMEE (Geeske). ActiGraph GT3X+ cut-points for identifying sedentary behaviour in older adults in free-living environments. *J Sci Med Sport* 2013;**17**:293–9. doi:10.1016/j.jsams.2013.07.002

117 Santos-Lozano A, Santín-Medeiros F, Cardon G, *et al.* Actigraph GT3X: validation and determination of physical activity intensity cut points. *Int J Sports Med* 2013;**34**:975–82. doi:10.1055/s-0033-1337945

118 Ellis K, Kerr J, Godbole S, *et al.* A random forest classifier for the prediction of energy expenditure and type of physical activity from wrist and hip accelerometers. *Physiol Meas* 2014;**35**:2191–203. doi:10.1088/0967-3334/35/11/2191

119 Kim Y, Lee JM, Peters BP, *et al.* Examination of different accelerometer cut-points for assessing sedentary behaviors in children. *PLoS One* 2014;**9**:1–8. doi:10.1371/journal.pone.0090630

120 Keadle SK, Shiroma EJ, Freedson PS, *et al.* Impact of accelerometer data processing decisions on the sample size, wear time and physical activity level of a large cohort study. *BMC Public Health* 2014;**14**:1210. doi:10.1186/1471-2458-14-1210

121 Cain KL, Sallis JF, Conway TL, *et al.* Using accelerometers in youth physical activity studies: a review of methods. *J Phys Act Health* 2013;**10**:437–50. doi:10.1123/jpah.10.3.437

122 Choi L, Ward SC, Schnelle JF, *et al.* Assessment of Wear/Nonwear Time Classification Algorithms for Triaxial Accelerometer. *Med Sci Sports Exerc* 2012;**44**:2009–16. doi:10.1249/MSS.0b013e318258cb36

123 Crouter SE, Flynn JI, Bassett DR. Estimating physical activity in youth using a wrist accelerometer. *Med Sci Sports Exerc* 2015;**47**:944–51. doi:10.1249/MSS.0000000000000502

124 Johansson E, Ekelund U, Nero H, *et al.* Calibration and cross-validation of a wrist-worn Actigraph in young preschoolers. *Pediatr Obes* 2015;**10**:1–6. doi:10.1111/j.2047-6310.2013.00213.x

125 Costa S, Barber SE, Cameron N, *et al.* Calibration and validation of the ActiGraph GT3X+ in 2-3 year olds. *J Sci Med Sport* 2014;**17**:617–22. doi:10.1016/j.jsams.2013.11.005

126 Jimmy G, Seiler R, Mäder U. Development and validation of GT3X accelero-meter cut-off points in 5- to 9-year-old children based on indirect calorimetry measurements. *Schweizerische Zeitschrift fur Sport und Sport* 2013;**61**:37–43.

127 Ellis K, Kerr J, Godbole S, *et al.* Hip and wrist accelerometer algorithms for free-living behavior classification. *Med Sci Sports Exerc* 2016;**48**:933–40. doi:10.1249/MSS.0000000000000840

128 Butte NF, Wong WW, Lee JS, *et al.* Prediction of energy expenditure and physical activity in preschoolers. *Med Sci Sports Exerc* 2014;**46**:1216–26. doi:10.1249/MSS.0000000000000209

129 Booth A, Clarke M, Dooley G, *et al.* The nuts and bolts of PROSPERO: An international prospective register of systematic reviews. *Syst Rev* 2012;**1**:2. doi:10.1186/2046-4053-1-2

130 Moher D, Liberati A, Tetzlaff J, *et al.* Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement (Reprinted from Annals of Internal Medicine). *Phys Ther* 2009;**89**:873–80.

doi:10.1371/journal.pmed.1000097
131 Kozey-Keadle S, Libertine A, Lyden K, *et al.* Validation of Wearable Monitors for Assessing Sedentary Behavior. *Med Sci Sports Exerc* 2011;**43**:1561–7. doi:10.1249/MSS.0b013e31820ce174
132 Ridgers ND, Salmon J, Ridley K, *et al.* Agreement between activPAL and ActiGraph for assessing children ' s sedentary time. *Int J Behav Nutr Phys Act* 2012;**9**:15. doi:10.1186/1479-5868-9-15
133 Matthews CE, Chen KY, Freedson PS, *et al.* Amount of time spent in sedentary behaviors in the United States, 2003-2004. *Am J Epidemiol* 2008;**167**:875–81. doi:10.1093/aje/kwm390
134 Treuth MS, Schmitz K, Catellier DJ, *et al.* Defining accelerometer thresholds for activity intensities in adolescent girls. *Med Sci Sports Exerc* 2004;**36**:1259–66.
135 Aadland E, Ylvisaker E, Ylvisåker E. Reliability of Objectively Measured Sedentary Time and Physical Activity in Adults. *PLoS One* 2015;**10**:1–13. doi:10.1371/journal.pone.0133296
136 Aadland E, Ylvisåker E. Reliability of the actigraph GT3X+ accelerometer in adults under free-living conditions. *PLoS One* 2015;**10**:1–10. doi:10.1371/journal.pone.0134606
137 Troiano RP, Berrigan D, Dodd KW, *et al.* Physical activity in the United States measured by accelerometer. *Med Sci Sports Exerc* 2008;**40**:181–8. doi:10.1249/mss.0b013e31815a51b3
138 Metzger JS, Catellier DJ, Evenson KR, *et al.* Patterns of objectively measured physical activity in the United States. *Med Sci Sports Exerc* 2008;**40**:630–8. doi:10.1249/MSS.0b013e3181620ebc
139 Pate RR, Almeida MJ, McIver KL, *et al.* Validation and calibration of an accelerometer in preschool children. *Obesity* 2006;**14**:2000–6. doi:10.1038/oby.2006.234
140 Freedson P, Pober D, Janz KF. Calibration of accelerometer output for children. *Med Sci Sports Exerc* 2005;**37**. doi:10.1249/01.mss.0000185658.28284.ba
141 Davis MG, Fox KR. Physical activity patterns assessed by accelerometry in older people. *Eur J Appl Physiol* 2007;**100**:581–9. doi:10.1007/s00421-006-0320-8
142 Reilly JJ, Coyle J, Kelly L, *et al.* An objective method for measurement of sedentary behavior in 3- to 4-year olds. *Obes Res* 2003;**11**:1155–8. doi:10.1038/oby.2003.158
143 Andersen LB, Harro M, Sardinha LB, *et al.* Physical activity and clustered cardiovascular risk in children: a cross-sectional study (The European Youth Heart Study). *Lancet* 2006;**368**:299–304. doi:10.1016/S0140-6736(06)69075-2
144 Puyau MR, Adolph AL, Vohra FA, *et al.* Validation and calibration of physical activity monitors in children. *Obes Res* 2002;**10**:150–7. doi:10.1038/oby.2002.24
145 Sirard JR, Trost SG, Pfeiffer KA, *et al.* Calibration and evaluation of an objective measure of physical activity in preschool children. *J Phys Act Heal* 2005;**2**:345–57. doi:10.1123/jpah.2.3.345
146 Van Cauwenberghe E, Labarque V, Trost SG, *et al.* Calibration and comparison of accelerometer cut points in preschool children. *Int J Pediatr Obes* 2011;**6**:582–9. doi:10.3109/17477166.2010.526223
147 Aittasalo M, Vähä-Ypyä H, Vasankari T, *et al.* Mean amplitude deviation calculated from raw acceleration data: a novel method for classifying the intensity of adolescents' physical activity irrespective of accelerometer brand. *BMC Sports Sci Med Rehabil* 2015;**7**:1–7. doi:10.1186/s13102-015-0010-0
148 Grydeland M, Bergh IH, Bjelland M, *et al.* Correlates of weight status among Norwegian 11-year-olds: The HEIA study. *BMC Public Health* 2012;**12**:1. doi:10.1186/1471-2458-12-1053
149 Matthews C. Calibration for Accelerometer Output for Adults. *Med Sci Sport Exerc* 2005;**S512**:S512–22. doi:10.1249/01.mss.0000185659.11982.3d
150 Mattocks C, Leary S, Ness A, *et al.* Calibration of an accelerometer during free-living activities in children. *Int J Pediatr Obes* 2007;**2**:218–26. doi:10.1080/17477160701408809
151 Pruitt LA, Ph D, Glynn NW, *et al.* Use of accelerometry to measure physical activity in older adults at risk for mobility disability. *J Aging Phys Act* 2010;**16**:416–34.
152 Pulakka A, Cheung YB, Ashorn U, *et al.* Feasibility and validity of the ActiGraph GT3X accelerometer in measuring physical activity of Malawian toddlers. *Acta Paediatr Int J Paediatr* 2013;**102**:1192–8. doi:10.1111/apa.12412
153 Vähä-Ypyä H, Vasankari T, Husu P, *et al.* A universal, accurate intensity-based classification of different physical activities using raw data of accelerometer. *Clin Physiol Funct Imaging* 2015;**35**:64–70. doi:10.1111/cpf.12127
154 Vanhelst J, Béghin L, Turck D, *et al.* New validated thresholds for various intensities of physical activity in adolescents using the Actigraph accelerometer. *Int J Rehabil Res* 2011;**34**:175–7. doi:10.1097/MRR.0b013e328340129e
155 Zhu Z, Chen P, Zhuang J. Intensity classification accuracy of accelerometer-measured physical activities in chinese children and youth. *Res Q Exerc Sport* 2013;**84**:37–41. doi:10.1080/02701367.2013.850919
156 Zisko N, Carlsen T, Salvesen Ø, *et al.* New relative intensity ambulatory accelerometer thresholds for elderly men and women: The Generation 100 study. *BMC Geriatr* 2015;**15**:1–10. doi:10.1186/s12877-015-0093-1
157 Crouter SE, Churilla JR, Bassett DR. Estimating energy expenditure using accelerometers. *Eur J Appl Physiol* 2006;**98**:601–12. doi:10.1007/s00421-006-0307-5
158 Crouter SE, Kuffel E, Haas JD, *et al.* Refined two-regression model for the ActiGraph

accelerometer. *Med Sci Sports Exerc* 2010;**42**:1029–37. doi:10.1249/MSS.0b013e3181c37458

159 Liu S, Gao RX, Freedson PS. Computational methods for estimating energy expenditure in human physical activities. *Med Sci Sports Exerc* 2012;**44**:2138–46. doi:10.1249/MSS.0b013e31825e825a

160 Schmitz KH, Treuth M, Hannan P, *et al.* Predicting energy expenditure from accelerometry counts in adolescents girls. *Med Sci Sports Exerc* 2005;**37**:155–61. doi:10.1249/01.MSS.0000150084.97823.F7

161 Stec MJ, Rawson ES. Estimation of resistance exercise energy expenditure using triaxial accelerometry. *J strength Cond Res* 2012;**26**:1413–22. doi:10.1519/JSC.0b013e318248d7b4

162 Zakeri IF, Adolph AL, Puyau MR, *et al.* Cross-Sectional Time Series and Multivariate Adaptive Regression Splines Models Using Accelerometry and Heart Rate Predict Energy Expenditure of Preschoolers. *J Nutr* 2013;**143**:114–22. doi:10.3945/jn.112.168542

163 Zhu Z, Chen P, Zhuang J. Predicting chinese children and youth's energy expenditure using actigraph accelerometers: A calibration and cross-validation study. *Res Q Exerc Sport* 2013;**84**:37–41. doi:10.1080/02701367.2013.850989

164 Hjorth MF, Chaput JP, Damsgaard CT, *et al.* Measure of sleep and physical activity by a single accelerometer: Can a waist-worn Actigraph adequately measure sleep in children? *Sleep Biol Rhythms* 2012;**10**:328–35. doi:10.1111/j.1479-8425.2012.00578.x

165 Ozemek C, Kirschner MM, Wilkerson BS, *et al.* Intermonitor reliability of the GT3X+ accelerometer at hip, wrist and ankle sites during activities of daily living. *Physiol Meas* 2014;**35**:129–38. doi:10.1088/0967-3334/35/2/129

166 Slater JA, Botsis T, Walsh J, *et al.* Assessing sleep using hip and wrist actigraphy. *Sleep Biol Rhythms* 2015;**13**:172–80. doi:10.1111/sbr.12103

167 Tudor-Locke C, Barreira T V., Schuna JM. Comparison of step outputs for waist and wrist accelerometer attachment sites. *Med Sci Sports Exerc* 2015;**47**:839–42. doi:10.1249/MSS.0000000000000476

168 Tudor-Locke C, Barreira T V., Schuna JM, *et al.* Improving wear time compliance with a 24-hour waist-worn accelerometer protocol in the International Study of Childhood Obesity, Lifestyle and the Environment (ISCOLE). *Int J Behav Nutr Phys Act* 2015;**12**. doi:10.1186/s12966-015-0172-x

169 Brønd JC, Arvidsson D. Sampling frequency affects the processing of Actigraph raw acceleration data to activity counts. *J Appl Physiol* 2015;**120**:362–9. doi:10.1152/japplphysiol.00628.2015

170 Toftager M, Kristensen PL, Oliver M, *et al.* Accelerometer data reduction in adolescents: Effects on sample retention and bias. *Int J Behav Nutr Phys Act* 2013;**10**. doi:10.1186/1479-5868-10-140

171 Freedson PS, John D. Comment on "Estimating Activity and Sedentary Behavior from an Accelerometer on the Hip and Wrist". *Med Sci Sport Exerc* 2013;**45**:962–3. doi:10.1249/MSS.0b013e31827f024d

172 Donaldson SC, Montoye AHK, Tuttle MS, *et al.* Variability of objectively measured sedentary behavior. *Med Sci Sports Exerc* 2016;**48**:755–61. doi:10.1249/MSS.0000000000000828

173 Lyden K, Kozey Keadle SL, Staudenmayer JW, *et al.* Validity of two wearable monitors to estimate breaks from sedentary time. *Med Sci Sports Exerc* 2012;**44**:2243–52. doi:10.1249/MSS.0b013e318260c477

174 Ried-Larsen M, Brønd JC, Brage S, *et al.* Mechanical and free living comparisons of four generations of the Actigraph activity monitor. *Int J Behav Nutr Phys Act* 2012;**9**:1–10. doi:10.1186/1479-5868-9-113

175 Cain KL, Conway TL, Adams M a, *et al.* Comparison of older and newer generations of ActiGraph accelerometers with the normal filter and the low frequency extension. *Int J Behav Nutr Phys Act* 2013;**10**:51. doi:10.1186/1479-5868-10-51

176 Cellini N, Buman MP, Mcdevitt EA, *et al.* Direct comparison of two actigraphy devices with polysomnographically recorded naps in healthy young adults. *Chronobiol Int* 2013;**30**:691–8. doi:10.3109/07420528.2013.782312

177 Wanner M, Martin BW, Meier F, *et al.* Effects of filter choice in GT3X accelerometer assessments of free-living activity. *Med Sci Sports Exerc* 2013;**45**:170–7. doi:10.1249/MSS.0b013e31826c2cf1

178 Barreira T V., Brouillette RM, Foil HC, *et al.* Comparison of older adults steps/day using NL-1000 pedometer and two GTX+ accelerometer filters. *J Aging Phys Act* 2012;**21**:402–16.

179 Kahan D, Nicaise V, Reuben K. Convergent validity of four accelerometer cutpoints with direct observation of preschool children's outdoor physical activity. *Res Q Exerc Sport* 2013;**84**:59–67. doi:10.1080/02701367.2013.762294

180 Mâsse LC, Fuemmeler BF, Anderson CB, *et al.* Accelerometer data reduction: A comparison of four reduction algorithms on select outcome variables. *Med Sci Sports Exerc* 2005;**37**. doi:10.1249/01.mss.0000185674.09066.8a

181 Aibar A, Bois JE, Zaragoza Casterad J, *et al.* Do epoch lengths affect adolescents' compliance with physical activity guidelines? *J Sports Med Phys Fitness* 2014;**54**:255–63.

