OULU 2021
D 1616

ACTA UNIVERSITATIS OULUENSIS

Vahid Farrahi

SEDENTARY TIME, PHYSICAL ACTIVITY AND CARDIOMETABOLIC HEALTH

ACCELEROMETRY-BASED STUDY IN THE NORTHERN FINLAND BIRTH COHORT 1966

*UNIVERSITY OF OULU GRADUATE SCHOOL;
UNIVERSITY OF OULU,
FACULTY OF MEDICINE;
MEDICAL RESEARCH CENTER OULU;
OULU UNIVERSITY HOSPITAL;
OULU DEACONESS INSTITUTE FOUNDATION SR*

ACTA UNIVERSITATIS OULUENSIS
D Medica 1616

VAHID FARRAHI

SEDENTARY TIME, PHYSICAL ACTIVITY AND CARDIOMETABOLIC HEALTH

Accelerometry-based study in the Northern Finland Birth Cohort 1966

Academic dissertation to be presented with the assent of the Doctoral Training Committee of Health and Biosciences of the University of Oulu for public defence in the Leena Palotie auditorium (101A) of the Faculty of Medicine (Aapistie 5 A), on 14 June 2021, at 12 noon

UNIVERSITY OF OULU, OULU 2021

Acta Univ. Oul. D 1616, 2021

Supervised by
Professor Timo Jämsä
Professor Raija Korpelainen
Doctor Maarit Kangas

Reviewed by
Docent Timo Rantalainen
Assistant Professor Alexander Montoye

Opponent
Associate Professor Francisco B. Ortega

ISBN 978-952-62-2945-4 (Paperback)
ISBN 978-952-62-2946-1 (PDF)

ISSN 0355-3221 (Printed)
ISSN 1796-2234 (Online)

Cover Design
Raimo Ahonen

PUNAMUSTA
TAMPERE 2021

Farrahi, Vahid, Sedentary time, physical activity and cardiometabolic health. Accelerometry-based study in the Northern Finland Birth Cohort 1966
University of Oulu Graduate School; University of Oulu, Faculty of Medicine; Medical Research Center Oulu; Oulu University Hospital; Oulu Deaconess Institute Foundation sr.
Acta Univ. Oul. D 1616, 2021
University of Oulu, P.O. Box 8000, FI-90014 University of Oulu, Finland

Abstract

The popularity of accelerometer-based activity monitors has been associated with several analytical challenges, including how to quantify accelerometer outputs in terms of sedentary behavior, light-intensity physical activity (LPA), and moderate-to-vigorous physical activity (MVPA). Recently, machine learning (ML) approaches have been coupled with raw accelerometry to classify activities by intensity, but the generalizability of ML models outside of the development datasets remains poorly understood. Currently, the health benefits of meeting the recommended amounts of sleep and MVPA in adults are well documented, but the cardiometabolic health implications of sedentary time and LPA are still unclear.

The present study reviewed studies calibrating and validating wearable accelerometers using ML approaches and preformed cross-dataset tests to evaluate the generalization performance of ML models for classifying activity intensities from raw acceleration data. Additionally, the latest follow-up in the Northern Finland Birth Cohort 1966 study (n = 5,840) at age 46 years included measurement of daily activities for two weeks with two accelerometers. This data was used to examine how the levels and patterns of accelerometer-estimated activity intensities (sedentary behavior, LPA, and MVPA) are associated with cardiometabolic health in this large sample of middle-aged adults, and to create a data-driven hierarchy predicting their activity behaviors.

Based on the study, ML techniques can classify activities in terms of type, category, or intensity with acceptable accuracy irrespective of accelerometer placement. However, ML models developed with raw acceleration data for classifying activity intensities (sedentary behavior, LPA, and MVPA) are not generalizable to other populations monitored with different accelerometers, suggesting that further strategies are needed to enhance their generalizability. The study suggests that adults, in addition to MVPA, may also gain cardiometabolic health benefits through LPA, particularly when it replaces sedentary time. Finally, the data-driven hierarchy of correlates created consisted of factors of relative importance, and can potentially be used to target and tailor interventions.

Keywords: adiposity, data mining, dyslipidemias, insulin resistance, machine learning, metabolic diseases

Farrahi, Vahid, Paikallaanolo, fyysinen aktiivisuus ja sydänterveys. Kiihtyvyysanturimittaukseen perustuva Pohjois-Suomen syntymäkohortti 1966 -tutkimus
Oulun yliopiston tutkijakoulu; Oulun yliopisto, Lääketieteellinen tiedekunta; Medical Research Center Oulu; Oulun yliopistollinen sairaala; Oulun Diakonissalaitoksen säätiö sr.
Acta Univ. Oul. D 1616, 2021
Oulun yliopisto, PL 8000, 90014 Oulun yliopisto

Tiivistelmä

Nykyään hyvin suosittujen kiihtyvyysanturiin perustuvien aktiivisuusmittareiden keräämän datan analysointiin liittyy monia haasteita, kuten paikallaanolon, kevyen liikunnan sekä keskiraskaan ja raskaan liikunnan tarkan määrän määrittäminen. Viime aikoina on otettu käyttöön koneoppimismenetelmiä kiihtyvyysanturin tuottaman raakasignaalin analysoinnissa luokittelemaan liikettä sen intensiteetin perusteella, mutta toistaiseksi näiden menetelmien yleistettävyys on huonosti tiedossa. Nykyisin tiedetään aika hyvin ne terveyshyödyt, joita saadaan, jos noudatetaan unen sekä keskiraskaan ja raskaan liikunnan suosituksia. Paikallaanolon ja kevyen liikunnan vaikutukset sydän- ja verisuoniterveyteen ovat kuitenkin heikommin tiedossa.

Tässä tutkimuksessa tehtiin systemaattinen kirjallisuuskatsaus koneoppimismenetelmien käytöstä kannettavien kiihtyvyysanturien kalibroinnissa ja validoinnissa. Työssä testattiin koneoppimismenetelmien yleistettävyyttä fyysisen aktiivisuuden intensiteetin luokitteluun kiihtyvyysanturin antaman raakadatan perusteella yhdistäen useita toisistaan riippumattomia mittausaineistoja. Pohjois-Suomen vuoden 1966 syntymäkohortin 46-vuotisaineistonkeruussa (n = 5,840) oli mitattu liikunta-aktiivisuutta kahdella kiihtyvyysanturilla. Tämän mittaustiedon avulla tutkittiin sitä, kuinka kiihtyvyysanturilla mitattu fyysisen aktiivisuuden intensiteetti (paikallaanolo, kevyt liikunta sekä keskiraskas ja raskas liikunta) ja eri intensiteetillä toteutetun aktiivisuuden jakautuminen vuorokauden sisällä ovat yhteydessä keski-ikäisten sydänterveyteen. Lisäksi luotiin aineiston perusteella hierarkinen malli ennustamaan liikuntakäyttäytymistä.