182 Peeters G, van Gellecum Y, Ryde G, *et al.* Is the pain of activity log-books worth the gain in precision when distinguishing wear and non-wear time for tri-axial accelerometers? *J Sci Med Sport* 2013;**16**:515–9. doi:10.1016/j.jsams.2012.12.002

183 Choi, L., Liu, Z., Matthews, E. and Buchowski MS, Choi L, Liu Z, *et al.* Validation of Accelerometer Wear and Nonwear Time Classification Algorithm. *Med Sci Sports Exerc* 2011;**43**:357–64. doi:10.1249/MSS.0b013e3181ed61a3

184 Katzmarzyk PT, Barreira T V., Broyles ST, *et al.* The international study of childhood obesity, lifestyle and the environment (ISCOLE): Design and methods. *BMC Public Health* 2013;**13**. doi:10.1186/1471-2458-13-900

185 Janssen X, Cliff DP, Reilly JJ, *et al.* Predictive validity and classification accuracy of actigraph energy expenditure equations and cut-points in young children. *PLoS One* 2013;**8**. doi:10.1371/journal.pone.0079124

186 Van Cauwenberghe E, Gubbels J, De Bourdeaudhuij I, *et al.* Feasibility and validity of accelerometer measurements to assess physical activity in toddlers. *Int J Behav Nutr Phys Act* 2011;**8**:67. doi:10.1186/1479-5868-8-67

187 Peterson NE, Sirard JR, Kulbok PA, *et al.* Validation of Accelerometer Thresholds and Inclinometry for Measurement of Sedentary Behavior in Young Adult University Students. *Res Nurs Heal* 2015;**38**:492–9. doi:10.1002/nur.21694

188 Rowlands A V., Rennie K, Kozarski R, *et al.* Children's physical activity assessed with wrist- and hip-worn accelerometers. *Med Sci Sports Exerc* 2014;**46**:2308–16. doi:10.1249/MSS.0000000000000365

189 Meredith-Jones K, Williams S, Galland B, *et al.* 24 h Accelerometry: impact of sleep-screening methods on estimates of sedentary behaviour and physical activity while awake. *J Sports Sci* 2016;**34**:679–85. doi:10.1080/02640414.2015.1068438

190 Rosenberger ME, Buman MP, Haskell WL, *et al.* Twenty-four Hours of Sleep, Sedentary Behavior, and Physical Activity with Nine Wearable Devices. *Med Sci Sports Exerc* 2016;**48**:457–65. doi:10.1249/MSS.0000000000000778

191 Kaplan RF, Wang Y, Loparo KA, *et al.* Performance evaluation of an automated single- channel sleep – wake detection algorithm. 2014;:113–22.

192 Zinkhan M, Berger K, Hense S, *et al.* Agreement of different methods for assessing sleep characteristics: A comparison of two actigraphs, wrist and hip placement, and self-report with polysomnography. *Sleep Med* 2014;**15**:1107–14. doi:10.1016/j.sleep.2014.04.015

193 Trost SG, Mciver KL, Pate RR. Conducting accelerometer-based activity assessments in field-based research. *Med Sci Sports Exerc* 2005;**37**:531–43. doi:10.1249/01.mss.0000185657.86065.98

194 Tryon WW, Williams R. Fully proportional actigraphy: A new instrument. *Behav Res Methods, Instruments, Comput* 1996;**28**:392–403. doi:10.3758/BF03200519

195 Cavagna GA, Franzetti P. The determinants of the step frequency in walking in humans. *J Physiol* 1986;**373**:235–42. doi:10.1113/jphysiol.1986.sp016044

196 Cavagna GA, Willems PA, Franzetti P, *et al.* The two power limits conditioning step frequency in human running. *J Physiol* 1991;**437**:95–108. doi:10.1113/jphysiol.1991.sp018586

197 John D, Miller R, Kozey-Keadle S, *et al.* Biomechanical examination of the 'plateau phenomenon' in ActiGraph vertical activity counts. *Physiol Meas* 2012;**33**:219–30. doi:10.1088/0967-3334/33/2/219

198 Robusto KM, Trost SG. Comparison of three generations of ActiGraph™ activity monitors in children and adolescents. *J Sports Sci* 2012;**30**:1429–35. doi:10.1080/02640414.2012.710761

199 Grydeland M, Hansen BH, Ried-Larsen M, *et al.* Comparison of three generations of ActiGraph activity monitors under free-living conditions: do they provide comparable assessments of overall physical activity in 9-year old children? *BMC Sport Sci Med Rehabil* 2014;**6**:26. doi:10.1186/2052-1847-6-26

200 Migueles JH, Rowlands A V., Huber F, *et al.* GGIR: A Research Community–Driven Open Source R Package for Generating Physical Activity and Sleep Outcomes From Multi-Day Raw Accelerometer Data. *J Meas Phys Behav* 2019;**2**:188–96. doi:10.1123/jmpb.2018-0063

201 Seinstra F, Wallom D, Keahey K. Editorial. *SoftwareX* 2015;**1–2**:1–2. doi:10.1016/j.softx.2015.08.001

202 van Hees VT, Sabia S, Jones SE, *et al.* Estimating sleep parameters using an accelerometer without sleep diary. *Sci Rep* 2018;**8**:1–11. doi:10.1038/s41598-018-31266-z

203 Sabia S, Cogranne P, van Hees VT, *et al.* Physical Activity and Adiposity Markers at Older Ages: Accelerometer Vs Questionnaire Data. *J Am Med Dir Assoc* 2015;**16**:438.e7-438.e13. doi:10.1016/j.jamda.2015.01.086

204 Doherty A, Jackson D, Hammerla N, *et al.* Large scale population assessment of physical activity using wrist worn accelerometers: The UK biobank study. *PLoS One* 2017;**12**:1–14. doi:10.1371/journal.pone.0169649

205 Gubelmann C, Heinzer R, Haba-Rubio J, *et al.* Physical Activity is Associated with Higher Sleep Efficiency in the General Population: The CoLaus Study. *Sleep* Published Online First: March 2018. doi:10.1093/sleep/zsy070

206 Jones SE, Lane JM, Wood AR, *et al.* Genome-wide association analyses of chronotype in 697,828 individuals provides new insights into circadian rhythms in humans and links to disease. Published Online First: 2018. doi:10.1101/303941

207 Jones SE, van Hees VT, Mazzotti DR, *et al.* Genetic studies of accelerometer-based sleep measures in 85,670 individuals yield new insights into human sleep behaviour. 2018.

208 Lane JM, Jones S, Dashti H, *et al.* Biological and clinical insights from genetics of insomnia

symptoms. 2018.
209 Bell JA, Hamer M, van Hees VT, *et al.* Healthy obesity and objective physical activity. *Am J Clin Nutr* 2015;**102**:268–75. doi:10.3945/ajcn.115.110924
210 Menai M, Van Hees VT, Elbaz A, *et al.* Accelerometer assessed moderate-To-vigorous physical activity and successful ageing: Results from the Whitehall II study. *Sci Rep* 2017;**8**:1–9. doi:10.1038/srep45772
211 Cassidy S, Fuller H, Chau J, *et al.* Accelerometer-derived physical activity in those with cardio-metabolic disease compared to healthy adults: a UK Biobank study of 52,556 participants. *Acta Diabetol* Published Online First: May 2018. doi:10.1007/s00592-018-1161-8
212 Stiles VH, Metcalf BS, Knapp KM, *et al.* A small amount of precisely measured high-intensity habitual physical activity predicts bone health in pre- and post-menopausal women in UK Biobank. *Int J Epidemiol* 2017;**46**:1847–56. doi:10.1093/ije/dyx080
213 Bachasson D, Landon-Cardinal O, Benveniste O, *et al.* Physical activity monitoring: A promising outcome measure in idiopathic inflammatory myopathies. *Neurology* 2017;**89**:101–3. doi:10.1212/WNL.0000000000004061
214 Sabia S, Van Hees VT, Shipley MJ, *et al.* Association between questionnaire-and accelerometer-assessed physical activity: The role of sociodemographic factors. *Am J Epidemiol* 2014;**179**:781–90. doi:10.1093/aje/kwt330
215 Da silva ICM, Van hees VT, Ramires V V., *et al.* Physical activity levels in three Brazilian birth cohorts as assessed with raw triaxial wrist accelerometry. *Int J Epidemiol* 2014;**43**:1959–68. doi:10.1093/ije/dyu203
216 Campbell C, Gott A, Langford J, *et al.* GENEAclassify. 2018.
217 Jackson D. OMGUI. 2018.
218 Jackson D. Export Re-sampled WAV. 2018.
219 White T. Pampro. 2018. doi:http://doi.org/10.5281/zenodo.1187043
220 Doherty A, Hollowell S, Willetts M, *et al.* BiobankAccelerometerAnalysis. 2017.
221 Willetts M, Hollowell S, Aslett L, *et al.* Statistical machine learning of sleep and physical activity phenotypes from sensor data in 96,220 UK Biobank participants. *Sci Rep* 2018;**8**:1–10. doi:10.1038/s41598-018-26174-1
222 van Hees VT, Renström F, Wright A, *et al.* Estimation of daily energy expenditure in pregnant and Non-Pregnant women using a Wrist-Worn Tri-Axial accelerometer. *PLoS One* 2011;**6**. doi:10.1371/journal.pone.0022922
223 Rowlands A V., Cliff DP, Fairclough SJ, *et al.* Moving forward with backward compatibility: translating wrist accelerometer data. *Med Sci Sports Exerc* 2016;**48**:2142–9. doi:10.1249/MSS.0000000000001015
224 Innerd P, Harrison R, Coulson M. Using open source accelerometer analysis to assess physical activity and sedentary behaviour in overweight and obese adults. *BMC Public Health* 2018;**18**:543. doi:10.1186/s12889-018-5215-1
225 Rowlands A V., Yates T, Davies M, *et al.* Raw accelerometer data analysis with GGIR R-package: does accelerometer brand matter? *Med Sci Sports Exerc* 2016;**48**:1935–41. doi:10.1249/MSS.0000000000000978
226 Eddelbuettel D. *Seamless R and C++ Integration with Rcpp*. New York, NY: : Springer New York 2013. doi:10.1007/978-1-4614-6868-4
227 Ligges U, Krey S, Mersmann O, *et al.* {tuneR}: Analysis of Music and Speech. 2018.
228 Ligges U, Short T, Kienzle P, *et al.* {signal}: Signal processing. 2014.
229 Zeileis A, Grothendieck G. zoo : S3 Infrastructure for Regular and Irregular Time Series. *J Stat Softw* 2005;**14**. doi:10.18637/jss.v014.i06
230 Afshin A, Forouzanfar MH, Reitsma MB, *et al.* Health effects of overweight and obesity in 195 countries over 25 years. *N Engl J Med* 2017;**377**:13–27. doi:10.1056/NEJMoa1614362
231 Seale JL, Conway JM, Canary JJ. Seven-day validation of doubly labeled water method using indirect room calorimetry. *J Appl Physiol* 1993;**74**:402 LP – 409.
232 Sasaki JE, Hickey AM, Staudenmayer JW, *et al.* Performance of Activity Classification Algorithms in Free-Living Older Adults. *Med Sci Sports Exerc* 2016;**48**:941–50. doi:10.1249/MSS.0000000000000844
233 Migueles JH, Cadenas-Sanchez C, Ekelund U, *et al.* Accelerometer Data Collection and Processing Criteria to Assess Physical Activity and Other Outcomes: A Systematic Review and Practical Considerations. *Sports Med* 2017;**47**:1821–45. doi:10.1007/s40279-017-0716-0
234 Henriksson H, Forsum E, Löf M. Evaluation of Actiheart and a 7Â d activity diary for estimating free-living total and activity energy expenditure using criterion methods in 1·5-and 3-year-old children. *Br J Nutr* 2014;**111**:1830–40. doi:10.1017/S0007114513004406
235 Sijtsma A, Schierbeek H, Goris AHC, *et al.* Validation of the TracmorD triaxial accelerometer to assess physical activity in preschool children. *Obesity* 2013;**21**:1877–83. doi:10.1002/oby.20401
236 Ojiambo R, Konstabel K, Veidebaum T, *et al.* Validity of hip-mounted uniaxial accelerometry with heart-rate monitoring vs. triaxial accelerometry in the assessment of free-living energy expenditure in young children: The IDEFICS Validation Study. *J Appl Physiol* 2012;**113**:1530–6. doi:10.1152/japplphysiol.01290.2011
237 Brage S, Westgate K, Franks PW, *et al.* Estimation of free-living energy expenditure by heart rate and movement sensing: A doubly-labelled water study. *PLoS One* 2015;**10**:1–19. doi:10.1371/journal.pone.0137206
238 Chomistek AK, Yuan C, Matthews CE, *et al.*

Physical Activity Assessment with the ActiGraph GT3X and Doubly Labeled Water. *Med Sci Sports Exerc* 2017;**49**:1935–44. doi:10.1249/MSS.0000000000001299

239 Van Hees VT. The importance of nuance in statements about methods for human energy expenditure estimation that use motion sensors. *Eur J Clin Nutr* 2017;**71**:1136–7. doi:10.1038/ejcn.2017.65

240 Nyström CD, Sandin S, Henriksson P, *et al.* Mobile-based intervention intended to stop obesity in preschool-aged children: The MINISTOP randomized controlled trial. *Am J Clin Nutr* 2017;**105**:1327–35. doi:10.3945/ajcn.116.150995

241 Jeran S, Steinbrecher A, Pischon T. Prediction of activity-related energy expenditure using accelerometer-derived physical activity under free-living conditions: A systematic review. *Int J Obes* 2016;**40**:1187–97. doi:10.1038/ijo.2016.14

242 Plasqui G, Bonomi AG, Westerterp KR. Daily physical activity assessment with accelerometers: New insights and validation studies. *Obes Rev* 2013;**14**:451–62. doi:10.1111/obr.12021

243 Zhang P. Model Selection Via Multifold Cross Validation. *Ann Stat* 1993;**21**:299–313. doi:10.1214/aos/1176349027

244 Bland JM, Altman DG. Statistical Methods for Assessing Agreement Between Two Methods of Clinical Measurement. *Lancet* 1986;**327**:307–10. doi:10.1016/S0140-6736(86)90837-8

245 Rosenberger ME, Haskell WL, Albinali F, *et al.* Estimating activity and sedentary behavior from an accelerometer on the hip or wrist. *Med Sci Sports Exerc* 2013;**45**:964–75. doi:10.1249/MSS.0b013e31827f0d9c

246 LaMunion SR, Bassett DR, Toth LP, *et al.* The effect of body placement site on ActiGraph wGT3X-BT activity counts. *Biomed Phys Eng Express* 2017;**3**:035026. doi:10.1088/2057-1976/aa777c

247 Wikland KA, Luo ZC, Niklasson A, *et al.* Swedish population-based longitudinal reference values from birth to 18 years of age for height, weight and head circumference. *Acta Paediatr Int J Paediatr* 2002;**91**:739–54. doi:10.1080/08035250213216

248 Wells JCK, Williams JE, Chomtho S, *et al.* Body-composition reference data for simple and reference techniques and a 4-component model : a new UK reference child 1 – 3. 2012;:1316–26. doi:10.3945/ajcn.112.036970.INTRODUCTION

249 International Atomic Energy Agency. Assessment of body composition and total energy expenditure in humans using stable isotope technique. *IAEA Hum Heal Ser* 2009;**3**.Assessment of body composition, and total energy expenditure in humans using stable isotope technique. International Atomic Energy Agency Vienna International Centre Human