Tutkimuksen perusteella koneoppimistekniikoiden avulla voidaan riittävällä tarkkuudella luokitella fyysistä aktiivisuutta liikuntamuodon, intensiteetin ja eri intensiteettien jakautumisen perusteella riippumatta kiihtyvyysanturin sijainnista. Kiihtyvyysanturin tuottamaan raakadataan perustuvat fyysisen aktiivisuuden intensiteetin luokitteluun kehitetyt koneoppimismallit eivät ole kuitenkaan yleistettävissä muihin väestöryhmiin, joissa on käytetty erilaisia kiihtyvyysantureita, vaan tarvitaan lisätutkimusta parantamaan mallien yleistettävyyttä. Tutkimuksen perusteella keskiraskaan ja raskaan liikunnan lisäksi kevytkin liikunta-aktiivisuus, erityisesti jos se korvaa paikallaan oloa, on yhteydessä aikuisten parempaan sydänterveyteen. Aineiston perusteella luotu hierarkinen malli antoi tietoa useista sydänterveyttä edistävistä tekijöistä ja sitä voidaan käyttää liikuntainterventioiden räätälöinnissä.

Asiasanat: insuliiniresistenssi, koneoppiminen, rasva-aineenvaihdunnan häiriöt, rasvakudos, tiedonlouhinta aineenvaihduntasairaudet

To my family and friends

Acknowledgements

This study was performed at the Research Unit of Medical Imaging, Physics, and Technology (MIPT) at the University of Oulu between 2017 and 2021.

I owe my deepest gratitude to my supervisors Professor Timo Jämsä, Professor Raija Korpelainen, and Maarit Kangas PhD, for all kinds of support they have provided since the day I started my doctoral research. Timo, Raija, and Maarit; the things that I have learned from you during my doctoral journey are countless, but most importantly you made me a more confident researcher. Thanks to your guidance and critical comments, I have learned how to perform independent research. From the first day I started working with you, every day I have felt that I am better at science and teaching than the day before.

I also wish to thank my colleagues and co-authors Maisa Niemelä PhD, Mikko Kärmeniemi MSc, Soile Puhakka MSc, and Antti Kiviniemi PhD for their significant contributions to my studies. Maisa, my research would not be possible without your help. Sometimes, I want to know too much too quickly. I want to thank you in particular for replying to all my inquiries and questions about the data and other issues promptly and precisely. Mikko, I appreciate your time and patience. Also, thanks for all the information you shared generously with me when I was doing my first study. Soile, you were of great help whenever I came to you with dozens of questions.

I would also like to thank other friends and colleagues Dr. Jérôme Thevenot, University Lecturer Mikko Finnilä, Robel Gebre, Bijay Shakya, Pragya Shrestha, Vesa Virtanen, Shuvashis Das Gupta, Anahita Gheisari, and Shahab Khakpour, who were not directly involved in my studies but surely made a great impact on my so-called PhD life. Jérôme and Mikko, you have always been supporting, and it is always nice to chat with you about science and other matters. Robel, you are an excellent friend and office mate. We have a shared office together, and chatting with you has been always fun. Bijay and Pragya, thanks for all your help and advice both in the workplace and outside, and also for being two of my good friends since the day I came to Finland. Shahab, I have always had great conversations with you, and I am thankful to you for being a good listener in all the discussions we had. I also wish to say a big thank you to all the other MIPT members for their inspirational work and support. It has been truly wonderful to be part of this excellent research unit.

A significant part of my academic career was shaped during my research visit to Bradford Research Institute for Health (BRIH) in the UK. I wish to thank Paul

J. Collings PhD, who hosted me in the UK and taught me many priceless statistical skills. Paul, my statistical skills reached another level after my visit to BRIH. I am truly thankful to you for being a friend and a mentor to me. I also express my deep gratitude to Daniel Bingham PhD from the BRIH, who was also kind and helpful during my research visit. During my stay in the UK I was also lucky to visit the University of Oxford. I would like to thank Rosemary Walmsley at the University of Oxford for all her statistical advice and for organizing my visit to her excellent research group.

I also sincerely appreciate my parents and brother for their honest love for me. Both of my parents and my brother have always been supportive throughout each stage of life and study. They have always trusted and believed in me. Mom and Dad, I am thankful for all the lessons you have given me, and especially for making me a stubborn person for learning new things.

This study received funding from the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska–Curie grant agreement no. 713645. This study would not have been possible without the BioMEP doctoral program. I would like to thank all my BioMEP fellows for giving me a chance to be part of this unique doctoral program. I also wish to thank all the participants of the Northern Finland Birth Cohorts (NFBC), especially the members of cohort NFBC1966. The NFBC cohorts are excellent sources of data for many scientists, and this is certainly thanks to the dedication of their members for donating high-quality data for medical research. Finally, I wish to thank Assistant Professor Alexander Montoye and Docent Timo Rantalainen for the pre-examination of this study and their valuable comments.

Oulu, January 2021 Vahid Farrahi

Abbreviations

3D	Three dimensional
ANN	Artificial neural network
ANOVA	One-way analysis of variance
β	Regression coefficient from compositional models
B	Regression coefficient from generalized linear mixed models
BMI	Body mass index
CHAID	Chi-squared Automatic Interaction Detection
CI	Confidence interval
cm	Centimeter
DSA	Daily and Sports Activities dataset
H	Hip
HDL	High-density lipoprotein
HOMA-IR	Homeostasis model assessment of insulin resistance
hours/day	Hours per day
kg	Kilogram
LDL	Low-density lipoprotein
LDL/HDL	LDL to HDL cholesterol ratio
LOSO	Leave-one-subject-out
LPA	Light-intensity physical activity
m	Meter
MAD	Mean amplitude deviation
MEMS	Micro electro-mechanical system
MET	Metabolic equivalents
min/day	Minutes per day
ML	Machine learning
MVPA	Moderate-to-vigorous physical activity
NFBC1966	Northern Finland Birth Cohort 1966
OSU	Oregon State University dataset
PAMAP	PAMAP2 Physical Activity Monitoring dataset
RI	Response index
ROC	Receiver operating characteristics
SB	Sedentary behavior
SD	Standard deviation
Total/HDL	Total to HDL cholesterol ratio
UOULU	University of Oulu dataset

W	Wrist
etc.	et cetera
i.e.	id est
e.g.	exempli gratia

List of original publications

This thesis is based on the following publications, which are referred to throughout the text by their Roman numerals:

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

II Farrahi, V., Niemelä, M., Tjurin, P., Kangas, M., Korpelainen, R., & Jämsä, T. (2020). Evaluating and enhancing the generalization performance of machine learning models for physical activity intensity prediction from raw acceleration data. *IEEE Journal of Biomedical and Health Informatics*, *24*(1), 27–38. doi: 10.1109/JBHI.2019.2917565

III Farrahi, V., Kangas, M., Walmsley, R., Niemelä, M., Kiviniemi, A., Puukka, K., Collings, P., Korpelainen, R., & Jämsä, T. (2021). Compositional associations of sleep and activities within the 24-h cycle with cardiometabolic health markers in adults. *Medicine and Science in Sports and Exercise*, *53*(2), 324–332. doi: 10.1249/MSS.0000000000002481

IV Farrahi, V., Kangas, M., Kiviniemi, A., Puukka, K., Korpelainen, R., & Jämsä, T. (2021). Accumulation patterns of sedentary time and breaks and their association with cardiometabolic health markers in adults. *Scandinavian Journal of Medicine & Science in Sports*, Online ahead of print. doi: 10.1111/sms.13958

V Farrahi, V., Niemelä, M., Kärmeniemi, M., Puhakka, S., Kangas, M., Korpelainen, R., & Jämsä, T. (2020). Correlates of physical activity behavior in adults: a data mining approach. *International Journal of Behavioral Nutrition and Physical Activity*, *17*(94). doi: 10.1186/s12966-020-00996-7

The author of this doctoral thesis designed all the studies (I–V), performed the statistical and data analyses (I–V), and drafted the manuscripts (I–V). In sub-study I, the author performed the initial literature search and screened all identified articles. In sub-studies II–V, the author performed all the analyses with previously collected data. In studies III and IV, the author was responsible for cleaning, processing, and analyzing the accelerometry data.