250 Speakman JR. *Double labelled water. Theory and practice*. 1st ed. Chapman & Hall: London, UK 1997.

251 Ekelund U, Steene-Johannessen J, Brown WJ, *et al.* Does physical activity attenuate, or even eliminate, the detrimental association of sitting time with mortality? A harmonised meta-analysis of data from more than 1 million men and women. *Lancet* 2016;**388**:1302–10. doi:10.1016/S0140-6736(16)30370-1

252 Czeisler CA. Impact of sleepiness and sleep deficiency on public health - Utility of biomarkers. *J Clin Sleep Med* 2011;**7**:5–7. doi:10.5664/JCSM.1340

253 Butte NF, Ekelund U, Westerterp KR. Assessing physical activity using weareable monitors: measures of physical activity. *Med Sci Sport Exerc* 2012;**44**:S5-12. doi:10.1249/MSS.ObO

254 Lee I-M, Shiroma EJ. Using accelerometers to measure physical activity in large-scale epidemiological studies: issues and challenges. *Br J Sports Med* 2014;**48**:197–201. doi:10.1136/bjsports-2013-093154

255 Rennie KL, Wareham NJ. The validation of physical activity instruments for measuring energy expenditure: problems and pitfalls. *Public Health Nutr* 1998;**1**:265–71. doi:10.1079/phn19980043

256 Koster A, Shiroma EJ, Caserotti P, *et al.* Comparison of Sedentary Estimates between activPAL and Hip- and Wrist-Worn ActiGraph. *Med Sci Sports Exerc* 2016;**48**:1514–22. doi:10.1249/MSS.0000000000000924

257 Barreira T V, Schuna JMJ, Mire EF, *et al.* Identifying Children's Nocturnal Sleep Using 24-h Waist Accelerometry. *Med Sci Sports Exerc* Published Online First: 2014. doi:10.1249/MSS.0000000000000486

258 Buchan DS, McSeveney F, McLellan G. A comparison of physical activity from Actigraph GT3X+ accelerometers worn on the dominant and non-dominant wrist. *Clin Physiol Funct Imaging* 2018;:1–6. doi:10.1111/cpf.12538

259 Shiroma EJ, Schepps MA, Harezlak J, *et al.* Daily physical activity patterns from hip- and wrist-worn accelerometers. *Physiol Meas* 2016;**37**:1852–61. doi:10.1088/0967-3334/37/10/1852

260 Dieu O, Mikulovic J, Fardy PS, *et al.* Physical activity using wrist-worn accelerometers: comparison of dominant and non-dominant wrist. *Clin Physiol Funct Imaging* 2017;**37**:525–9. doi:10.1111/cpf.12337

261 Kamada M, Shiroma EJ, Harris TB, *et al.* Comparison of physical activity assessed using hip- and wrist-worn accelerometers. *Gait Posture* 2016;**44**:23–8. doi:10.1016/j.gaitpost.2015.11.005

262 Vähä-Ypyä H, Vasankari T, Husu P, *et al.* Validation of cut-points for evaluating the intensity of physical activity with accelerometry-based Mean Amplitude Deviation (MAD). *PLoS One* 2015;**10**:1–13. doi:10.1371/journal.pone.0134813

263 Arnardottir NY, Koster A, Domelen D Van, *et al.*

Objective measurements of daily physical activity patterns and sedentary behaviour in older adults: Age, Gene/Environment Susceptibility-Reykjavik Study. *Age Ageing* 2013;**42**:222–9. doi:10.1093/ageing/afs160

264 Troiano RP, McClain JJ, Brychta RJ, *et al.* Evolution of accelerometer methods for physical activity research. *Br J Sports Med* 2014;**48**:1019–23. doi:10.1136/bjsports-2014-093546

265 Rowlands A V., Stiles VH. Accelerometer counts and raw acceleration output in relation to mechanical loading. *J Biomech* 2012;**45**:448–54. doi:10.1016/j.jbiomech.2011.12.006

266 Migueles JH, Cadenas-Sanchez C, Tudor-Locke C, *et al.* Comparability of published cut-points for the assessment of physical activity: Implications for data harmonization. *Scand J Med Sci Sport* 2019;**29**:566–74. doi:10.1111/sms.13356

267 Robinson MA, Vanrenterghem J, Pataky TC. Statistical Parametric Mapping (SPM) for alpha-based statistical analyses of multi-muscle EMG time-series. *J Electromyogr Kinesiol* 2015;**25**:14–9. doi:10.1016/j.jelekin.2014.10.018

268 Lee I, Cook NR, Gaziano JM, *et al.* Vitamin E in the primary prevention of cardiovascular disease and cancer: the Women's Health Study: a randomized controlled trial. *JAMA* 2005;**294**:56–65. doi:10.1001/jama.294.1.56

269 Rowlands A V., Fraysse F, Catt M, *et al.* Comparability of measured acceleration from accelerometry-based activity monitors. *Med Sci Sports Exerc* 2015;**47**:201–10. doi:10.1249/MSS.0000000000000394

270 John D, Sasaki J, Staudenmayer J, *et al.* Comparison of raw acceleration from the genea and actigraph™ GT3X+ activity monitors. *Sensors (Switzerland)* 2013;**13**:14754–63. doi:10.3390/s131114754

271 Smith MP, Standl M, Heinrich J, *et al.* Accelerometric estimates of physical activity vary unstably with data handling. *PLoS One* 2017;**12**:1–12. doi:10.1371/journal.pone.0187706

272 Kerr J, Marinac CR, Ellis K, *et al.* Comparison of Accelerometry Methods for Estimating Physical Activity. *Med Sci Sports Exerc* 2017;**49**:617–24. doi:10.1249/MSS.0000000000001124

273 Rowlands A V, Mirkes EM, Yates T, *et al.* Accelerometer-assessed physical activity in epidemiology: are monitors equivalent? *Med Sci Sports Exerc* 2018;**50**:257–65. doi:10.1249/MSS.0000000000001435

274 Lee I-M, Shiroma EJ, Evenson KR, *et al.* Accelerometer-Measured Physical Activity and Sedentary Behavior in Relation to All-Cause Mortality: The Women's Health Study. *Circulation* 2018;**137**:203–5. doi:10.1161/CIRCULATIONAHA.117.031300

275 Van Loo CMT, Okely AD, Batterham MJ, *et al.* Wrist Acceleration Cut-points for Moderate-to-Vigorous Physical Activity in Youth. *Med Sci Sport Exerc* 2018;**50**:609–16. doi:10.1249/MSS.0000000000001449

276 Leinonen AM, Ahola R, Kulmala J, *et al.* Measuring physical activity in free-living conditions-Comparison of three accelerometry-based methods. *Front Physiol* 2017;**7**:1–9. doi:10.3389/fphys.2016.00681

277 Booth FW, Laye MJ, Lees SJ, *et al.* Reduced physical activity and risk of chronic disease: the biology behind the consequences. *Eur J Appl Physiol* 2008;**102**:381–90. doi:10.1007/s00421-007-0606-5

278 Forouzanfar MH, Afshin A, Alexander LT, *et al.* Global, regional, and national comparative risk assessment of 79 behavioural, environmental and occupational, and metabolic risks or clusters of risks, 1990–2015: a systematic analysis for the Global Burden of Disease Study 2015. *Lancet* 2016;**388**:1659–724. doi:10.1016/S0140-6736(16)31679-8

279 Myers J, McAuley P, Lavie CJ, *et al.* Physical Activity and Cardiorespiratory Fitness as Major Markers of Cardiovascular Risk: Their Independent and Interwoven Importance to Health Status. *Prog Cardiovasc Dis* 2015;**57**:306–14. doi:10.1016/j.pcad.2014.09.011

280 Kohl HW, Craig CL, Lambert EV, *et al.* The pandemic of physical inactivity: Global action for public health. *Lancet* 2012;**380**:294–305. doi:10.1016/S0140-6736(12)60898-8

281 Guthold R, Stevens GA, Riley LM, *et al.* Worldwide trends in insufficient physical activity from 2001 to 2016: a pooled analysis of 358 population-based surveys with 1·9 million participants. *Lancet Glob Heal* 2018;**S2214-109X**:30357–7. doi:10.1016/S2214-109X(18)30357-7

282 Althoff T, Sosič R, Hicks JL, *et al.* Large-scale physical activity data reveal worldwide activity inequality. *Nature* 2017;**547**:336–9. doi:10.1038/nature23018

283 Dennison BA, Erb TA, Jenkins PL. Television viewing and television in bedroom associated with overweight risk among low-income preschool children. *Pediatrics* 2002;**109**:1028–35.

284 Vale S, Santos R, Silva P, *et al.* Relationship of objective measurement of physical activity during school hours and BMI in preschool children. *Int J Pediatr Obes* 2011;**6**:37–8. doi:10.3109/17477166.2011.613663

285 Ding D, Nguyen B, Learnihan V, *et al.* Moving to an active lifestyle? A systematic review of the effects of residential relocation on walking, physical activity and travel behaviour. *Br J Sports Med* 2018;**52**:789–99. doi:10.1136/bjsports-2017-098833

286 Kobel S, Kettner S, Kesztyüs D, *et al.* Correlates of habitual physical activity and organized sports in German primary school children. *Public Health* 2015;**129**:237–43. doi:10.1016/j.puhe.2014.12.002

287 Cliff DP, Okely AD, Morgan PJ, *et al.* Proficiency deficiency: Mastery of fundamental movement

skills and skill components in overweight and obese children. *Obesity* 2012;**20**:1024–33. doi:10.1038/oby.2011.241

288 Tudor-Locke C, Rowe DA. Using cadence to study free-living ambulatory behaviour. *Sport Med* 2012;**42**:381–98. doi:10.2165/11599170-000000000-00000

289 Barreira T V., Katzmarzyk PT, Johnson WD, *et al.* Cadence patterns and peak cadence in US children and adolescents: NHANES, 2005-2006. *Med Sci Sports Exerc* 2012;**44**:1721–7. doi:10.1249/MSS.0b013e318254f2a3

290 Tudor-Locke C, Han H, Aguiar EJ, *et al.* How fast is fast enough? Walking cadence (steps/min) as a practical estimate of intensity in adults: a narrative review. *Br J Sports Med* 2018;**52**:776–88. doi:10.1136/bjsports-2017-097628

291 Ortega FB. ActiveBrains. http://profith.ugr.es/activebrains?lang=en

292 Fairclough SJ, Taylor S, Rowlands A V., *et al.* Average acceleration and intensity gradient of primary school children and associations with indicators of health and well-being. *J Sports Sci* 2019;**37**:2159–67. doi:10.1080/02640414.2019.1624313

293 Gomez-Marcos MA, Recio-Rodŕiguez JI, Patino-Alonso MC, *et al.* Relationship between physical activity and plasma fibrinogen concentrations in adults without chronic diseases. *PLoS One* 2014;**9**:1–7. doi:10.1371/journal.pone.0087954

294 Tudor-Locke C, Camhi SM, Leonardi C, *et al.* Patterns of adult stepping cadence in the 2005-2006 NHANES. *Prev Med (Baltim)* 2011;**53**:178–81. doi:10.1016/j.ypmed.2011.06.004

295 De Craemer M, De Decker E, Santos-Lozano A, *et al.* Validity of the Omron pedometer and the actigraph step count function in preschoolers. *J Sci Med Sport* 2015;**18**:289–93. doi:10.1016/j.jsams.2014.06.001

296 Tudor-Locke C, Schuna JM, Han H, *et al.* Cadence (steps/min) and intensity during ambulation in 6-20 year olds: the CADENCE-kids study. *Int J Behav Nutr Phys Act* 2018;**15**:20. doi:10.1186/s12966-018-0651-y

297 Tudor-Locke C, Brashear MM, Katzmarzyk PT, *et al.* Peak stepping cadence in free-living adults: 2005-2006 NHANES. *J Phys Act Health* 2012;**9**:1125–9. doi:10.1123/jpah.9.8.1125

298 Barreira T V., Schuna JM, Mire EF, *et al.* Normative steps/day and peak cadence values for united states children and adolescents: National Health and Nutrition Examination Survey 2005-2006. *J Pediatr* 2015;**166**:139-143.e3. doi:10.1016/j.jpeds.2014.09.014

299 Barreira T V., Katzmarzyk PT, Johnson WD, *et al.* Walking cadence and cardiovascular risk in children and adolescents: NHANES, 2005-2006. *Am J Prev Med* 2013;**45**:e27-34. doi:10.1016/j.amepre.2013.08.005

300 Martínez-Gómez D, Ruiz JR, Gómez-Martínez S, *et al.* Active commuting to school and cognitive performance in adolescents: The AVENA study. *Arch Pediatr Adolesc Med* 2011;**165**:300–5. doi:10.1001/archpediatrics.2010.244

301 Mendoza JA, Liu Y. Active Commuting to Elementary School and Adiposity: An Observational Study. *Child Obes* 2014;**10**:34–41. doi:10.1089/chi.2013.0133

302 Migueles JH, Cadenas-Sanchez C, Rowlands A V, *et al.* Comparability of accelerometer signal aggregation metrics across placements and dominant wrist cut points for the assessment of physical activity in adults. *Sci Rep* 2019;**9**:18235. doi:10.1038/s41598-019-54267-y

303 Togo F, Watanabe E, Park H, *et al.* How many days of pedometer use predict the annual activity of the elderly Reliably? *Med Sci Sports Exerc* 2008;**40**:1058–64. doi:10.1249/MSS.0b013e318167469a

304 Konstabel K, Veidebaum T, Verbestel V, *et al.* Objectively measured physical activity in European children: The IDEFICS study. *Int J Obes* 2014;**38**:S135–43. doi:10.1038/ijo.2014.144

305 Rääsk T, Lätt E, Jürimäe T, *et al.* Association of subjective ratings to objectively assessed physical activity in pubertal boys with differing BMI. *Percept Mot Skills* 2015;**121**:245–59. doi:10.2466/10.03.PMS.121c13x7

306 Burchartz A, Anedda B, Auerswald T, *et al.* Assessing physical behavior through accelerometry – State of the science, best practices and future directions. *Psychol Sport Exerc* 2020;**49**:101703. doi:10.1016/j.psychsport.2020.101703

307 John D, Tang Q, Albinali F, *et al.* An Open-Source Monitor-Independent Movement Summary for Accelerometer Data Processing. *J Meas Phys Behav* 2019;**2**:268–81. doi:10.1123/jmpb.2018-0068

308 Rowlands A V., Edwardson CL, Davies MJ, *et al.* Beyond Cut Points: Accelerometer Metrics that Capture the Physical Activity Profile. *Med Sci Sports Exerc* 2018;**50**:1323–32. doi:10.1249/MSS.0000000000001561

309 Mora-Gonzalez J, Esteban-Cornejo I, Cadenas-Sanchez C, *et al.* Fitness, physical activity, working memory, and neuroelectric activity in children with overweight/obesity. *Scand J Med Sci Sports* 2019;**29**:1352–63. doi:10.1111/sms.13456

310 Aadland E, Kvalheim OM, Anderssen SA, *et al.* The Triaxial Physical Activity Signature Associated with Metabolic Health in Children. *Med Sci Sports Exerc* 2019;**51**:2173–9. doi:10.1249/MSS.0000000000002021

311 Aadland E, Kvalheim OM, Anderssen SA, *et al.* The multivariate physical activity signature associated with metabolic health in children. *Int J Behav Nutr Phys Act* 2018;**15**:1–11. doi:10.1186/s12966-018-0707-z

312 Rowlands A V., Dawkins NP, Maylor B, *et al.* Enhancing the value of accelerometer-assessed physical activity: meaningful visual comparisons of data-driven translational accelerometer metrics. *Sport Med - Open* 2019;**5**. doi:10.1186/s40798-019-0225-9

313 Augustin NH, Mattocks C, Cooper AR, *et al.* Modelling fat mass as a function of weekly physical activity profiles measured by Actigraph accelerometers. *Physiol Meas* 2012;**33**:1831–9. doi:10.1088/0967-3334/33/11/1831

314 Augustin NH, Mattocks C, Faraway JJ, *et al.* Modelling a response as a function of high-frequency count data: The association between physical activity and fat mass. *Stat Methods Med Res* 2017;**26**:2210–26. doi:10.1177/0962280215595832

315 Benadjaoud MA, Menai M, van Hees VT, *et al.* The association between accelerometer-assessed physical activity and respiratory function in older adults differs between smokers and non-smokers. *Sci Rep* 2019;**9**:1–9. doi:10.1038/s41598-019-46771-y

316 Bouten C V, Verboeket-van de Venne WP, Westerterp KR, *et al.* Daily physical activity assessment: comparison between movement registration and doubly labeled water. *J Appl Physiol* 1996;**81**:1019–26. doi:10.1152/jappl.1996.81.2.1019

317 Garber CE, Blissmer B, Deschenes MR, *et al.* Quantity and quality of exercise for developing and maintaining cardiorespiratory, musculoskeletal, and neuromotor fitness in apparently healthy adults: Guidance for prescribing exercise. *Med Sci Sports Exerc* 2011;**43**:1334–59. doi:10.1249/MSS.0b013e318213fefb

318 Narayanan A, Desai F, Stewart T, *et al.* Application of Raw Accelerometer Data and Machine-Learning Techniques to Characterize Human Movement Behavior: A Systematic Scoping Review. *J Phys Act Health* 2020;**17**:360–83. doi:10.1123/jpah.2019-0088

319 Doherty A, Smith-Byrne K, Ferreira T, *et al.* GWAS identifies 14 loci for device-measured physical activity and sleep duration. *Nat Commun* 2018;**9**:5257. doi:10.1038/s41467-018-07743-4

320 van Kuppevelt D, Heywood J, Hamer M, *et al.* Segmenting accelerometer data from daily life with unsupervised machine learning. *PLoS One* 2019;**14**:e0208692. doi:10.1371/journal.pone.0208692

321 Ekelund U, Tarp J, Steene-Johannessen J, *et al.* Dose-response associations between accelerometry measured physical activity and sedentary time and all cause mortality: systematic review and harmonised meta-analysis. *BMJ* 2019;**366**:l4570. doi:10.1136/bmj.l4570

322 Goldsmith J, Zipunnikov V, Schrack J. Generalized multilevel function-on-scalar regression and principal component analysis. *Biometrics* 2015;**71**:344–53. doi:10.1111/biom.12278

323 Pearson K. Mathematical contributions to the theory of evolution.-On a form of spurious correlation which may arise when indices are used in the measurement of organs. *Proc R Soc London* 1897;**60**:489–98. doi:10.1098/rspl.1896.0076

324 Aitchison J. The statistical analysis of compositional data. *J R Stat Soc* 1982;**44**:139–77.

325 McGregor DE, Palarea-Albaladejo J, Dall PM, *et al.* Cox regression survival analysis with compositional covariates: Application to modelling mortality risk from 24-h physical activity patterns. *Stat Methods Med Res* 2019;**29**:096228021986412. doi:10.1177/0962280219864125

326 Chastin SFM, Palarea-Albaladejo J, Dontje ML, *et al.* Combined effects of time spent in physical activity, sedentary behaviors and sleep on obesity and cardio-metabolic health markers: A novel compositional data analysis approach. *PLoS One* 2015;**10**:e0139984. doi:10.1371/journal.pone.0139984

327 Dumuid D, Stanford TE, Martin-Fernández JA, *et al.* Compositional data analysis for physical activity, sedentary time and sleep research. *Stat Methods Med Res* 2018;**27**:3726–38. doi:10.1177/0962280217710835

328 Aadland E, Kvalheim OM, Anderssen SA, *et al.* Multicollinear physical activity accelerometry data and associations to cardiometabolic health: challenges, pitfalls, and potential solutions. *Int J Behav Nutr Phys Act* 2019;**16**:1–14. doi:10.1186/s12966-019-0836-z

329 Hinkle J, Rayens W. Partial least squares and compositional data: problems and alternatives. *Chemom Intell Lab Syst* 1995;**30**:159–72. doi:10.1016/0169-7439(95)00062-3

330 Egozcue JJ, Pawlowsky-Glahn V, Mateu-Figueras G, *et al.* Isometric Logratio Transformations for Compositional Data Analysis. *Math Geol* 2003;**35**:279–300. doi:10.1023/A:1023818214614

331 McGregor DE, Palarea-Albaladejo J, Dall PM, *et al.* Compositional analysis of the association between mortality and 24-hour movement behaviour from NHANES. *Eur J Prev Cardiol* 2019;:204748731986778. doi:10.1177/2047487319867783

332 Migueles JH, Cadenas-Sanchez C, Esteban-Cornejo I, *et al.* Associations of objectively-assessed physical activity and sedentary time with hippocampal gray matter volume in children with overweight/obesity. *J Clin Med* 2020;**9**. doi:10.3390/jcm9041080

333 Skala W. Some effects of the constant-sum problem in geochemistry. *Chem Geol* 1979;**27**:1–9. doi:10.1016/0009-2541(79)90099-8

334 Rook A. An investigation into the longevity of Cambridge sportsmen. *Br Med J* 1954;**1**:773–7. doi:10.1136/bmj.1.4865.773

335 Ekelund U, Dalene KE, Tarp J, *et al.* Physical activity and mortality: What is the dose response and how big is the effect. *Br J Sports Med* 2020;**0**:5–6. doi:10.1136/bjsports-2019-101765

336 Henson J, Yates T, Biddle SJHH, *et al.* Associations of objectively measured sedentary behaviour and physical activity with markers of cardiometabolic health. *Diabetologia* 2013;**56**:1012–20.

doi:10.1007/s00125-013-2845-9

337 Mora-Gonzalez J, Esteban-Cornejo I, Cadenas-Sanchez C, *et al.* Physical Fitness, Physical Activity, and the Executive Function in Children with Overweight and Obesity. *J Pediatr* 2019;**208**:50-56.e1. doi:10.1016/j.jpeds.2018.12.028

338 Grgic J, Dumuid D, Bengoechea EG, *et al.* Health outcomes associated with reallocations of time between sleep, sedentary behaviour, and physical activity: A systematic scoping review of isotemporal substitution studies. *Int J Behav Nutr Phys Act* 2018;**15**:1–68. doi:10.1186/s12966-018-0691-3

339 Mekary RA, Willett WC, Hu FB, *et al.* Isotemporal substitution paradigm for physical activity epidemiology and weight change. *Am J Epidemiol* 2009;**170**:519–27. doi:10.1093/aje/kwp163

340 Kvalheim OM, Arneberg R, Grung B, *et al.* Determination of optimum number of components in partial least squares regression from distributions of the root-mean-squared error obtained by Monte Carlo resampling. *J Chemom* 2018;**32**:1–12. doi:10.1002/cem.2993

341 Rajalahti T, Arneberg R, Berven FS, *et al.* Biomarker discovery in mass spectral profiles by means of selectivity ratio plot. *Chemom Intell Lab Syst* 2009;**95**:35–48. doi:10.1016/j.chemolab.2008.08.004

342 Rajalahti T, Kvalheim OM. Multivariate data analysis in pharmaceutics: A tutorial review. *Int J Pharm* 2011;**417**:280–90. doi:10.1016/j.ijpharm.2011.02.019

343 Aadland E, Andersen LB, Resaland GKKK, *et al.* Interpretation of Multivariate Association Patterns between Multicollinear Physical Activity Accelerometry Data and Cardiometabolic Health in Children-A Tutorial. *Metabolites* 2019;**9**:1–14. doi:10.3390/metabo9070129

344 Ramsay J, Silverman B. *Functional Data Analysis*. 2nd Editio. New York: 2005. doi:10.1007/978-3-540-32691-5_16

345 Sørensen H, Goldsmith J, Sangalli LM. An introduction with medical applications to functional data analysis. *Stat Med* 2013;**32**:5222–40. doi:10.1002/sim.5989

346 Bi Q, Goodman KE, Kaminsky J, *et al.* What Is Machine Learning: a Primer for the Epidemiologist. *Am J Epidemiol* Published Online First: 11 September 2019. doi:10.1093/aje/kwz189

347 Alaa AM, Bolton T, Angelantonio E Di, *et al.* Cardiovascular disease risk prediction using automated machine learning: A prospective study of 423,604 UK Biobank participants. *PLoS One* 2019;**14**:1–17. doi:10.1371/journal.pone.0213653

348 Hua A, Quicksall Z, Di C, *et al.* Accelerometer-based predictive models of fall risk in older women: a pilot study. *npj Digit Med* 2018;**1**. doi:10.1038/s41746-018-0033-5

349 Stanton JM. Galton, Pearson, and the Peas: A Brief History of Linear Regression for Statistics Instructors. *J Stat Educ* 2001;**9**. doi:10.1080/10691898.2001.11910537

350 Tarp J, Hansen BH, Fagerland MW, *et al.* Accelerometer-measured physical activity and sedentary time in a cohort of US adults followed for up to 13 years: The influence of removing early follow-up on associations with mortality. *Int J Behav Nutr Phys Act* 2020;**17**:1–8. doi:10.1186/s12966-020-00945-4

351 Wold S, Ruhe A, Wold H, *et al.* The Collinearity Problem in Linear Regression. The Partial Least Squares (PLS) Approach to Generalized Inverses. *SIAM J Sci Stat Comput* 1984;**5**:735–43. doi:10.1137/0905052

352 Di J, Spira A, Bai J, *et al.* Joint and Individual Representation of Domains of Physical Activity, Sleep, and Circadian Rhythmicity. *Stat Biosci* 2019;**11**:371–402. doi:10.1007/s12561-019-09236-4

353 Kvalheim OM, Karstang T V. Interpretation of latent-variable regression models. *Chemom Intell Lab Syst* 1989;**7**:39–51. doi:10.1016/0169-7439(89)80110-8

354 Rajalahti T, Arneberg R, Kroksveen AC, *et al.* Discriminating variable test and selectivity ratio plot: Quantitative tools for interpretation and variable (biomarker) selection in complex spectral or chromatographic profiles. *Anal Chem* 2009;**81**:2581–90. doi:10.1021/ac802514y

355 Reiss PT, Goldsmith J, Shang HL, *et al.* Methods for scalar-on-function regression. *Int Stat Rev* 2017;**85**:228–49. doi:10.1111/insr.12163

356 Morris JS. Comparison and Contrast of Two General Functional Regression Modeling Frameworks. *Stat Modelling* 2017;**17**:59–85. doi:10.1177/1471082X16681875

357 Goldsmith J, Liu X, Jacobson JS, *et al.* New Insights into Activity Patterns in Children, Found Using Functional Data Analyses. *Med Sci Sports Exerc* 2016;**48**:1723–9. doi:10.1249/MSS.0000000000000968

358 Chacón JE, Duong T. Multivariate plug-in bandwidth selection with unconstrained pilot bandwidth matrices. *Test* 2010;**19**:375–98. doi:10.1007/s11749-009-0168-4

359 Rowlands A, Davies M, Dempsey P, *et al.* Wrist-worn accelerometers: recommending ~1.0 m g as the minimum clinically important difference (MCID) in daily average acceleration for inactive adults. *Br J Sports Med* 2020;**0**:bjsports-2020-102293. doi:10.1136/bjsports-2020-102293

360 Jakicic JM, Kraus WE, Powell KE, *et al.* Association between Bout Duration of Physical Activity and Health: Systematic Review. *Med Sci Sports Exerc* 2019;**51**:1213–9. doi:10.1249/MSS.0000000000001933

361 Matricciani L, Paquet C, Galland B, *et al.* Children's sleep and health: A meta-review. *Sleep Med Rev* 2019;**46**:136–50. doi:10.1016/j.smrv.2019.04.011

362 Biddle GJHH, Edwardson CL, Henson J, *et al.* Associations of Physical Behaviours and

Behavioural Reallocations with Markers of Metabolic Health: A Compositional Data Analysis. *Int J Environ Res Public Health* 2018;**15**:1–14. doi:10.3390/ijerph15102280

363 Aadland E, Nilsen AKO, Andersen LB, *et al.* A comparison of analytical approaches to investigate associations for accelerometry-derived physical activity spectra with health and developmental outcomes in children. *J Sports Sci* 2020;:1–9. doi:10.1080/02640414.2020.1824341

364 Collins GS, Moons KGM. Reporting of artificial intelligence prediction models. *Lancet* 2019;**393**:1577–9. doi:10.1016/S0140-6736(19)30037-6

365 Monti JM, Hillman CH, Cohen NJ. Aerobic fitness enhances relational memory in preadolescent children: the FITKids randomized control trial. *Hippocampus* 2012;**22**:1876–82. doi:10.1002/hipo.22023

366 Kaplan MS, Hinds JW. Neurogenesis in the adult rat: Electron microscopic analysis of light radioautographs. *Science (80-)* 1977;**197**:1092–4. doi:10.1126/science.887941

367 Neve RL, Finch EA, Bird ED, *et al.* Growth-associated protein GAP-43 is expressed selectively in associative regions of the adult human brain. *Proc Natl Acad Sci U S A* 1988;**85**:3638–42. doi:10.1073/pnas.85.10.3638

368 Kandola A, Hendrikse J, Lucassen PJ, *et al.* Aerobic exercise as a tool to improve hippocampal plasticity and function in humans: Practical implications for mental health treatment. *Front Hum Neurosci* 2016;**10**:1–25. doi:10.3389/fncel.2016.00373

369 Chaddock L, Erickson KI, Prakash RS, *et al.* A neuroimaging investigation of the association between aerobic fitness, hippocampal volume, and memory performance in preadolescent children. *Brain Res* 2010;**1358**:172–83. doi:10.1016/j.brainres.2010.08.049

370 Ruotsalainen I, Renvall V, Gorbach T, *et al.* Aerobic fitness, but not physical activity, is associated with grey matter volume in adolescents. *Behav Brain Res* 2019;**362**:122–30. doi:10.1016/j.bbr.2018.12.041

371 Herting MM, Keenan MF, Nagel BJ. Aerobic Fitness Linked to Cortical Brain Development in Adolescent Males: Preliminary Findings Suggest a Possible Role of BDNF Genotype. *Front Hum Neurosci* 2016;**10**:327. doi:10.3389/fnhum.2016.00327

372 Sorrells SF, Paredes MF, Cebrian-Silla A, *et al.* Human hippocampal neurogenesis drops sharply in children to undetectable levels in adults. *Nature* 2018;**555**:377–81. doi:10.1038/nature25975

373 Wafa SWW bte SST, Shahril MR bin, Ahmad A bte, *et al.* Association between physical activity and health-related quality of life in children: A cross-sectional study. *Health Qual Life Outcomes* 2016;**14**:1–6. doi:10.1186/s12955-016-0474-y

374 Ou X, Andres A, Pivik RT, *et al.* Brain gray and white matter differences in healthy normal weight and obese children. *J Magn Reson Imaging* 2015;**42**:1205–13. doi:10.1002/jmri.24912

375 Patenaude B, Smith SM, Kennedy DN, *et al.* A Bayesian model of shape and appearance for subcortical brain segmentation. *Neuroimage* 2011;**56**:907–22. doi:10.1016/j.neuroimage.2011.02.046

376 Erickson KI, Voss MW, Prakash RS, *et al.* Exercise training increases size of hippocampus and improves memory. *Proc Natl Acad Sci U S A* 2011;**108**:3017–22. doi:10.1073/pnas.1015950108

377 Erickson KI, Prakash RS, Voss MW, *et al.* Aerobic fitness is associated with hippocampal volume in elderly humans. *Hippocampus* 2009;**19**:1030–9. doi:10.1002/hipo.20547

378 Crova C, Struzzolino I, Marchetti R, *et al.* Cognitively challenging physical activity benefits executive function in overweight children. *J Sports Sci* 2014;**32**:201–11. doi:10.1080/02640414.2013.828849