Contents

1 Introduction

Cardiometabolic diseases, including cardiovascular diseases, stroke, and diabetes are reportedly among the leading causes of deaths globally (WHO, 2013). Physical activity is shown to reduce the risk of cardiometabolic diseases, mortality, and many other adverse health conditions (Grgic et al., 2018; Guthold et al., 2018; Lee et al., 2012; WHO, 2010). Much of the evidence generated to date—on which current physical activity guidelines have largely been based—has been derived from self-reported moderate-to-vigorous physical activity (MVPA), typically defined as an energy expenditure of three metabolic equivalents (MET) or more (Guthold et al., 2018; Kohl 3rd et al., 2012; Rosenberger et al., 2019).

Recently it has been recognized that movement intensities represent a continuum, suggesting that activities across a more broadened spectrum including light-intensity physical activities (LPA; activities at 1.5 to 3 MET) and sedentary behaviors (activities at under 1.5 MET in sitting or reclined posture), and their patterns may also be related to cardiometabolic health in adults (Chastin et al., 2019; Ekelund et al., 2019; Tremblay et al., 2017). This is, in large part, due to the availability of wearable accelerometer-based activity monitors (Lee et al., 2014).

Although advantageous, the proliferation of accelerometers has also been associated with several considerable challenges regarding how to collect, use, and most importantly quantify accelerometer outputs (signals) across the entire intensity continuum in terms of sedentary behaviors, LPA, and MVPA (Migueles et al., 2017; Wijndaele et al., 2015). There are many algorithms for translating accelerometer outputs into activity intensities, yet none has been recognized as reliable enough to be adopted as a standardized method, partly because validation of those algorithms outside the laboratory have typically shown poor performance (Bassett Jr et al., 2012; Freedson et al., 2012; Wijndaele et al., 2015). A reliable data processing methodology for accelerometer data has continued to remain a challenge even after the promising shift toward machine learning (ML) approaches (Bassett Jr et al., 2012), despite their better predictive capabilities (Ellis et al., 2016; Montoye, Westgate, et al., 2018).

Interstudy comparison among existing studies is currently limited due to the inconsistency of methodologies used for processing accelerometer signals (Migueles et al., 2019; Wijndaele et al., 2015). Hence, based on systematic review of literature, even small amounts of physical activity from light-intensity upwards may confer cardiometabolic health benefits in adults (Chastin et al., 2019; Füzéki et al., 2017). Research has also recognized that time spent in each activity may modify the health-related influence of time spent in any of the other activities (Pedišić et al., 2017; Rosenberger et al., 2019). For instance, breaking up sedentary time with physical activity could reduce the detrimental effects on cardiometabolic markers caused by sedentary behavior (Healy et al., 2008, 2011). However, it is still unclear how and when sedentary time should be interrupted before it becomes detrimental to cardiometabolic health in adults (Chastin, Egerton, et al., 2015). Sleep, sedentary behavior, LPA, and MVPA over a 24-hour sleep and wake cycle seems to be codependently related to health (Migueles et al., 2021; Pedišić et al., 2017). Yet little is known about the combined associations of these activities and their patterns with cardiometabolic health in adults (Rosenberger et al., 2019).

Despite considerable health benefits, few adults meet the current recommendation for physical activity (Guthold et al., 2018). Many studies have therefore focused on identifying the factors associated with activity behaviors, and examined the association between various factors including personal, societal, and environmental factors with different indices such as the daily amount of MVPA or sedentariness (Bauman et al., 2012; O'Donoghue et al., 2016). However, activity behavior has been found to be a multidimensional behavior that is explained by a complex web of factors, followed by requests for more complex statistical approaches in correlates research to better understand why some adults are active but others are not (Bauman et al., 2012; Chastin et al., 2016).

The present study focused on the in-depth understanding of the levels and patterns of activity across the entire intensity continuum (sedentary behavior, LPA, and MVPA) and their associations with cardiometabolic health markers in a large population-based sample of middle-aged Finnish adults. Additionally, the present study reviewed studies calibrating and validating wearable accelerometers using ML approaches, developed robust ML models for classifying activity intensities from raw accelerometry, and performed comprehensive and powerful analyses of the correlates of activity behaviors and identification of target subgroups for future intervention.

2 Review of the literature

2.1 Physical activity, inactivity, and health in adults

Any bodily movement produced by skeletal muscles that requires energy expenditure above the basal is defined as physical activity (WHO, 2010). Regular physical activity has considerable health benefits such as reduced risk of mortality and cardiometabolic diseases (Guthold et al., 2018; Lee et al., 2012; Warburton et al., 2017), that outweighs potential risk of adverse events, for example through accidents (WHO, 2010). Physical activity can be done at any level of skill and for enjoyment (Bauman et al., 2012), and common ways to remain active are by walking, cycling, sports, and active games (WHO, 2010). Even people with poor mobility may benefit from physical activity (three or more times per week), for example to enhance balance and prevent falls (WHO, 2010).

Conversely, physical inactivity (insufficient physical activity) is the fourth leading cause of death worldwide, and was responsible for more than five million of the 57 million deaths that occurred worldwide in 2008 (Lee et al., 2012; WHO, 2010). It is also a risk factor for major noncommunicable diseases (Lee et al., 2012). However, more than one in four adults worldwide (in total more than 1.4 billion adults) fail to meet the minimum recommended level of physical activity that is 150 minutes of moderate-intensity aerobic exercise or 75 minutes of vigorous-intensity exercise per week (Bull et al., 2020; Guthold et al., 2018).

Active adults compared to their inactive peers have improved quality of life and well-being (O'Donovan et al., 2017; Wen et al., 2011), and thus an increased likelihood of a healthier ageing (Daskalopoulou et al., 2017). The world is currently experiencing a considerable increase in the number of people living longer than 60 years. Reportedly, the number of people aged 60 years or older worldwide is projected to grow more than twofold by 2050, reaching a total number of two billion and outnumbering the number of children by 2047 (Chatterji et al., 2015). However, it appears that increased longevity is not necessarily accompanied by an extended period of good health, and little evidence exists that older people today are experiencing an extended period of health compared to previous generations at the same age (Beard et al., 2016).

The higher risk of mortality and adverse health conditions in later life has, to a great extent, been attributable to a hazardous accumulated lifestyle, including insufficient physical activity during adulthood (Chatterji et al., 2015; Khaw et al., 2008). Research has shown that habitual patterns of physical activity starts to develop very early during preschool age (Telama et al., 2014). Although the causal relationship between high physical activity in adulthood and reduced risk of premature death is yet to be proven (Kujala, 2018), it is already well-documented that physical activity is modifiable and health-enhancing behavior for nearly everyone at every life stage (Bauman et al., 2012; Bull et al., 2020; Wen et al., 2011; WHO, 2010). The importance of physical activity during adulthood therefore seems to be twofold: An active lifestyle throughout adulthood may improve quality of life and reduce the risks of major diseases and health conditions such as cardiometabolic diseases, heart disease, stroke and diabetes, and it helps to preserve this well-being and health into old age (Bull et al., 2020; Chatterji et al., 2015; Lee et al., 2012; Wen et al., 2011).