379 Xu Q-S, Liang Y-Z. Monte Carlo cross validation. *Chemom Intell Lab Syst* 2001;**56**:1–11. doi:10.1016/S0169-7439(00)00122-2

380 Kuzik N, Carson V, Andersen LB, *et al.* Physical Activity and Sedentary Time Associations with Metabolic Health Across Weight Statuses in Children and Adolescents. *OBESITY* 2017;**25**:1762–9. doi:10.1002/oby.21952

381 Drollette ES, Hillman CH. Walking effects on memory in children: Implications for individual differences in BMI. *Ment Health Phys Act* 2020;**18**. doi:10.1016/j.mhpa.2020.100317

382 Kang EB, Koo JH, Jang YC, *et al.* Neuroprotective Effects of Endurance Exercise Against High-Fat Diet-Induced Hippocampal Neuroinflammation. *J Neuroendocrinol* 2016;**28**. doi:10.1111/jne.12385

383 Fotuhi M, Do D, Jack C. Modifiable factors that alter the size of the hippocampus with ageing. *Nat Rev Neurol* 2012;**8**:189–202. doi:10.1038/nrneurol.2012.27

384 Park H-S, Cho H-S, Kim T-W. Physical exercise promotes memory capability by enhancing hippocampal mitochondrial functions and inhibiting apoptosis in obesity-induced insulin resistance by high fat diet. *Metab Brain Dis* 2018;**33**:283–92. doi:10.1007/s11011-017-0160-8

385 Stranahan AM, Mattson MP. Bidirectional metabolic regulation of neurocognitive function. *Neurobiol Learn Mem* 2011;**96**:507–16. doi:10.1016/j.nlm.2011.01.004

386 Stranahan AM. Models and mechanisms for hippocampal dysfunction in obesity and diabetes. *Neuroscience* 2015;**309**:125–39. doi:10.1016/j.neuroscience.2015.04.045

387 Burgess N, Maguire EA, O'Keefe J. The human hippocampus and spatial and episodic memory. *Neuron* 2002;**35**:625–41. doi:10.1016/S0896-6273(02)00830-9

388 Gold JJ, Trauner DA. Hippocampal volume and memory performance in children with perinatal stroke. *Pediatr Neurol* 2014;**50**:18–25. doi:10.1016/j.pediatrneurol.2013.08.029

389 Lynch KM, Shi Y, Toga AW, *et al.* Hippocampal Shape Maturation in Childhood and Adolescence. *Cereb Cortex* 2019;**29**:3651–65. doi:10.1093/cercor/bhy244

390 Leppänen MH, Migueles JH, Cadenas-Sanchez C, *et al.* Hip and wrist accelerometers showed consistent associations with fitness and fatness in children aged 8-12 years. *Acta Paediatr* 2020;**109**:995–1003. doi:10.1111/apa.15043

391 Eaton DK, McKnight-Eily LR, Lowry R, *et al.* Prevalence of Insufficient, Borderline, and Optimal Hours of Sleep Among High School Students - United States, 2007. *J Adolesc Heal* 2010;**46**:399–401. doi:10.1016/j.jadohealth.2009.10.011

392 Basch CE, Basch CH, Ruggles K V., *et al.* Prevalence of sleep duration on an average school night among 4 nationally representative successive samples of American high school students, 2007-2013. *Prev Chronic Dis* 2014;**11**:6–10. doi:10.5888/pcd11.140383

393 Owens JA, Au R, Carskadon M, *et al.* School start times for adolescents. *Pediatrics* 2014;**134**:642–9. doi:10.1542/peds.2014-1697

394 Dewald JF, Meijer AM, Oort FJ, *et al.* The influence of sleep quality, sleep duration and sleepiness on school performance in children and adolescents: A meta-analytic review. *Sleep Med Rev* 2010;**14**:179–89. doi:10.1016/j.smrv.2009.10.004

395 Tonetti L, Natale V, Randler C. Association between circadian preference and academic achievement: A systematic review and meta-analysis. *Chronobiol Int* 2015;**32**:792–801. doi:10.3109/07420528.2015.1049271

396 Wheaton AG, Chapman DP, Croft JB. School Start Times, Sleep, Behavioral, Health, and Academic Outcomes: A Review of the Literature. *J Sch Health* 2016;**86**:363–81. doi:10.1111/josh.12388

397 Hui Tee JY, Gan WY, Tan KA, *et al.* Obesity and unhealthy lifestyle associated with poor executive function among Malaysian adolescents. *PLoS One* 2018;**13**:1–17. doi:10.1371/journal.pone.0195934

398 Kuula L, Pesonen A-K, Heinonen K, *et al.* Naturally occurring circadian rhythm and sleep duration are related to executive functions in early adulthood. *J Sleep Res* 2018;**27**:113–9. doi:10.1111/jsr.12581

399 Ortega F, Ruiz J, Castillo R, *et al.* Sleep duration and cognitive performance in adolescence. the AVENA study. *Acta Paediatr Int J Paediatr* 2010;**99**:454–6. doi:10.1111/j.1651-2227.2009.01618.x

400 Urrila AS, Artiges E, Massicotte J, *et al.* Sleep habits, academic performance, and the adolescent brain structure. *Sci Rep* 2017;**7**:1–9. doi:10.1038/srep41678

401 Taki Y, Hashizume H, Thyreau B, *et al.* Sleep duration during weekdays affects hippocampal gray matter volume in healthy children. *Neuroimage* 2012;**60**:471–5. doi:10.1016/j.neuroimage.2011.11.072

402 Dudai Y, Karni A, Born J. The Consolidation and Transformation of Memory. *Neuron* 2015;**88**:20–32. doi:10.1016/j.neuron.2015.09.004

403 Axmacher N, Draguhn A, Elger CE, *et al.* Memory processes during sleep: beyond the standard consolidation theory. *Cell Mol Life Sci* 2009;**66**:2285–97. doi:10.1007/s00018-009-0019-1

404 Owens JA, Spirito A, McGuinn M. The Children's Sleep Habits Questionnaire (CSHQ): psychometric properties of a survey instrument for school-aged children. *Sleep* Published Online First: 2000. doi:10.1111/j.1469-8749.2001.tb00204.x

405 Diamond A. The early development of executive functions. In: *Lifespan cognition: mechanisms of change*. 2006. 603–10. doi:10.1093/acprof

406 Diamond A. Exectuve functions. *Annu Rev Psychol* 2013;**64**:135–68. doi:10.1146/annurev-psych-113011-143750.Executive

407 McCullagh MC, Rosemberg MA. Social Desirability Bias in Self-Reporting of Hearing Protector Use among Farm Operators. *Ann Occup Hyg* 2014;**59**:1200–7. doi:10.1093/annhyg/mev046

408 Hébert JR. Social Desirability Trait: Biaser or Driver of Self-Reported Dietary Intake? *J Acad Nutr Diet* 2016;**116**:1895–8. doi:10.1016/j.jand.2016.08.007

409 Mullola S, Ravaja N, Lipsanen J, *et al.* Gender differences in teachers' perceptions of students' temperament, educational competence, and teachability. *Br J Educ Psychol* 2012;**82**:185–206. doi:10.1111/j.2044-8279.2010.02017.x

410 Miller AL, Lumeng JC, Lebourgeois MK. Sleep patterns and obesity in childhood. *Curr Opin Endocrinol Diabetes Obes* 2015;**22**:41–7. doi:10.1097/MED

411 Moon JY, Wang T, Sofer T, *et al.* Objectively measured physical activity, sedentary behavior, and genetic predisposition to obesity in U.S. Hispanics/Latinos: Results from the hispanic community health study/study of Latinos (HCHS/SOL). *Diabetes* 2017;**66**:3001–12. doi:10.2337/db17-0573

412 Lang C, Kalak N, Brand S, *et al.* The relationship between physical activity and sleep from mid adolescence to early adulthood. A systematic review of methodological approaches and meta-analysis. *Sleep Med Rev* 2016;**28**:32–45. doi:10.1016/j.smrv.2015.07.004

413 Ruiz JRR, Castillo R, Labayen I, *et al.* Individual and combined effects of ApoE and MTHFR 677C/T polymorphisms on cognitive performance in Spanish adolescents: the AVENA study. *J Pediatr* 2010;**156**:978–84, 984 e1. doi:10.1016/j.jpeds.2009.12.018

414 Swanson J. The Delis-Kaplan Executive Function System:A Review. *Can J Sch Psychol* 2005;**20**:117–

28. doi:10.1177/0829573506295469
415 Stroop J. Studies of interference in serial verbal reactions. *J Exp Psychol* 1935;**121**:15–23.
416 Song X-WW, Dong Z-YY, Long X-YY, *et al.* REST: a toolkit for resting-state functional magnetic resonance imaging data processing. *PLoS One* 2011;**6**:e25031. doi:10.1371/journal.pone.0025031
417 Hayasaka S, Phan KL, Liberzon I, *et al.* Nonstationary cluster-size inference with random field and permutation methods. *Neuroimage* 2004;**22**:676–87. doi:10.1016/j.neuroimage.2004.01.041
418 Gross CG. Single neuron studies of inferior temporal cortex. *Neuropsychologia* 2008;**46**:841–52. doi:10.1016/j.neuropsychologia.2007.11.009
419 Levy J, Vidal JR, Oostenveld R, *et al.* Alpha-band suppression in the visual word form area as a functional bottleneck to consciousness. *Neuroimage* 2013;**78**:33–45. doi:10.1016/j.neuroimage.2013.04.020
420 Lewis GJ, Bates TC. Task-context-dependent linear representation of multiple visual objects in hyman parietal cortex. *Psychologist* 2013;**26**:194–8. doi:10.1162/jocn
421 Lundstrom BN, Ingvar M, Petersson KM. The role of precuneus and left inferior frontal cortex during source memory episodic retrieval. *Neuroimage* 2005;**27**:824–34. doi:10.1016/j.neuroimage.2005.05.008
422 Lundstrom BN, Petersson KM, Andersson J, *et al.* Isolating the retrieval of imagined pictures during episodic memory: Activation of the left precuneus and left prefrontal cortex. *Neuroimage* 2003;**20**:1934–43. doi:10.1016/j.neuroimage.2003.07.017
423 Lien A, Sampasa-Kanyinga H, Colman I, *et al.* Adherence to 24-hour movement guidelines and academic performance in adolescents. *Public Health* 2020;**183**:8–14. doi:10.1016/j.puhe.2020.03.011
424 Walsh JJ, Barnes JD, Cameron JD, *et al.* Associations between 24 hour movement behaviours and global cognition in US children: a cross-sectional observational study. *Lancet Child Adolesc Heal* 2018;**2**:783–91. doi:10.1016/S2352-4642(18)30278-5
425 Cliff DP, McNeill J, Vella SA, *et al.* Adherence to 24-Hour Movement Guidelines for the Early Years and associations with social-cognitive development among Australian preschool children. *BMC Public Health* 2017;**17**. doi:10.1186/s12889-017-4858-7
426 Glod CA, Teicher MH, Polcari A, *et al.* Circadian rest-activity disturbances in children with seasonal affective disorder. *J Am Acad Child Adolesc Psychiatry* 1997;**36**:188–95. doi:10.1097/00004583-199702000-00009
427 Van Der Heijden KB, De Sonneville LMJ, Althaus M. Time-of-day effects on cognition in preadolescents: A trails study. *Chronobiol Int* 2010;**27**:1870–94. doi:10.3109/07420528.2010.516047
428 Díaz-Morales JF, Escribano C. Social jetlag, academic achievement and cognitive performance: Understanding gender/sex differences. *Chronobiol Int* 2015;**32**:822–31. doi:10.3109/07420528.2015.1041599
429 Migueles JH, Cadenas-Sanchez C, Esteban-Cornejo I, *et al.* Associations of sleep with gray matter volume and their implications for academic achievement, executive function and intelligence in children with overweight/obesity. *Pediatr Obes* 2020;:e12707. doi:10.1111/ijpo.12707
430 Zavala-Crichton JP, Esteban-Cornejo I, Solis-Urra P, *et al.* Association of sedentary behavior with brain structure and intelligence in children with overweight or obesity: The ActiveBrains project. *J Clin Med* 2020;**9**:1101. doi:10.3390/jcm9041101
431 Martinez-Nicolas A, Madrid JA, Rol MA. Day-night contrast as source of health for the human circadian system. *Chronobiol Int* 2014;**31**:382–93. doi:10.3109/07420528.2013.861845
432 Martinez-Nicolas A, Martinez-Madrid MJ, Almaida-Pagan PF, *et al.* Assessing Chronotypes by Ambulatory Circadian Monitoring. *Front Physiol* 2019;**10**:1396. doi:10.3389/fphys.2019.01396
433 Roenneberg T, Kuehnle T, Pramstaller PP, *et al.* A marker for the end of adolescence. *Curr Biol* 2004;**14**:1038–9. doi:10.1016/j.cub.2004.11.039
434 Esteban-Cornejo I, Mora-Gonzalez J, Cadenas-Sanchez C, *et al.* Fitness, cortical thickness and surface area in overweight/obese children: The mediating role of body composition and relationship with intelligence. *Neuroimage* 2019;**186**:771–81. doi:10.1016/j.neuroimage.2018.11.047
435 Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J R Stat Soc Ser B* 1995;**57**:289–300.http://www.jstor.org/stable/2346101
436 Wu JQ, Li P, Stavitsky Gilbert K, *et al.* Circadian Rest-Activity Rhythms Predict Cognitive Function in Early Parkinson's Disease Independently of Sleep. *Mov Disord Clin Pract* 2018;**5**:614–9. doi:10.1002/mdc3.12692
437 Sun W, Li SX, Wang G, *et al.* Association of Sleep and Circadian Activity Rhythm with Emotional Face Processing among 12-month-old Infants. *Sci Rep* 2018;**8**:1–11. doi:10.1038/s41598-018-21448-0
438 de Greeff JW, Bosker RJ, Oosterlaan J, *et al.* Effects of physical activity on executive functions, attention and academic performance in preadolescent children: a meta-analysis. *J Sci Med Sport* 2018;**21**:501–7. doi:10.1016/j.jsams.2017.09.595
439 Watson A, Timperio A, Brown H, *et al.* Effect of classroom-based physical activity interventions on academic and physical activity outcomes : a systematic review and meta-analysis. Published Online First: 2017. doi:10.1186/s12966-017-

0569-9

440 Saeedi P, Petersohn I, Salpea P, *et al.* Global and regional diabetes prevalence estimates for 2019 and projections for 2030 and 2045: Results from the International Diabetes Federation Diabetes Atlas, 9th edition. *Diabetes Res Clin Pract* 2019;**157**:107843. doi:10.1016/j.diabres.2019.107843

441 Sarwar N, Gao P, Kondapally Seshasai SR, *et al.* Diabetes mellitus, fasting blood glucose concentration, and risk of vascular disease: A collaborative meta-analysis of 102 prospective studies. *Lancet* 2010;**375**:2215–22. doi:10.1016/S0140-6736(10)60484-9

442 Copeland KC, Zeitler P, Geffner M, *et al.* Characteristics of adolescents and youth with recent-onset type 2 diabetes: The TODAY cohort at baseline. *J Clin Endocrinol Metab* 2011;**96**:159–67. doi:10.1210/jc.2010-1642

443 Thivel D, Ring-Dimitriou S, Weghuber D, *et al.* Muscle Strength and Fitness in Pediatric Obesity: a Systematic Review from the European Childhood Obesity Group. *Obes Facts* 2016;**9**:52–63. doi:10.1159/000443687

444 Alberga AS, Prud'homme D, Kenny GP, *et al.* Effects of aerobic and resistance training on abdominal fat, apolipoproteins and high-sensitivity C-reactive protein in adolescents with obesity: The HEARTY randomized clinical trial. *Int J Obes* 2015;**39**:1494–500. doi:10.1038/ijo.2015.133

445 Alberga AS, Prud'homme D, Sigal RJ, *et al.* Effects of aerobic training, resistance training, or both on cardiorespiratory and musculoskeletal fitness in adolescents with obesity: the HEARTY trial. *Appl Physiol Nutr Metab* 2016;**41**:255–65. doi:10.1139/apnm-2015-0413

446 Goldfield GS, Alberga AS, Hadjiyannakis S, *et al.* Effects of Aerobic Training, Resistance Training, or Both on Psychological Health in Adolescents with Obesity: The HEARTY Randomized Controlled Trial. *J Consult Clin Psychol* 2015;**83**:1123–35. doi:10.2471/BLT.