2.1.1 Light and moderate-to-vigorous intensity physical activity and cardiometabolic health

Recent evidence shows an L-shaped dose response between the total volume of physical activity and risk of cardiometabolic diseases and mortality in adults (Kyu et al., 2016; Samitz et al., 2011). Existing studies have typically considered achieving a higher physical activity volume through MVPA (Guthold et al., 2018; Rosenberger et al., 2019; Warburton et al., 2017), which also appears to be the most potent health-enhancing activity intensity, even after accounting for other activities during the day (Chastin, Palarea-Albaladejo, et al., 2015; McGregor, Palarea-Albaladejo, Dall, del Pozo Cruz, et al., 2019). However, previous studies have mostly used self-reported MVPA to investigate the associations between this movement intensity and different health indicators including the markers of cardiometabolic health in adults (Guthold et al., 2018; Warburton et al., 2017).

Measurement method may have a significant impact on the observed levels of physical activity in adults (Prince et al., 2008). Mostly based on self-reported physical activity data, guidelines for physical activity have generally recommended 150 minutes per week of MVPA for health benefits in adults (Rosenberger et al., 2019; Warburton et al., 2016). However, with the technological advances and the availability of device-based methods that allow for relatively more accurate and detailed measurement of daily activities than self-reported measures, studies have shown that even a lower level of volume and/or intensity of physical activity may confer significant health benefits especially for inactive adults (Chastin, Palarea-Albaladejo, et al., 2015; Strain et al., 2020; Warburton et al., 2016, 2017). This has resulted in the emergence of studies challenging the current threshold-based messaging related to physical activity and health, and suggesting that any amount of physical activity could be beneficial for adults (Bull et al., 2020; Ding et al., 2020; Warburton et al., 2016, 2017). Still, more studies with device-based measurement of physical activity are required to warrant these findings.

More recently, studies have shown that activities at lighter intensities may also confer considerable health benefits (Amagasa et al., 2018; Chastin et al., 2019; Füzéki et al., 2017). This is important because LPA promotion could potentially be a more feasible and sustainable means to increase volume of physical activity, especially for inactive people (Pulsford et al., 2017). High-intensity physical activity such as brisk walking might not be feasible in many situations (e.g., at work or home) or may even be strenuous for some (Chastin et al., 2019). On the other hand, LPA does not necessarily require planning or dedicated time commitment as it usually involves incidental daily living and is accessible regardless of physical fitness and even during occupational time (e.g., ambulation, and convenient and causal walking) (Buman et al., 2017). Still, missing from most physical activity guidelines are detailed recommendations for light-intensity physical activities (Lee et al., 2014; Rosenberger et al., 2019). Although more research is still required, there are studies indicating that even replacing sedentary time with standing could confer some cardiometabolic health benefits for adults (Healy et al., 2015; Saeidifard et al., 2018), and may therefore be a potential solution for a sedentary lifestyle.

Adults spend considerably more daily time in LPA than in MVPA (Chaput et al., 2014; Spittaels et al., 2012), yet little is known about the health-enhancing potential of LPA, and even less about biological mechanistic information explaining the observed associations and effects of this intensity of physical activity in humans (Chastin et al., 2019; Füzéki et al., 2017). To date, several population-based studies have found that higher levels of LPA on a daily basis may confer mortality and cardiometabolic health benefits for adults (Ekblom-Bak et al., 2016; Fishman et al., 2016; Howard et al., 2015; Pulsford et al., 2017), but detailed information on the timing and frequency is still lacking (Rosenberger et al., 2019). This appears to be in large part due to insufficient and often contradictory findings about the long-term health-enhancing potential of LPA. For instance, a recent meta-analysis of both experimental and epidemiological studies has found that light-intensity activity is associated with beneficial acute and long-term cardiometabolic responses and lower risk of mortality in adults, but a definitive conclusion could not be drawn due to the moderate consistency between experimental and epidemiological studies (Chastin et al., 2019). Other systematic reviews have also found that LPA might confer cardiometabolic health benefits in adults (Füzéki et al., 2017), also after accounting for MVPA (Amagasa et al., 2018), but similarly with some degree of indefinite conclusions, requiring further research on the potential cardiometabolic health benefits of LPA.

Limited technologies for accurate estimation of LPA in daily life is another possible reason for relatively less available information about how timing and frequency of LPA is related to cardiometabolic health (Lee et al., 2014). Considering that LPA are generally incidental and unplanned compared to MVPA, recalling and self-reporting daily LPA could be relatively less accurate or even infeasible (Lee et al., 2014). In recent years, many studies have started to use wearable activity monitors for monitoring daily activities that could yield relatively more accurate estimates of daily LPA compared to self-reported measures (Amagasa et al., 2018; Chastin et al., 2019). However, even device-based estimation of LPA could be relatively less accurate when compared to the activities with higher intensity. Studies focusing on classifying activities by intensities from wearable activity monitors have shown that estimation of low-intensity activities from such devices could be challenging, requiring further studies to improve the estimation of LPA from wearable devices (Alberto et al., 2017; Carr et al., 2011; Montoye, Westgate, et al., 2018).

2.1.2 Sedentary time and cardiometabolic health

According to the consensus definition (Tremblay et al., 2017), all waking activities that involve minimal movement and low energy expenditure (≤1.5 MET) in a sitting, reclining, or lying posture are sedentary behaviors. Sedentary time, that is defined as the total time spent in sedentary behaviors in any context (e.g., at school or work), includes a major part of adults' activity profile. On the basis of self-reported data, in Europe more than 60%, and globally more than 40% of adults spend four or more hours per day in merely sitting (Hallal et al., 2012). Time spent in sedentary behaviors typically tend to increase with age (Ortega et al., 2013; Spittaels et al., 2012), and device-estimated time-use data indicates that some adults spend in excess of 7–10 hours or more per day in sedentary behavior (Matthews et al., 2008). However, although sedentary behaviors are modifiable (Chastin et al., 2016; Owen et al., 2011), the potential cardiometabolic health risks of this ubiquitous behavior are relatively unclear when compared to other modifiable health behaviors, such as alcohol intake, tobacco use, and lack of physical activity or exercise (Vincent et al., 2017).

Overall, sedentary physiology is a less well-established discipline than physical activity research (Rosenberger et al., 2019; Vincent et al., 2017). This is in large part because sedentary behaviors are more complex than merely a lack of physical activity (Rhodes et al., 2012). The health associations of sedentary behaviors are typically domain-specific and dependent on how sedentary time is accumulated and interrupted (Chastin, Egerton, et al., 2015; de Rezende et al., 2014). Even how adults spend their sedentary time may be important. For example, studies have shown that television viewing time is negatively associated with cardiometabolic health in adults, but this could partially be due to the possible poor dietary choices made while sitting and watching television (Cleland et al., 2008; Thorp et al., 2013). Yet, it is still unclear whether it is the sitting posture, the dietary intake, or the combination of both that is related to deleterious changes in the cardiometabolic health markers (Cleland et al., 2008; Thorp et al., 2013). Additionally, the exact physiological responses by which sedentary time is related to deleterious changes in cardiometabolic health markers in adults is still unclear (Proper et al., 2011; Thorp et al., 2011). Nevertheless, findings of epidemiological studies suggest that total time spent in sedentary behavior, particularly when accumulated in uninterrupted, prolonged bouts (e.g., 60 minutes), may be detrimentally associated with cardiometabolic health and mortality risk in adults (Brocklebank et al., 2015; Rezende et al., 2016; Thorp et al., 2011; Wilmot et al., 2012). Studies have also shown that these detrimental associations may be largely independent of the level of physical activity (de Rezende et al., 2014; Ekelund et al., 2016; Thorp et al., 2011).