447 Labayen I, Medrano M, Arenaza L, *et al.* Effects of Exercise in Addition to a Family-Based Lifestyle Intervention Program on Hepatic Fat in Children With Overweight. *Diabetes Care* 2020;**43**:306–13. doi:10.2337/dc19-0351

448 Davis CL, Pollock NK, Waller JL, *et al.* Exercise dose and diabetes risk in overweight and obese children: A randomized controlled trial. *JAMA - J Am Med Assoc* 2012;**308**:1103–12. doi:10.1001/2012.jama.10762

449 Davis CL, Litwin SE, Pollock NK, *et al.* Exercise effects on arterial stiffness and heart health in children with excess weight: The SMART RCT. *Int J Obes* 2019;:1152–63. doi:10.1038/s41366-019-0482-1

450 Farpour-Lambert NJ, Aggoun Y, Marchand LM, *et al.* Physical Activity Reduces Systemic Blood Pressure and Improves Early Markers of Atherosclerosis in Pre-Pubertal Obese Children. *J Am Coll Cardiol* 2009;**54**:2396–406. doi:10.1016/j.jacc.2009.08.030

451 Rodriguez-Ayllon M, Cadenas-Sánchez C, Estévez-López F, *et al.* Role of Physical Activity and Sedentary Behavior in the Mental Health of Preschoolers, Children and Adolescents: A Systematic Review and Meta-Analysis. *Sports Med* 2019;**49**:1383–410. doi:10.1007/s40279-019-01099-5

452 Martínez-Vizcaíno V, Martínez MS, Aguilar FS, *et al.* Validity of a single-factor model underlying the metabolic syndrome in children: a confirmatory factor analysis. *Diabetes Care* 2010;**33**:1370–2. doi:10.2337/dc09-2049

453 Alberti KGMM, Eckel RH, Grundy SM, *et al.* Harmonizing the metabolic syndrome: a joint interim statement of the International Diabetes Federation Task Force on Epidemiology and Prevention; National Heart, Lung, and Blood Institute; American Heart Association; World Heart Federation; International. *Circulation* 2009;**120**:1640–5. doi:10.1161/CIRCULATIONAHA.109.192644

454 Medrano M, Arenaza L, Ramírez-Vélez R, *et al.* Prevalence of responders for hepatic fat, adiposity and liver enzyme levels in response to a lifestyle intervention in children with overweight/obesity: EFIGRO randomized controlled trial. *Pediatr Diabetes* 2020;**21**:215–23. doi:10.1111/pedi.12949

455 Lubans DR, Lonsdale C, Cohen K, *et al.* Framework for the design and delivery of organized physical activity sessions for children and adolescents: rationale and description of the 'SAAFE' teaching principles. *Int J Behav Nutr Phys Act* 2017;**14**:24. doi:10.1186/s12966-017-0479-x

456 Blüher S, Schwarz P. Metabolically healthy obesity from childhood to adulthood - Does weight status alone matter? *Metabolism* 2014;**63**:1084–92. doi:10.1016/j.metabol.2014.06.009

457 Jolliffe CJ, Janssen I. Development of Age-Specific Adolescent Metabolic Syndrome Criteria That Are Linked to the Adult Treatment Panel III and International Diabetes Federation Criteria. *J Am Coll Cardiol* 2007;**49**:891–8. doi:10.1016/j.jacc.2006.08.065

458 *The International Society for the Advancement in Kinanthropometry (ISAK) International Standards for Anthropometric Assessment. ISAK*. Underdale, SA, Australia: 2001.

459 Lang JJ, Tomkinson GR, Janssen I, *et al.* Making a Case for Cardiorespiratory Fitness Surveillance Among Children and Youth. *Exerc Sport Sci Rev* 2018;**46**.

460 Martínez-Vizcaino V, Ortega FB, Solera-Martínez M, *et al.* Stability of the factorial structure of metabolic syndrome from childhood to adolescence: a 6-year follow-up study. *Cardiovasc Diabetol* 2011;**10**:81. doi:10.1186/1475-2840-10-81

461 Sink KM, Espeland MA, Castro CM, *et al.* Effect of a 24-month physical activity intervention vs health

education on cognitive outcomes in sedentary older adults: The LIFE randomized trial. *JAMA* 2015;**314**:781–90. doi:10.1001/jama.2015.9617

462 Little RJA. Missing-Data Adjustments in Large Surveys. *J Bus Econ Stat* 1988;**6**:287–96. doi:10.1080/07350015.1988.10509663

463 Katzmarzyk PT, Martin CK, Newton RL, *et al.* Weight Loss in Underserved Patients - A Cluster-Randomized Trial. *N Engl J Med* 2020;**383**:909–18. doi:10.1056/NEJMoa2007448

464 Kelley GA, Kelley KS, Pate RR. Exercise and adiposity in overweight and obese children and adolescents: A systematic review with network meta-analysis of randomised trials. *BMJ Open* 2019;**9**:8–12. doi:10.1136/bmjopen-2019-031220

465 García-Hermoso A, Ramírez-Vélez R, Ramírez-Campillo R, *et al.* Concurrent aerobic plus resistance exercise versus aerobic exercise alone to improve health outcomes in paediatric obesity: A systematic review and meta-Analysis. *Br J Sports Med* 2018;**52**:161–6. doi:10.1136/bjsports-2016-096605

466 Seabra AC, Seabra AF, Brito J, *et al.* Effects of a 5-month football program on perceived psychological status and body composition of overweight boys. *Scand J Med Sci Sports* 2014;**24 Suppl 1**:10–6. doi:10.1111/sms.12268

467 Monserrat Romero-Perez E, Antonio De Paz-Fernandez J, Antonio Camberos-Castaneda N, *et al.* Assessment of depression and anxiety states of obese children after participating in an exerecise program. *Biotecnia* 2015;**17**. doi:10.18633/bt.v17i0.162

468 Costigan SA, Eather N, Plotnikoff RC, *et al.* High-Intensity Interval Training on Cognitive and Mental Health in Adolescents. *Med Sci Sport Exerc* Published Online First: 2016. doi:10.1249/MSS.0000000000000993

469 Henriksson P, Henriksson H, Tynelius P, *et al.* Fitness and Body Mass Index During Adolescence and Disability Later in Life. *Ann Intern Med* 2019;**170**:230. doi:10.7326/M18-1861

470 Henriksson H, Henriksson P, Tynelius P, *et al.* Cardiorespiratory fitness, muscular strength, and obesity in adolescence and later chronic disability due to cardiovascular disease: a cohort study of 1 million men. *Eur Heart J* 2020;**41**:1503–10. doi:10.1093/eurheartj/ehz774

CURRICULUM VITAE

“Yo soy yo y mis circunstancias”

Ortega y Gasset

CONTENTS

CV at a glance

Jairo Hidalgo Migueles

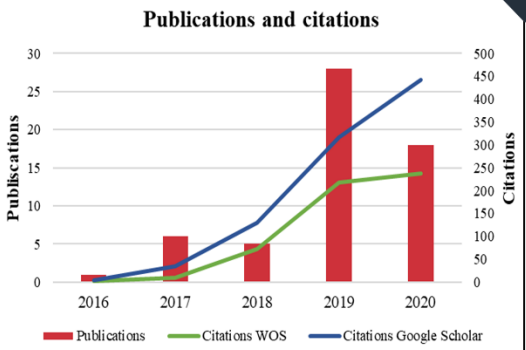
Born | 30/07/1991 (29 years old)
e-mail | jairohm@ugr.es
Phone | +34610959940

Research

Role in projects

- **ActiveBrains**
  - Management and organization
  - Data collection
  - Accelerometer data processing
  - Databases
  - Data analysis
  - Manuscripts writing
- **MINISTOP**
  - Accelerometer data processing
  - Data analysis
  - Manuscripts writing
- **Pilot study on accelerometry**
  - Study design
  - Project manager
  - Data collection
  - Data analysis
  - Manuscripts writing

Additionally, I have been in charge of the **accelerometer data processing** in more than 15 national and international projects



International Stays

15/10/2015
Start PhD studies

25/11 – 19/12
Research
Northeastern university (Boston, MA, USA)
Prof CH Hillman

29/05 – 08/06
Research
Karolinska Institutet / Linköping University (Stockholm, Sweden)
Prof M Löf

01/08 - 28/10
Research
National Institute on Aging (Baltimore, MD, USA)
PhD EJ Shiroma

31/05 – 27/07
Research
Karolinska Institutet / Linköping University (Stockholm, Sweden)
Prof M Löf

26/01 – 28/02
Research
University of Pittsburgh (Pittsburgh, PA, USA)
PhD KI Erickson

2015/16 | 2017 | 2018 | 2019 | 2020

28/08 - 25/11
Research
University of Massachusetts Amherst (Amherst, MA, USA)
Prof C Tudor-Locke

21/04 – 28/04
Teaching
University of Krakow (Krakow, Poland)

08/07 – 14/07
Research
eScience center (Amsterdam, the Netherlands)
PhD V van Hees

16/09 - 23/09
Research
University of Pittsburgh (Pittsburgh, PA, USA)
PhD KI Erickson

26/01 - 02/02
Teaching
Universitá Degli Studi di Roma "Foro Italico" (Rome, Italy)

Outreach

- I have contributed to some of the **GGIR** functionalities and addressed some bugs. I have led the GGIR description manuscript, assisted to a some research groups in the use of GGIR. I am now developing a click-based interface for GGIR (**GGIRshiny**)
- Finalist in the '**3-minute thesis**' contest, press releases and interviews about **Study IX**, organization of international workshop and symposium.

Teaching

I have had the opportunity to teach in the Faculty of Physical Activity and Sports Sciences in the University of Granada. I have been involved in the ***Fundamentals of Handball***, ***Physical activity and health***, and ***New trends in fitness***.

Academic experience and background

Experience

15/09/2016 – 18/12/2020 | Predoctoral fellow at University of Granada (FPU15/02645)
Funder: Spanish Ministry of Education, Culture and Sport
Supervisor: Francisco B. Ortega

15/10/2015 – 14/09/2016 | Research assistant associated to the ActiveBrains project
Funder: Spanish Ministry of Economy and Competitiveness
Supervisor: Francisco B. Ortega

06/10/2014 – 14/10/2015 | Research starting grant
Funder: University of Granada
Supervisor: Francisco B. Ortega

Background

15/10/2015 – 18/12/2020 | Predoctoral student in Biomedicine
Organization: University of Granada
Supervisor: Francisco B. Ortega

01/10/2014 – 30/06/2015 | MSc in research on physical activity and sport
Organization: University of Granada

15/09/2009 – 30/06/2014 | BSc in physical activity and sport sciences
Organization: University of Granada

Publications in peer-reviewed journals

Journal publications derived from this thesis

1 **Migueles JH**, Cadenas-Sanchez C, Ekelund U, Delisle Nyström C, Mora-Gonzalez J, Löf M, Labayen I, Ruiz JR, Ortega FB. Accelerometer Data Collection and Processing Criteria to Assess Physical Activity and Other Outcomes: A Systematic Review and Practical Considerations. Sports Med 2017 Sep;47(9):1821–1845. PMID: 28303543
Citations Google Scholar: 502 | **Citations WOS:** 308
Highly cited paper (top-1% in citations)

2 **Migueles JH**, Rowlands A V., Huber F, Sabia SS, van Hees VT. GGIR: A Research Community–Driven Open Source R Package for Generating Physical Activity and Sleep Outcomes From Multi-Day Raw Accelerometer Data. J Meas Phys Behav 2019;2(3):188–196.
Citations Google Scholar: 34

3 **Migueles JH**, Delisle Nyström C, Henriksson P, Cadenas-Sanchez C, Ortega FB, Löf M. Accelerometer Data Processing and Energy Expenditure Estimation in Preschoolers. Med Sci Sports Exerc 2019;51(3):590–598. PMID: 30303935
Citations Google Scholar: 4 | **Citations WOS: 3**

4 **Migueles JH**, Cadenas-Sanchez C, Rowlands A V, Henriksson P, Shiroma EJ, Acosta FM, Rodriguez-Ayllon M, Esteban-Cornejo I, Plaza-Florido A, Gil-Cosano JJ, Ekelund U, van Hees VT, Ortega FB. Comparability of accelerometer signal aggregation metrics across placements and dominant wrist cut points for the assessment of physical activity in adults. Sci Rep 2019 Dec 3;9(1):18235. PMID: 31796778
Citations Google Scholar: 8 | **Citations WOS: 3**

5 **Migueles JH**, Cadenas-Sanchez C, Tudor-Locke C, Löf M, Esteban-Cornejo I, Molina-Garcia P, Mora-Gonzalez J, Rodriguez-Ayllon M, Garcia-Marmol E, Ekelund U, Ortega FB. Comparability of published cut-points for the assessment of physical activity: Implications for data harmonization. Scand J Med Sci Sport 2019 Apr;29(4):566–574. PMID: 30548545
Citations Google Scholar: 40 | Citations WOS: 28
Top downloaded paper in 2018-2019 in the journal

6 **Migueles JH**, Cadenas-Sanchez C, Aguiar EJ, Molina-garcia P, Solis-Urra P, Mora-Gonzalez J, Garcia-Marmol E, Shiroma EJ, Labayen I, Chillon P, Lof M, Tudor-Locke C, Ortega FB. Step-Based Metrics and Overall Physical Activity in Children With Overweight or Obesity: Cross-Sectional Study. JMIR mHealth uHealth 2020 Apr;8(4). PMID: 32343251
Citations Google Scholar: 0 | Citations WOS: 0

7 **Migueles JH**, …, Ortega FB. The GRANADA consensus on analytical approaches to assess associations with accelerometer-determined physical behaviours in epidemiological studies
Submitted to Br J Sports Med

8 **Migueles JH**, Cadenas-Sanchez C, Esteban-Cornejo I, Torres-Lopez LV, Aadland E, Chastin SF, Erickson KI, Catena A, Ortega FB. Associations of objectively-assessed physical activity and sedentary time with hippocampal gray matter volume in children with overweight/obesity. J Clin Med 2020 Apr 10;9(4). PMID: 32290290
Citations Google Scholar: 1 | Citations WOS: 1

9 **Migueles JH**, Cadenas-Sanchez C, Esteban-Cornejo I, Mora-Gonzalez J, Rodriguez-Ayllon M, Solis-Urra P, Erickson KI, Kramer AF, Hillman CH, Catena A, Ortega FB. Associations of sleep with gray matter volume and their implications for academic achievement, executive function and intelligence in children with overweight/obesity. Pediatr Obes 2020 Aug 13;e12707. PMID: 32790234
Citations Google Scholar: 0 | Citations WOS: 0

10 **Migueles JH**, Martinez-Nicolas A, Cadenas-Sanchez C, Esteban-Cornejo I, Muntaner-Mas A, Mora-Gonzalez J, Rodriguez-Ayllon M, Madrid JA, Rol MA, Hillman CH, Catena A, Ortega FB. Activity-rest circadian pattern and academic achievement, executive function and intelligence in children with obesity. Scand J Med Sci Sports 2020 Oct.
Citations Google Scholar: 0 | Citations WOS: 0