Existing evidence is still insufficient for suggesting an appropriate time-based limit for the daily amount of sedentary time required to minimize cardiometabolic health risks in adults. Most guidelines encourage adults to minimize their sedentary time without detailed instructions (Bull et al., 2020; Rosenberger et al., 2019; Vincent et al., 2017). The findings of existing studies are somewhat mixed (Ekelund et al., 2009; Proper et al., 2011; Thorp et al., 2011), especially the findings of those studies with statistical adjustment for adiposity level (mostly body mass index (BMI) and waist circumference) or for some measures of physical activity level (mostly MVPA) (Brocklebank et al., 2015; Thorp et al., 2011). Part of this problem has been attributed to incomplete adjusting for other activity intensities and sleep duration (typically only MVPA but not LPA and/or sleep) among the existing studies (Brocklebank et al., 2015; Powell et al., 2018; Rosenberger et al., 2019).

Another reason for this problem appears to be related to poor estimation of sedentary time with self-reported measures (Thorp et al., 2011). However, inconsistent results have also been observed among the studies with device-based measures that are known to be more precise than self-assessment/questionnaires. One possible explanation for this inconsistency may the restrictive abilities of existing methodologies for differentiating between sedentary behaviors, sleep, and non-wear time from wearable activity monitor signals (Janssen, Basterfield, et al., 2015; Kozey-Keadle et al., 2014; McVeigh et al., 2016), which can potentially result in misclassifications among these components and subsequently overestimation or underestimation of sedentary time. For instance, a previous study of adults found significant associations between total device-estimated sedentary time and adiposity measures (Healy et al., 2011). However, this association was not replicated in another study with the same study sample when combined associations of device-estimated sedentary and MVPA time with obesity risk were examined (Maher et al., 2013). It appears that a high level of sedentary time, especially when combined with low MVPA (Maher et al., 2013), is associated with a higher level of cardiometabolic risk (Powell et al., 2018). However, additional research with device-estimated sedentary time is needed to understand how sedentary time is associated with cardiometabolic health in adults, after proper accounting for other device-estimated activity intensities in waking hours (Chaput et al., 2014; Powell et al., 2018; Vincent et al., 2017).

2.1.3 Interrelationships among sleep, waking activities, and cardiometabolic health

According to expert consensus (Hirshkowitz et al., 2015), the recommended amount of sleep for adults is 7–9 hours per night, with reasonably strong agreement that sleeping regularly outside this recommended range may have adverse effects on health. In line with this, studies relating sleep duration to cardiometabolic health have often observed a U-shaped relationship between sleep duration and cardiometabolic outcomes, with both short and long durations exhibiting adverse associations with cardiometabolic outcomes (Knutson, 2010; Vincent et al., 2017; Xi et al., 2014).

Sleep time and quality appear to be indirectly related to cardiometabolic health through affecting activity patterns and intensities during waking hours (Atkinson et al., 2007; Vincent et al., 2017). For instance, there is already a well-understood relationship between sleep quality and MVPA. Moderate-to-vigorous physical activity is associated with greater ease in falling asleep and deeper sleep (Buman et al., 2015; Rosenberger et al., 2019), which may in turn result in meeting the recommended amount of sleep duration for adults and potentially better cardiometabolic health (Tsunoda et al., 2015). Additionally, insufficient sleep appears to be associated with lower levels of physical activity and higher sedentariness during the day (Bromley et al., 2012; Lakerveld et al., 2016), both of which are related to adults' cardiometabolic health (Brocklebank et al., 2015; Vincent et al., 2017).

Sedentary behaviors and physical activities may also be interrelated. Based on a systematic review of literature, most studies investigating the relationship between sedentary behaviors and physical activities found an association between higher television viewing and lower leisure-time physical activity, but this association was not found between general sitting and leisure-time physical activity (Rhodes et al., 2012). This inverse association appears to be partially because leisure time spent on watching television may potentially displace leisure-time MVPA (Rhodes et al., 2012). Few studies have investigated whether sedentary time is linked to sleep duration independent of time spent in MVPA and/or LPA (Mansoubi et al., 2014). Research has shown that sedentary time might be associated with both shortened and long sleep duration (Basner et al., 2007; Lakerveld et al., 2016; Štefan et al., 2019), but those studies have typically neglected the effects of physical activity levels.

Additional relationships, especially about how LPA affects cardiometabolic health separately from MVPA, sedentary behavior, or sleep, are largely unexplored (Rosenberger et al., 2019). A recent systematic review showed that LPA could be beneficial for cardiometabolic health after adjustments for MVPA (Amagasa et al., 2018), but sleep and sedentary behaviors were not considered in most of the reviewed studies. Combined relationships between LPA and MVPA independent of sedentary behaviors are also largely unknown (Rosenberger et al., 2019).

In general, the associations seen between sleep, sedentary time, and/or physical activities with cardiometabolic diseases and mortality risk are open to the possibility of a reverse causality pathway. Many previous studies investigating how patterns and levels of sleep and movement intensities are related to the markers of cardiometabolic health were cross-sectional (Knutson, 2010; Powell et al., 2018). Due to the cross-sectional study design, inference about the direction of causality between sleep, sedentary time, and/or physical activities with markers of cardiometabolic health remains limited (Henson et al., 2013; Kanagasabai et al., 2017). On the other hand, longitudinal studies have generally taken multiple proper measures to account for reverse causation. This includes adjustment for baseline diseases, excluding adults with diseases or short follow-up in sensitivity analyses, and having moderately long follow-ups (Mezick et al., 2011; Ramakrishnan et al., 2021; Stamatakis, Gale, et al., 2019). Although concerns about reverse causality in previous findings cannot be completely ruled out due to unmeasured and clinically undiagnosed diseases, which can potentially alter the patterns and levels of sleep, sedentary behaviors, and/or physical activities, it seems that the associations of sleep and movement intensities with cardiometabolic diseases tend to remain after accounting for measures of reverse causality (Mezick et al., 2011; Ramakrishnan et al., 2021; Stamatakis, Gale, et al., 2019). Hence, it appears that activity intensities are interrelated and all may be related to cardiometabolic health in adults, and that their implications to cardiometabolic health may be affected by sleep duration at least through modifying the daily time-use (Rosenberger et al., 2019; Vincent et al., 2017). For example, adults with shorter sleep time during a 24-hour period may spend their extra time in sedentary behavior rather than in physical activity (Lakerveld et al., 2016). However, previous studies have not generally considered the potential interrelationships among the activity intensities and sleep time, partly due to limited technologies for precise measurement and assessment of sleep and activities across the entire activity intensity continuum (Chaput et al., 2014; Pedišić, 2014; Rosenberger et al., 2019).