11 **Migueles JH**, …, Ortega FB. Effects of exercise on cardiometabolic and mental health in children with overweight or obesity: The ActiveBrains randomized controlled trial.
In process to be submitted

Other journal publications as first author not included in this Thesis

1 **Migueles JH**, Cadenas-Sanchez C, Ortega FB. Critique of: "Physical Activity Assessment Between Consumer- and Research-Grade Accelerometers: A Comparative Study in Free-Living Conditions". JMIR mHealth uHealth [Internet] 2017 Feb 27;5(2):e15. PMID: 28242591
Citations Google Scholar: 4 | Citations WOS: 1

2 **Migueles JH**, Lee I-M, Cadenas-Sanchez C, Ortega FB, Buring JE, Shiroma EJ. Investigating an "ideal" combination of time spent in physical activity and sedentary behavior that is associated with greatest reduction in mortality rates among older women.
Under review in IJBNPA

Co-authored journal publications

1 Torres-Lopez L V, Cadenas-Sanchez C, **Migueles JH**, Adelantado-Renau M, Plaza-Florido A, Solis-Urra P, Molina-Garcia P, Ortega FB. Associations of Sedentary Behaviour, Physical Activity, Cardiorespiratory Fitness and Body Composition with Risk of Sleep-Related Breathing Disorders in Children with Overweight/Obesity: A Cross-Sectional Study. J Clin Med 2020 May 20;9(5). PMID: 32443799

2 Tenorio-Jiménez C, Martínez-Ramírez MJ, Tercero-Lozano M, Arraiza-Irigoyen C, Del Castillo-Codes I, Olza J, Plaza-Díaz J, Fontana L, **Migueles JH**, Olivares M, Gil Á, Gomez-Llorente C. Evaluation of the effect of Lactobacillus reuteri V3401 on biomarkers of inflammation, cardiovascular risk and liver steatosis in obese adults with metabolic syndrome: a randomized clinical trial (PROSIR). BMC Complement Altern Med 2018 Nov 20;18(1):306. PMID: 30453950

3 Sõritsa D, Mäestu E, Nuut M, Mäestu J, **Migueles JH**, Läänelaid S, Ehrenberg A, Sekavin A, Sõritsa A, Salumets A, Ortega FB, Altmäe S. Maternal physical activity and sedentary behaviour before and during in vitro fertilization treatment: a longitudinal study exploring the associations with controlled ovarian stimulation and pregnancy outcomes. J Assist Reprod Genet 2020 Aug;37(8):1869–1881. PMID: 32578030

4 Solis-Urra P, Esteban-Cornejo I, Cadenas-Sanchez C, Rodriguez-Ayllon M, Mora-Gonzalez J, **Migueles JH**, Labayen I, Verdejo-Román J, Kramer AF, Erickson KI, Hillman CH, Catena A, Ortega FB. Early life factors, gray matter brain volume and academic performance in overweight/obese children: The ActiveBrains project. Neuroimage 2019;202(August):116130. PMID: 31465844

5 Rodriguez-Ayllon M, Esteban-Cornejo I, Verdejo-Román J, Muetzel RL, **Migueles JH**, Mora-Gonzalez J, Solis-Urra P, Erickson KI, Hillman CH, Catena A, Tiemeier H, Ortega FB. Physical Activity, Sedentary Behavior, and White Matter Microstructure in Children with Overweight or Obesity. Med Sci Sports Exerc 2020;52(5):1218–1226. PMID: 31876665

6 Rodriguez-Ayllon M, Cadenas-Sánchez C, Estévez-López F, Muñoz NE, Mora-Gonzalez J, **Migueles JH**, Molina-García P, Henriksson H, Mena-Molina A, Martínez-Vizcaíno V, Catena A, Löf M, Erickson KI, Lubans DR, Ortega FB, Esteban-Cornejo I. Role of Physical Activity and Sedentary Behavior in the Mental Health of Preschoolers, Children and Adolescents: A Systematic Review and Meta-Analysis. Sports Med; 2019 Sep;49(9):1383–1410. PMID: 30993594

7 Rodriguez-Ayllon M, Cadenas-Sanchez C, Esteban-Cornejo I, **Migueles JH**, Mora-Gonzalez J, Henriksson P, Martín-Matillas M, Mena-Molina A, Molina-García P, Estévez-López F, Enriquez GM, Perales JC, Ruiz JR, Catena A, Ortega FB. Physical fitness and psychological health in overweight/obese children: A cross-sectional study from the ActiveBrains project. J Sci Med Sport 2018 Feb;21(2):179–184. PMID: 29031643

8 Plaza-Florido A, **Migueles JH**, Sacha J, Ortega FB. The role of heart rate in the assessment of cardiac autonomic modulation with heart rate variability. Clin Res Cardiol 2019;108(12):1408–1409. PMID: 31139891

9 Plaza-Florido A, **Migueles JH**, Mora-Gonzalez J, Molina-Garcia P, Rodriguez-Ayllon M, Cadenas-Sanchez C, Esteban-Cornejo I, Solis-Urra P, de Teresa C, Gutiérrez Á, Michels N, Sacha J, Ortega FB. Heart Rate Is a Better Predictor of Cardiorespiratory Fitness Than Heart Rate Variability in Overweight/Obese Children: The ActiveBrains Project. Front Physiol 2019;10(MAY):510. PMID: 31133870

10 Plaza-Florido A, **Migueles JH**, Mora-Gonzalez J, Molina-Garcia P, Rodriguez-Ayllon M, Cadenas-Sanchez C, Esteban-Cornejo I, Navarrete S, Maria Lozano R, Michels N, Sacha J, Ortega FB. The Role of Heart Rate on the Associations Between Body Composition and Heart Rate Variability in Children With Overweight/Obesity: The ActiveBrains Project. Front Physiol 2019;10(JUL):895. PMID: 31379602

11 Plaza-Florido A, Alcantara JMA, **Migueles JH**, Amaro-Gahete FJ, Acosta FM, Mora-Gonzalez J, Sacha J, Ortega FB. Inter- and intra-researcher reproducibility of heart rate variability parameters in three human cohorts. Sci Rep Nature Publishing Group UK; 2020 Jul 9;10(1):11399. PMID: 32647148

12 Ortolá R, García-Esquinas E, Cabanas-Sánchez V, **Migueles JH**, Martínez-Gómez D, Rodríguez-Artalejo F. Association of Physical Activity, Sedentary Behavior, and Sleep With Unhealthy Aging: Consistent Results for Device-Measured and Self-reported Behaviors Using Isotemporal Substitution Models. Journals Gerontol Ser A 2020;XX(Xx):1–10. [doi: 10.1093/gerona/glaa177]

13 Ortega FB, Cadenas-Sanchez C, **Migueles JH**, Labayen I, Ruiz JR, Sui X, Blair SN, Martínez-Vizcaino V, Lavie CJ. Role of Physical Activity and Fitness in the Characterization and Prognosis of the Metabolically Healthy Obesity Phenotype: A Systematic Review and Meta-analysis. Prog Cardiovasc Dis Elsevier Inc.; 2018;61(2):190–205. PMID: 30122522

14 Mora-Gonzalez J, **Migueles JH**, Esteban-Cornejo I, Cadenas-Sanchez C, Pastor-Villaescusa B, Molina-García P, Rodriguez-Ayllon M, Rico MC, Gil A, Aguilera CM, Escolano-Margarit MV, Gejl AK, Andersen LBO, Catena A, Ortega FB. Sedentarism, Physical

Activity, Steps, and Neurotrophic Factors in Obese Children. Med Sci Sports Exerc 2019;51(11):2325–2333. PMID: 31634295

15 Mora-Gonzalez J, Esteban-Cornejo I, Solis-Urra P, **Migueles JH**, Cadenas-Sanchez C, Molina-Garcia P, Rodriguez-Ayllon M, Hillman CH, Catena A, Pontifex MB, Ortega FB. Fitness, physical activity, sedentary time, inhibitory control, and neuroelectric activity in children with overweight or obesity: The ActiveBrains project. Psychophysiology 2020;57(6):e13579. PMID: 32249933

16 Mora-Gonzalez J, Esteban-Cornejo I, Cadenas-Sanchez C, **Migueles JH**, Rodriguez-Ayllon M, Molina-García P, Hillman CH, Catena A, Pontifex MB, Ortega FB. Fitness, physical activity, working memory, and neuroelectric activity in children with overweight/obesity. Scand J Med Sci Sports 2019 Sep;29(9):1352–1363. PMID: 31058358

17 Mora-Gonzalez J, Esteban-Cornejo I, Cadenas-Sanchez C, **Migueles JH**, Molina-Garcia P, Rodriguez-Ayllon M, Henriksson P, Pontifex MB, Catena A, Ortega FB. Physical Fitness, Physical Activity, and the Executive Function in Children with Overweight and Obesity. J Pediatr 2019;208:50-56.e1. PMID: 30902422

18 Molina-Garcia P, Mora-Gonzalez J, **Migueles JH**, Rodriguez-Ayllon M, Esteban-Cornejo I, Cadenas-Sanchez C, Plaza-Florido A, Gil-Cosano JJ, Pelaez-Perez MA, Garcia-Delgado G, Vanrenterghem J, Ortega FB. Effects of Exercise on Body Posture, Functional Movement, and Physical Fitness in Children With Overweight/Obesity. J strength Cond Res 2020;34(8):2146–2155. PMID: 32459738

19 Molina-Garcia P, Miranda-Aparicio D, Molina-Molina A, Plaza-Florido A, **Migueles JH**, Mora-Gonzalez J, Cadenas-Sanchez C, Esteban-Cornejo I, Rodriguez-Ayllon M, Solis-Urra P, Vanrenterghem J, Ortega FB. Effects of Exercise on Plantar Pressure during Walking in Children with Overweight/Obesity. Med Sci Sports Exerc 2020;52(3):654–662. PMID: 31524828

20 Molina-Garcia P, **Migueles JH**, Cadenas-Sanchez C, Esteban-Cornejo I, Mora-Gonzalez J, Rodriguez-Ayllon M, Plaza-Florido A, Vanrenterghem J, Ortega FB. A systematic review on biomechanical characteristics of walking in children and adolescents with overweight/obesity: Possible implications for the development of musculoskeletal disorders. Obes Rev 2019;20(7):1033–1044. PMID: 30942558

21 Medrano M, Arenaza L, **Migueles JH**, Rodríguez-Vigil B, Ruiz JR, Labayen I. Associations of physical activity and fitness with hepatic steatosis, liver enzymes, and insulin resistance in children with overweight/obesity. Pediatr Diabetes 2020;21(4):565–574. PMID: 32237015

22 Leppänen MH, **Migueles JH**, Cadenas-Sanchez C, Henriksson P, Mora-Gonzalez J, Henriksson H, Labayen I, Löf M, Esteban-Cornejo I, Ortega FB. Hip and wrist accelerometers showed consistent associations with fitness and fatness in children aged 8-12 years. Acta Paediatr 2020 Oct 4;109(5):995–1003. PMID: 31583747

23 Jurado-Fasoli L, De-la-O A, Molina-Hidalgo C, **Migueles JH**, Castillo MJ, Amaro-Gahete FJ. Exercise training improves sleep quality: A randomized controlled trial. Eur J Clin Invest 2020 Mar;50(3):e13202. PMID: 31989592

24 Henriksson P, Sandborg J, Blomberg M, Alexandrou C, Maddison R, Silfvernagel K, Henriksson H, Leppänen MH, **Migueles JH**, Widman L, Thomas K, Trolle Lagerros Y, Löf M. A Smartphone App to Promote Healthy Weight Gain, Diet, and Physical Activity During Pregnancy (HealthyMoms): Protocol for a Randomized Controlled Trial. JMIR Res Protoc 2019 Mar 1;8(3):e13011. PMID: 30821695

25 Gomez-Bruton A, Arenaza L, Medrano M, Mora-Gonzalez J, Cadenas-Sanchez C, **Migueles JH**, Muñoz-Hernández V, Merchan-Ramirez E, Martinez-Avila WD, Maldonado J, Oses M, Tobalina I, Gracia-Marco L, Vicente-Rodriguez G, Ortega FB, Labayen I. Associations of dietary energy density with body composition and cardiometabolic risk in children with overweight and obesity: role of energy density calculations, under-reporting energy intake and physical activity. Br J Nutr 2019;121(9):1057–1068. PMID: 30724143

26 Gil-Cosano JJ, Gracia-Marco L, Ubago-Guisado E, **Migueles JH**, Mora-Gonzalez J, Escolano-Margarit M V., Gómez-Vida J, Maldonado J, Ortega FB. Muscular Fitness Mediates the Association between 25-Hydroxyvitamin D and Areal Bone Mineral Density in Children with Overweight/Obesity. Nutrients 2019 Nov 14;11(11):1–10. PMID: 31739435

27 Esteban-Cornejo I, Mora-Gonzalez J, Cadenas-Sanchez C, Contreras-Rodriguez O, Verdejo-Román J, Henriksson P, **Migueles JH**, Rodriguez-Ayllon M, Molina-García P, Suo C, Hillman CH, Kramer AF, Erickson KI, Catena A, Verdejo-García A, Ortega FB. Fitness, cortical thickness and surface area in overweight/obese children: The mediating role of body composition and relationship with intelligence. Neuroimage 2019;186(November 2018):771–781. PMID: 30500426

28 Esteban-Cornejo I, Cadenas-Sanchez C, Contreras-Rodriguez O, Verdejo-Roman J, Mora-Gonzalez J, **Migueles JH**, Henriksson P, Davis CL, Verdejo-Garcia A, Catena A, Ortega FB. A whole brain volumetric approach in overweight/obese children: Examining the association with different physical fitness

components and academic performance. The Active-Brains project. Neuroimage 2017;159(July):346–354. PMID: 28789992

29 Delisle Nyström C, Pomeroy J, Henriksson P, Forsum E, Ortega FB, Maddison R, **Migueles JH**, Löf M. Evaluation of the wrist-worn ActiGraph wGT3x-BT for estimating activity energy expenditure in preschool children. Eur J Clin Nutr Nature Publishing Group; 2017;71(10):1212–1217. PMID: 28745334

30 Chillón P, Herrador-Colmenero M, **Migueles JH**, Cabanas-Sánchez V, Fernández-Santos JR, Veiga ÓL, Castro-Piñero J, up&down study group. Convergent validation of a questionnaire to assess the mode and frequency of commuting to and from school. Scand J Public Health Sweden; 2017 Aug;45(6):612–620. PMID: 30747037

31 Cadenas-Sanchez C, **Migueles JH**, Ortega FB. Further Evidence on Cardiorespiratory Fitness as a Key Factor for the Metabolically Healthy Obese Phenotype Independent of the Race. J Adolesc Health Society for Adolescent Health and Medicine; 2019;64(3):290–291. PMID: 30819330

32 Cadenas-Sanchez C, **Migueles JH**, Esteban-Cornejo I, Mora-Gonzalez J, Henriksson P, Rodriguez-Ayllon M, Molina-García P, Löf M, Labayen I, Hillman CH, Catena A, Ortega FB. Fitness, physical activity and academic achievement in overweight/obese children. J Sports Sci Routledge; 2020 Apr 24;38(7):731–740. PMID: 32091309

33 Cadenas-Sanchez C, Esteban-Cornejo I, **Migueles JH**, Labayen I, Verdejo-Román J, Mora-Gonzalez J, Henriksson P, Maldonado J, Gómez-Vida J, Hillman CH, Erickson KI, Kramer AF, Catena A, Ortega FB. Differences in Brain Volume between Metabolically Healthy and Unhealthy Overweight and Obese Children: The Role of Fitness. J Clin Med 2020;9(4):1059. [doi: 10.3390/jcm9041059]

34 Cabanas-Sánchez V, Esteban-Cornejo I, **Migueles JH**, Banegas JR, Graciani A, Rodríguez-Artalejo F, Martínez-Gómez D. Twenty four-hour activity cycle in older adults using wrist-worn accelerometers: The seniors-ENRICA-2 study. Scand J Med Sci Sports 2020 Apr;30(4):700–708. PMID: 31834945

35 Béghin L, Vanhelst J, Drumez E, **Migueles JH**, Androutsos O, Widhalm K, Julian C, Moreno LA, De Henauw S, Gottrand F, HELENA study group. Gender influences physical activity changes during adolescence: The HELENA study. Clin Nutr 2019;38(6):2900–2905. PMID: 30718097

36 Amaro-Gahete FJ, Acosta FM, **Migueles JH**, Ponce González JG, Ruiz JR. Association of sedentary and physical activity time with maximal fat oxidation during exercise in sedentary adults. Scand J Med Sci Sports Denmark; 2020 Sep;30(9):1605–1614. PMID: 32335956

37 Alcantara JMA, Plaza-Florido A, Amaro-Gahete FJ, Acosta FM, **Migueles JH**, Molina-Garcia P, Sacha J, Sanchez-Delgado G, Martinez-Tellez B. Impact of Using Different Levels of Threshold-Based Artefact Correction on the Quantification of Heart Rate Variability in Three Independent Human Cohorts. J Clin Med 2020 Jan 23;9(2):325. PMID: 31979367

38 Adelantado-Renau M, Beltran-Valls MR, **Migueles JH**, Artero EG, Legaz-Arrese A, Capdevila-Seder A, Moliner-Urdiales D. Associations between objectively measured and self-reported sleep with academic and cognitive performance in adolescents: DADOS study. J Sleep Res 2019;28(4):e12811. PMID: 30609171

39 Acosta FM, Sanchez-Delgado G, Martinez-Tellez B, **Migueles JH**, Amaro-Gahete FJ, Rensen PCN, Llamas-Elvira JM, Blondin DP, Ruiz JR. Sleep duration and quality are not associated with brown adipose tissue volume or activity-as determined by 18F-FDG uptake, in young, sedentary adults. Sleep 2019;42(12):1–10. PMID: 31555815

40 Acosta FM, Martinez-Tellez B, Sanchez-Delgado G, **Migueles JH**, Contreras-Gomez MA, Martinez-Avila WD, Merchan-Ramirez E, Alcantara JMAA, Amaro-Gahete FJ, Llamas-Elvira JM, Ruiz JR. Association of Objectively Measured Physical Activity With Brown Adipose Tissue Volume and Activity in Young Adults. J Clin Endocrinol Metab United States; 2019 Feb 1;104(2):223–233. PMID: 30137350

41 Aadland E, Andersen LB, **Migueles JH**, Ortega FB, Kvalheim OM. Interpretation of associations between the accelerometry physical activity spectrum and cardiometabolic health and locomotor skills in two cohorts of children using raw, normalized, log-transformed, or compositional data. J Sports Sci 2020 Jul 29;1–12. PMID: 32723017

42 Ubago-Guisado E, Gracia-Marco L, Medrano M, Cadenas-Sanchez C, Arenaza L, **Migueles JH**, Mora-Gonzalez J, Tobalina I, Escolano-Margarit MV, Oses M, Martín-Matillas M, Labayen I, Ortega FB. Differences in areal bone mineral density between metabolically healthy and unhealthy overweight/obese children: the role of physical activity and cardiorespiratory fitness. Pediatr Res Springer US; 2020;87(7):1219–1225. PMID: 31822016

43 Solis-Urra P, Olivares-Arancibia J, Suarez-Cadenas E, Sanchez-Martinez J, Rodríguez-Rodríguez F, Ortega FB, Esteban-Cornejo I, Cadenas-Sanchez C, Castro-Piñero J, Veloz A, Chabert S, Saradangani KP, Zavala-Crichton JP, **Migueles JH**, Mora-Gonzalez J, Quiroz-Escobar M, Almonte-Espinoza D, Urzúa A, Dragicevic CD, Astudillo A, Méndez-Gassibe E, Riquelme-Uribe

D, Azagra MJ, Cristi-Montero C. Study protocol and rationale of the "Cogni-action project" a cross-sectional and randomized controlled trial about physical activity, brain health, cognition, and educational achievement in schoolchildren. BMC Pediatr BMC Pediatrics; 2019;19(1):260. PMID: 31349791

44 Martín-Matillas M, Mora-Gonzalez J, **Migueles JH**, Ubago-Guisado E, Gracia-Marco L, Ortega FB. Validity of Slaughter Equations and Bioelectrical Impedance Against Dual-Energy X-Ray Absorptiometry in Children. Obesity (Silver Spring) 2020;28(4):803–812. PMID: 32144886

45 Cadenas-Sánchez C, Mora-González J, **Migueles JH**, Martín-Matillas M, Gómez-Vida J, Escolano-Margarit MV, Maldonado J, Enriquez GM, Pastor-Villaescusa B, de Teresa C, Navarrete S, Lozano RM, de Dios Beas-Jiménez J, Estévez-López F, Mena-Molina A, Heras MJ, Chillón P, Campoy C, Muñoz-Hernández V, Martínez-Ávila WD, Merchan ME, Perales JC, Gil Á, Verdejo-García A, Aguilera CM, Ruiz JR, Labayen I, Catena A, Ortega FB. An exercise-based randomized controlled trial on brain, cognition, physical health and mental health in overweight/obese children (ActiveBrains project): Rationale, design and methods. Contemp Clin Trials Elsevier Inc.; 2016 Mar;47:315–24. PMID: 26924671

46 Mora-Gonzalez J, Esteban-Cornejo I, **Migueles JH**, Rodriguez-Ayllon M, Molina-Garcia P, Cadenas-Sanchez C, Solis-Urra P, Plaza-Florido A, Kramer AF, Erickson KI, Hillman CH, Catena A, Ortega FB. Physical fitness and brain source localization during a working memory task in children with overweight/obesity: The ActiveBrains project. Dev Sci 2020 Oct 9;e13048. PMID: 33037758

Books and book chapters

1 **Migueles JH**, …, Ortega FB. Guía para la prevención y tratamiento de la obesidad infantil: experiencias de los proyectos ActiveBrains y EFIGRO.
In process

2 Patricio Solis-Urra, **Migueles JH**. *Medidas de evaluación del nivel de actividad física*. En Actividad física en la salud y enfermedad del niño y el adolescente.
In production

Distinguished communications in conferences (total = 66)

Oral communications

26/06/2019 | *How do body attachment site and aggregation metrics affect accelerometer physical activity?*
Migueles JH, …, Ortega FB
International Conference on Ambulatory Monitoring of Physical Activity and Movement (ICAMPAM), Maastricht (the Netherlands)

15/02/2019 | *Is physical activity associated with gray matter volume in children with overweight/obesity?*
Migueles JH, …, Ortega FB
National congress of researchers PTS, Granada (Spain)

23/06/2017 | *Which is the best accelerometer-based metric to predict free-living energy expenditure?*
Migueles JH, …, **Ortega FB**, Löf M
International Conference on Ambulatory Monitoring of Physical Activity and Movement (ICAMPAM), Bethesda MD (USA)

19/07/2017 | *Collaborator speaker at the International Symposium ActiveBrains for all* (1 hour)
Migueles JH, …, **Ortega FB**
Activebrains for all, Granada (Spain)

Poster presentations

19/10/2017 | *An evaluation of the accuracy of 4 different motion sensors at self-paced walking overground*
Migueles JH, ..., Tudor-Locke C
New England American College of Sports Medicine, Providence RI (USA)

12/06/2017 | *Combined effects of sleep behavior, sedentary time and physical activity on academic achievement in overweight/obese children: Results from the ActiveBrains project*
Migueles JH, ..., **Ortega FB**
ActiveBrains for All, Granada (Spain)

Participation in research projects

2019-2022 | *HealthyMoms: smartphone app to promote healthy weight during pregnancy*
PI: M Löf | **Role:** adviser | **Funding:** €1,500,000

2019-2021 | *Cogni-Action: physical activity, brain health, cognition and educational achievement in children*
PI: C Cristi-Montero | **Role:** research assistant | **Funding:** €90,000

2018-2020 | *CoCA: Comorbid Conditional of Attention-deficit Hyperactivity Disorder (Granada, EU-funded project)*
Co-PI: FB. Ortega | **Role:** research assistant | **Funding:** €105,000

2018-2020 | *Australian based randomized controlled trial of physical activity for people with mental illness*
PI: Justin Chapman | **Role:** adviser | **Funding:** $115,000

2018-2020 | *SmarterMove: Exercise to prevent and treat obesity and insulin resistance*
PI: FB Ortega / JR Ruiz | **Role:** research assistant | **Funding:** €121,000

2017-2019 | *MINISTOP: Mobile-based intervention intended to stop obesity in Preschoolers*
PI: M Löf | **Role:** adviser | **Funding:** €265,000

2017-2019 | *ActiveBrains: Effects of exercise on brain and physical health in overweight/obese children*
PI: FB. Ortega | **Role:** research assistant | **Funding:** €145,200

2017-2019 | *ACTIBATE: Activating brown adipose tissue through exercise in young adults*
PI: JR. Ruiz | **Role:** research assistant | **Funding:** €120,000

2017-2018 | *Pilot study on accelerometer-based assessment of physical activity*
PI: FB. Ortega | **Role:** project manager | **Funding:** Non-funded

Teaching experience

2017-2019 | Fundamentals of handball (90 hours)
BSc in physical activity and sport sciences

2019-2020 | *Physical activity and health* (70 hours)
BSc in physical activity and sport sciences

2019-2020 | *New trends in fitness* (20 hours)
BSc in physical activity and sport sciences

International stays

Research stays

26/01/2020 – 28/02/2020 | University of Pittsburgh
Pittsburgh PA (USA)
Supervisor: PhD Kirk I Erickson

31/05/2019 – 07/07/2019 | Karolinska Institutet / Linköping University
Stockholm and Linköping (Sweden)
Supervisor: Prof Marie Löf

16/09/2018 – 23/09/2018 | University of Pittsburgh
Pittsburgh PA (USA)
Supervisor: PhD Kirk I Erickson

01/08/2018 – 28/10/2018 | National Institute on Aging
Baltimore MD (USA)
Supervisor: PhD Eric J Shiroma

08/07/2018 – 14/07/2018 | eScience Center
Amsterdam (the Netherlands)
Supervisor: PhD Vincent T van Hees

29/05/2018 – 08/06/2018 | Karolinska Institutet / Linköping University
Stockholm and Linköping (Sweden)
Supervisor: Prof Marie Löf

25/11/2017 – 19/12/2017 | Northeastern University
Boston MA (USA)
Supervisor: Prof Charles H Hillman

28/08/2017 – 25/11/2017 | University of Massachusetts Amherst
Amherst MA (USA)
Supervisor: Prof Catrine Tudor-Locke

Teaching stays

26/01/2019 – 02/02/2019 | Universitá Degli Studi di Roma "Foro Italico"
Rome (Italy)

21/04/2018 – 28/04/2018 | University of Krakow
Krakow (Poland)

Invited lectures, seminars and workshops

17/04/2020 | *Physical activity and cognition in children with overweight/obesity* (8 hours)
BSc in physical education, Prague (Czech Republic) – Online lecture

16/04/2020 | *Physical activity and sleep assessment with accelerometers* (2 hours)
MSc in research on physical activity and sport (UGR), Granada (Spain)

28/02/2020 | *Introductory workshop to accelerometer assessment of physical activity* (4 hours)
MSc in research on physical activity and sport (UCLM), Toledo (Spain)

26/04/2019 | *Sedentary time, physical activity and sleep with accelerometers* (2 hours)

MSc in research on physical activity and sport (UGR), Granada (Spain)

02/02/2019 | *Effects of physical activity interventions on brain, cognition and academic achievement* (8 hours)
BSc in physical activity (Universitá Degli Studi "Foro Italico"), Rome (Italy)

22/04/2018 | *Brain, cognition and fitness and physical activity* (8 hours)
BSc in physical education (University of Krakow), Krakow (Poland)

17/04/2018 | *Objective assessment of physical activity and sleep* (2 hours)
MSc in research on physical activity and sport (UGR), Granada (Spain)

19/02/2018 | *Physical activity: concept and measurement* (4 hours)
MSc in food, exercise and sport for health (UGR), Granada (Spain)

19/07/2017 | *Workshop on assessment of physical activity with accelerometers* (5 hours)
Open seminar in University of Almería, Almería (Spain)

26/04/2017 | *Objective assessment of physical activity and sleep* (2 hours)
MSc in research on physical activity and sport (UGR), Granada (Spain)

22/04/2017 | *Physical activity metrics from accelerometer: raw and count based data* (2 hours)
Research group seminar (UGR), Granada (Spain)

16/02/2017 | *Workshop on accelerometer-based measurement of physical activity* (2 hours)
PREVIENE project training program (UGR), Granada (Spain)

22/04/2016 | *Objective assessment of physical activity and sleep* (2 hours)
MSc in research on physical activity and sport (UGR), Granada (Spain)

Outreach activities

Present | Co-developer of *"GGIR: raw accelerometer data analysis"* software (R package)

30/09/2020 | Press release about the study on sleep and grey matter volume in children with overweight or obesity published in Pediatric Obesity. The press release appeared on the major national media and I was interviewed by the National and some local radio stations about this study.

21/10/2019 | Organization of the *"International Workshop: a focus on statistical methods to analyse accelerometer-determined physical activity"* (University of Granada, Spain)

12/06/2017 | Organization of the *"International Symposium ActiveBrains for all: exercise, cognition and mental health"* (University of Granada, Spain)

29/01/2018 - 28/02/2018 | Attendance to the "*Scientific communication and dissemination*" course and participation in the "*Three-Minute Thesis*" contest (classified as finalist) intended to communicate the thesis rationale and findings to the general population (video here).

2018 - 2019 | Writing a novel with a background on Alzheimer Disease and participation in a national contest (awarded with the first prize).

13/05/2020 | Press release on a communication activity using social media to promote physical activity during the COVID19 confinement (press note here). I performed this

activity within the *"Physical activity and Health"* module in the BSc in physical activity and sport sciences (University of Granada).

Funding obtained

2020 | *Collaboration project with Accelting on GGIR functionalities development*
Funder: Accelting© | **Funding:** 4,000 €

2019-2020 | *Erasmus+ teaching*
Funder: ERASMUS | **Funding:** 1,200 €

2018-2019 | *Erasmus+ teaching*
Funder: ERASMUS | **Funding:** 1,200 €

2018-2019 | *International mobility grant*
Funder: Ministry of Education, Culture and Sports |**Funding:** 3,406 €

2018-2019 | *Funding for congresses participation*
Funder: University of Granada | **Funding:** 600 €

2017-2018 | *Erasmus+ teaching*
Funder: ERASMUS | **Funding:** 1,200 €

2017-2018 | *International mobility grant*
Funder: Ministry of Education, Culture and Sports |**Funding:** 5,248 €

2016-2017 | *Funding for congresses participation*
Funder: University of Granada | **Funding:** 600 €

2016-2017 | *International mobility grant*
Funder: University of Granada | **Funding:** 2,600 €

2016-2020 | *University Teachers Training (main grant for PhD studies from the Spanish Ministry)*
Funder: Ministry of Education, Culture and Sports |**Funding:** 80,000 €

2015-2016 | *Research assistant in the ActiveBrains project*
Funder: University of Granada | **Funding:** 17,000 €

2014-2015 | *Starting research grant (for MSc students)*
Funder: University of Granada | **Funding:** 1,800 €

Awards

2019 | Winner of the *"Athenea"* short stories contest

2018 | Special award for BSc in physical activity and sport sciences

2018 | Finalist in the University of Granada phase of the *"Three-minute thesis"* contest

2016 | Special recognition for publishing a study in a 1st-decile JCR journal using the starting research grant funds