2.2 Assessment of physical activities and sedentary behaviors

The measurement of physical activities and sedentary behaviors is key to understanding their true health implications (Warren et al., 2010). Poor estimation of physical activities and sedentary behaviors can potentially mask or distort the true underlying relationships between activity behaviors and health markers, and can even lead to biased inferences and findings (Celis-Morales et al., 2012). The accurate measurement of physical activities and sedentary time is now even more important and challenging than before. Research has shown that movement intensities represent a continuum and are interrelated, suggesting that all movement and non-movement behaviors may be related to health and wellness in adults (Chaput et al., 2014; Rosenberger et al., 2019). This has created a challenge for accurate estimation activity intensities across the entire continuum in terms of sedentary behaviors, LPA, and MVPA for examining how the activity intensities are related to cardiometabolic health (Rosenberger et al., 2019). In general, there are two approaches for the assessment of sedentary behaviors and physical activities, namely subjective methods and device-based methods with wearable activity monitors (Celis-Morales et al., 2012; Yang et al., 2010).

2.2.1 Subjective methods

Subjective methods are the self-reported levels of physical activities and sedentary behaviors according to participants' own evaluation and perception, often reported through a set of questions but also can be maintained in the form of logs or diaries (Helmerhorst et al., 2012). These methods are generally cost-effective and easy to use in large-scale populations. Different types of questionnaires are currently available in several languages (Helmerhorst et al., 2012; Lee et al., 2011), and have been widely used across different populations and countries and even on a global scale (Guthold et al., 2018). However, subjective methods of measurement of physical activity and sedentary behaviors have several limitations. While subjective methods continue to provide useful evidence, such as on the context of these behaviors and relative indicators of activity intensities (DiPietro et al., 2020; Lee et al., 2014), their validity and inaccuracy remain a serious concern (Helmerhorst et al., 2012; Shephard, 2003). The main problem with the subjective methods is that the self-reporting of the amount and intensity of daily activities is potentially subject to response bias and recall problems (Shephard, 2003). Furthermore, self-reported data does not typically allow researchers to explore timing and precise intensities of movements (Lee et al., 2014).

2.2.2 Device-based methods

Compared to subjective methods, device-based methods do not have response bias and recall problems, and thus are likely to provide better estimates of actual physical activities and sedentary behaviors (Celis-Morales et al., 2012). Studies comparing measures of physical activity and sedentary behaviors in the same adult participants have often reported spurious differences for subjective measures in contrast to device-based measures that are more likely to be accurate (Dyrstad et al., 2014; Lagersted-Olsen et al., 2014). Device-based techniques typically include wearable motion and physiological sensors, such as pedometers, heart rate recorders, actometers, goniometers, gyroscopes, accelerometers, or a combination of these, for the measurement of activity behaviors (Yang et al., 2010). These devices generally work by sensing either physiological or mechanical responses to bodily movement, and use these signals to estimate variables that reflect physical activities and sedentary behaviors (Bassett Jr et al., 2012). In addition to these devices, the accurate measurement of physical activities and sedentary behaviors could be done by other means such as magnetic systems, optical systems, and video recording, yet such methods have rarely been used on a large scale due to complexity of setup and privacy issues in free-living settings (Yang et al., 2010).

2.2.3 Accelerometry

Accelerometers are electromechanical sensing devices that mechanically measure any moving object's acceleration (rate of change in time), whether caused by gravity or motion, and convert this data into an electrical signal (Kavanagh et al., 2008; Mathie et al., 2004). The operation principle for most accelerometers is based on Hooke's and Newton's laws (Kavanagh et al., 2008).

According to Hooke's law of extension and compression, the displacement of the spring is linearly proportional to the force. Mathematically, Hooke's law is stated as:

$$F = k * x \tag{1}$$

in which F is the force exerted on the spring, x is the displacement (distance) of the spring's end from its equilibrium position, and k is the spring constant. Newton's second law states that the force (F) created by a moving object is equal to its mass (m) times acceleration (a), giving the equation:

$$F = m * a \tag{2}$$

Combining these two equations, the acceleration (a) caused by the body mass (m) on a spring (whose elongation property is characterized by k) can be known by quantifying the distance (x), giving the equation:

$$F = k * x = m * a, and\ thus\ a = kx/m \tag{3}$$

In practice, there are several techniques and technology for measuring acceleration. Piezoelectric, piezoresistive, and capacitive sensors are the most common sensors found within commercial accelerometers, and they all operate by converting mechanical energy into an electrical signal but with different technologies, as detailed elsewhere (Yang et al., 2010). Currently, most accelerometers are available in small-to-miniature sizes at low cost. They are also capable of measuring high-frequency, triaxial acceleration in the three orthogonal axes (vertical, medio-lateral, and anterior-posterior in body axis), thanks to micro electro-mechanical system

(MEMS) that enable mechanical sensing structures of microscopic size (Yang et al., 2010). These features, combined with other advances in sensor technologies that enable storing high frequency accelerometry signals over longer time periods, have made accelerometers appealing for many applications requiring the objective measurement of bodily movement, including human activity sensing devices (Kavanagh et al., 2008; Mathie et al., 2004).

2.3 Device-estimated physical activities and sedentary behaviors using accelerometers

In recent years, studies in the field of physical activity and sedentary behavior research have increasingly relied on body-worn accelerometer-based activity monitors (Lee et al., 2014; Wijndaele et al., 2015). It is widely acknowledged that much of the recent advances about how physical activity and sedentary time are related to health is in large part due to the availability of accelerometer-based activity monitors and their associative advantages (Troiano et al., 2014; Trost et al., 2005). For instance, continuous, high-resolution accelerometer signals have allowed for studying the patterns of physical activity and sedentary time in terms of movement behavior profiles, and examining how these profiles are associated with cardiometabolic and mortality risk (Lee et al., 2013; Niemelä et al., 2019). In particular, the popularity of accelerometer-based activity monitors in epidemiological and population-based studies is mainly because they are lightweight, reliable, feasible, and inexpensive sensors that can record acceleration *continuously* over extended periods of time (Lee et al., 2014). Additionally, accelerometers are non-invasive means of measurement of human movement that do not alter or interfere with natural body movement (Yang et al., 2010).

2.3.1 Accelerometer placement

Accelerometer placement is one of the main sources of inconsistencies among existing accelerometry studies (Rosenberger et al., 2016; Trost et al., 2005). Although studies have shown that the assessment of physical activities and sedentary behaviors could be more accurate with multiple accelerometers combining acceleration signals acquired from different body locations (Cleland et al., 2013; Ellis et al., 2014; Trost et al., 2017), most large-scale, population-based studies have continued to use a single wear location (Bassett Jr et al., 2012). This is mainly because carrying multiple sensors in different body locations could be cumbersome for participants, and even infeasible for some. Multi-sensor devices that include multiple sensors such as magnetometer, gyroscopes, heart-rate, and/or accelerometers in one device appear to also have increased validity and accuracy for estimation of sedentary behaviors and physical activities compared to a single accelerometer (Nweke et al., 2019). With the current trend, it appears that the technical and computing improvements continue to increase the validity and acceptability of multi-accelerometer systems and multi-sensor devices (Nweke et al., 2019).

The hip, wrist, thigh, ankle, arm, knee, chest, and even ear have been all proposed and studied as an accelerometer attachment site (Atallah et al., 2011; de Almeida Mendes et al., 2018). Although there is still no consensus on a single accelerometer placement within the literature, it appears that large-scale studies have continued to select the hip and wrist (Wijndaele et al., 2015), and most recently also the thigh (Hamer et al., 2020) as the placement of choice.

Wrist-worn accelerometers have relatively higher compliance and feasibility (Troiano et al., 2014), but relating a wrist-based acceleration signal to activity types and body posture could be difficult (Atallah et al., 2011; Migueles et al., 2017). This may be because the wrist could be in various orientations (e.g., upward and downward) during similar activities such as sitting or standing. Besides, in many activities such as cycling, the wrist is generally at rest and does not reflect the whole-body acceleration (Atallah et al., 2011). Performing such activities at higher intensities may not necessarily result in a higher or different wrist-based acceleration signal, and therefore the estimation of types and intensity of those activities may be less accurate from wrist-worn accelerometers (Cleland et al., 2013; Ellis et al., 2014; Montoye et al., 2016b).

The thigh seems to be the most common wear location for body posture detection, specifically when accurate differentiation between sitting/lying, standing, and ambulatory (steps) activities is of interest (Berendsen et al., 2014). Practically, sitting versus an upright posture can be discriminated using static acceleration and determining the orientation of the thigh, and further from ambulatory activity (steps) by measuring dynamic acceleration (Edwardson et al., 2017). Due to the high accuracy of thigh-worn accelerometry and criterion measures for capturing ground truth data under free-living conditions, thigh-worn accelerometry has often been used as a ground truth in free-living settings, against which other accelerometers worn in other body locations have been validated for posture detection (Rosenberger et al., 2016; Vähä-Ypyä et al., 2018). However, current measurement protocols require fixed positioning of the sensor on the thigh (typically with breathable bandages) by research assistants, and typically participants are not allowed to remove the sensor during the monitoring period, making the thigh a relatively less convenient wear location (Berendsen et al., 2014).

The hip has continued to be the mainstay of activity monitoring (Matthews et al., 2012; Strath et al., 2012). It is close to the center of mass, and the whole-body movement is generally reflected in the acceleration signal collected from hip (Yang et al., 2010). Studies have shown that the hip could be an appropriate single location for an accelerometer for measuring the intensity of physical activities (Cleland et al., 2013). It also appears that the hip is a relatively acceptable wear location for participants (Berendsen et al., 2014). Hence, hip-worn accelerometer may be less accurate for identifying stationary and sedentary behaviors (Hart et al., 2011; Kozey-Keadle et al., 2011), requiring more advanced approaches for the estimation of all movement intensities from hip-based acceleration signals.

2.3.2 Activity counts and traditional statistical approaches

Traditionally, accelerometers have provided proprietary integrated signals known as “activity counts.” Previous studies have therefore generally used regression analysis to relate activity counts with activity energy expenditure, and subsequently established a set of cut points for converting activity counts into activity intensities (Bassett Jr et al., 2012; Crouter et al., 2013). Receiver Operating Characteristics (ROC) curve analysis is another approach that has been used for establishing activity intensity cut points (Kim et al., 2012). Using these relatively simple methods (referred to as traditional statistical methods in the literature), several regression-based equations and sets of thresholds (cut-points) have been established for estimating activity energy expenditure and activity intensities (Bassett Jr et al., 2012; Kim et al., 2012; Migueles et al., 2019). These regression-based equations and cut-points have been widely used in the existing literature for the assessment of physical activity intensities, sedentary time, and activity energy expenditure (Migueles et al., 2017), mainly due to their simplicity of implementation (Troiano et al., 2014). However, several studies have emphasized that these methods are not accurate in the presence of a wide range of activity types under free-living conditions, followed by requests to apply more sophisticated data modelling approaches for developing robust models capable of accurate assessment of all activity intensities (Bassett Jr et al., 2012; Freedson et al., 2012; Kim et al., 2012).

2.3.3 Raw accelerometry and machine learning approaches

Machine learning approaches are alternative analytic techniques for calibration and validation of accelerometer-based activity monitors (Bassett Jr et al., 2012; de Almeida Mendes et al., 2018). Machine learning approaches are advanced statistical techniques with the ability to capture complicated relationships and nonlinearities in data, making them excellent candidates for the calibration and validation of accelerometer-based activity monitors producing either activity counts or raw acceleration data (Bassett Jr et al., 2012; van Hees et al., 2016). A unique advantage of ML approaches over traditional statistical method is the possibility of including dozens or even hundreds of input features describing characteristics of acceleration signal in both time and frequency domains (Bassett Jr et al., 2012; Freedson et al., 2012; Wang et al., 2019). This has essentially allowed to create more sophisticated prediction models based on multiple acceleration signal characteristics for predicting activity types, categories, intensities, or energy expenditure, as opposed to regression-based approaches which typically use only mean acceleration values and linearly relate them to energy expenditure (Bassett Jr et al., 2012; Freedson et al., 2012).

To date, various ML approaches have been explored for prediction of activity types, intensities, categories, and energy expenditure (de Almeida Mendes et al., 2018; Wang et al., 2019), among them the majority of techniques have used random forests, support vector machines, artificial neural networks (ANN), decision trees, and different boosting approaches (de Almeida Mendes et al., 2018). Those previous studies have typically shown that ML approaches have increased accuracy when compared to cut-points or regression equations in predicting activity intensities and energy expenditure (Ellis et al., 2016; Montoye et al., 2015). However, although better in prediction accuracy (Ellis et al., 2016), the development and use of ML models could be relatively more complicated than other statistical approaches (Bassett Jr et al., 2012). There is still no consensus in the literature on which ML technique is best, mainly because their performance can depend on several factors such as wear location, population age range, window size, and even defined activity categories that can vary from one study to another (Bassett Jr et al., 2012; de Almeida Mendes et al., 2018; Mannini et al., 2017; Zhang, Murray, et al., 2012). Such parameters and choices may greatly influence the predictive ability of ML-based models, but their effects on the predictive ability of ML-based models remain unclear. This is currently obscuring better understanding of the potential advantages of ML approaches for the calibration and validation of accelerometer-based activity monitors (Bassett Jr et al., 2012; van Hees et al., 2016).

More recently, with the popularity of accelerometers capable of measuring and storing raw acceleration data (van Hees et al., 2016), it has been observed that there is a shift from traditional statistical approaches towards ML-based modeling for the calibration and validation of accelerometer-based activity monitors (de Almeida Mendes et al., 2018). Raw accelerometry, unlike activity counts, are unintegrated signals, offering an unprecedented opportunity for providing comparable accelerometry results (van Hees et al., 2016), and accordingly enabling interstudy comparison (Migueles et al., 2019; Wijndaele et al., 2015). Studies focusing on the comparability of accelerometry data have found that raw acceleration signals from various accelerometers are not equivalent but could be highly comparable often with some further calibration strategies (Bassett Jr et al., 2012; John et al., 2013; Montoye, Nelson, et al., 2018). It is therefore likely that increased output comparability provided by raw accelerometry together with the predictive ability of ML approaches may help to develop more reliable data processing techniques capable of predicting physical activities and sedentary behaviors in various population groups, independent of accelerometer type and brand beyond existing regression-based equations and cut points (Matthews et al., 2012; van Hees et al., 2016). Such techniques may eventually enable the comparability of results across studies and provide opportunities to pool data from different studies (Wijndaele et al., 2015).

2.3.4 Machine learning for activity intensity estimation

In practice, accelerometer-produced signals/outputs can be used for recognizing activity types and intensities, estimating activity energy expenditure, and detecting body postures (de Almeida Mendes et al., 2018; Yang et al., 2010). Still, categorizing activities by intensity (i.e., sedentary behaviors, light, moderate, and vigorous) seems to be the most common measure among the existing accelerometry-based studies (Wijndaele et al., 2015). Hence, a universally accepted method for predicting activity intensity from acceleration data across the entire intensity spectrum is still lacking (Migueles et al., 2021). This may be partly because of the limited generalization capability of existing data processing methods, including ML-based methods, when applied to populations different from the one used for model development (Bassett Jr et al., 2012; de Almeida Mendes et al., 2018).

Few studies to date have investigated the generalization performance of ML-based models for activity type and intensity prediction (de Almeida Mendes et al., 2018). Although those studies have consistently exhibited performance deterioration when ML models were cross-tested on independent populations, there is still no consensus on its cause. For example, while a previous study has identified that the differences between acceleration data collected in free-living and laboratory settings is the main reason for the performance degradation of laboratory-calibrated ML models for activity prediction (Bastian et al., 2015), another study has reported that the differences in sample characteristics were the main reasons for accuracy degradation (Mannini et al., 2017). These discrepancies among previous studies seem to be due to focusing on only one factor at a time, whereas there would be several contributing factors in real-world applications, ranging from acceleration data and sample characteristics to unseen activities, which could all affect the generalization capability of ML models (Bastian et al., 2015; John et al., 2013; Mannini et al., 2017). This has created a need for more realistic tests to better understand the generalization capability of ML models.

Recent evidence has suggested developing ML-based classification models for classifying activities directly according to their intensity categories. The validity of this method (Montoye et al., 2016b; Tjurin et al., 2017) as well as its generalizability on independent populations have been presented in previous studies (Montoye, Westgate, et al., 2018). Although better than simple regression equations, ML models developed for energy expenditure estimation models may also exhibit high error and bias (Montoye, Westgate, et al., 2018; Staudenmayer et al., 2015). Additionally, research has shown that the performance of classification models deteriorates as the number of activity categories increases (Ellis et al., 2014; Montoye et al., 2016a). Developing a direct method for predicting activity intensities may therefore be preferred over indirect methods, which first predict the energy expenditure of activities and then classify activity intensity based on energy expenditure thresholds, or first predict activity types and then collapse the categories into intensity categories (Montoye, Westgate, et al., 2018). However, further research is required to understand how well ML models developed for activity intensity classification would perform when applied to other populations.

2.3.5 Limitations and challenges

Despite numerous advantages, accelerometers have also some inherent limitations. Accelerometers have limited abilities in the assessment of body posture and resistance training exercises, and they cannot distinguish whether a person is carrying any weight (e.g., walking carrying a heavy bag requires more energy compared to walking with no load) (Lee et al., 2014). Traditionally, physical activity outcomes in studies calibrating and validating accelerometer-based activity monitors have been expressed as absolute intensity measures (e.g., 5 MET, 30 minutes in 3- to 6-MET range) with no consideration for the functional capacity of the individual (Freedson et al., 2012; Kujala et al., 2017). The relative intensity of physical activity with consideration of individual fitness and functional status may indeed be a better measure than absolute activity intensities for providing tailored physical activity recommendations (Kujala et al., 2017). This is because maximal exercise capacity in low-fit individuals, in particular among those who are obese or have chronic diseases, may be lower than the recommended absolute intensity level of MVPA, while fit individuals generally reach this intensity of movement relatively easier (Kujala et al., 2017). Consequently, individuals who cannot reach the recommended intensity level may not have this intensity of physical activity recorded based on absolute intensity values, and generally self-reported physical activities could be better means to determine the relative indicators of activity intensity (Freedson et al., 2012). Hence, how to incorporate measures of an individual's fitness level into the accelerometry-estimated physical activity metrics has remained an issue (Freedson et al., 2012).

In general, accelerometers could provide good estimates of the intensity of certain types of physical activity depending on the wear location (de Almeida Mendes et al., 2018). However, accelerometers generally have relatively lower accuracy for estimating activity intensities that fall at the lower end of the intensity spectrum (i.e., sedentary to light activities) (Alberto et al., 2017; Carr et al., 2011). In particular, measurement of different components of sedentary behaviors from accelerometry signal may also be challenging and related to the wear location (Alberto et al., 2017; Janssen & Cliff, 2015; Kozey-Keadle et al., 2011). For instance, the fact that wrist orientation is not always aligned with whole body makes it difficult to assess different types of sedentary behaviors from wrist data (Staudenmayer et al., 2015).

Although advantageous, the proliferation of accelerometers has also led to multiple analytical and practical challenges. On the surface, all body-fixed accelerometer-based activity monitors seem to perform with the same basic principle—monitoring total body acceleration. However, there are significant differences in sensor properties, measurement protocols, and data processing (Bassett Jr et al., 2012; Freedson et al., 2012; Migueles et al., 2017, 2019; Welk et al., 2012). These differences have made it difficult to directly compare outputs from different accelerometers (Rosenberger et al., 2016; Rowlands et al., 2017), and in turn, comparison of the results of accelerometer-based studies across different populations has remained limited (Wijndaele et al., 2015). In particular, inconsistencies among existing studies could, to a large extent, be related to different data processing methods used for interpreting accelerometry data (Freedson et al., 2012; Migueles et al., 2017, 2019).

2.4 Combined effects of physical activities, sedentary behaviors and sleep on cardiometabolic health

Currently, time-based recommendations for adults are only available for sleep duration (7–9 hours per night) and MVPA (150 minutes per week) for which there is reasonably strong evidence, whereas only general advice to minimize sedentary behavior and perform more LPA have been made (Bull et al., 2020; Knutson, 2010; Rosenberger et al., 2019). It remains unclear how time over a full 24-hour cycle should be distributed between sleep, sedentary time, and physical activities for optimal cardiometabolic health in adulthood (Vincent et al., 2017). This may partially be because of the fact that nearly all published studies to date have examined the relationship between specific health indicators of interest and time spent on only one activity during a daily 24-hour cycle (Chaput et al., 2014; Pedišić, 2014; Rosenberger et al., 2019), including when cardiometabolic health markers have been studied (Amagasa et al., 2018; Brocklebank et al., 2015; Füzéki et al., 2017; St-Onge et al., 2016).

Most recently, research has shown that time spent in each activity may modify the health-related influence of time spent in any of the other activities (Rosenberger et al., 2019). For example, increasing time spent in MVPA may significantly reduce the negative effects of sedentary time on cardiometabolic health (Chastin, Palarea-Albaladejo, et al., 2015; McGregor et al., 2018) and mortality risk (Ekelund et al., 2016). Sleep time and accelerometer-estimated sedentary time and physical activities may indeed be codependently associated with adults' cardiometabolic health (Chaput et al., 2014; Vincent et al., 2017). However, previous studies examining the cardiometabolic health implications of sleep and movement intensities have generally not considered the potential interrelationships among the activities over a 24-hour sleep and wake cycle (Brocklebank et al., 2015; Pedišić, 2014; Vincent et al., 2017).

2.4.1 Time use in the 24-hour cycle

Historically, movement behaviors have been assumed to be independently associated with health outcomes (Chaput et al., 2014; Chastin, Palarea-Albaladejo, et al., 2015). Most studies in adults have therefore typically used traditional regression methods and separately dealt with sleep, sedentary behavior, LPA, and MVPA when examining their cardiometabolic health implications (Pedišić, 2014). Few of those studies have considered accounting for other activities, but only partially accounted for one or two activities in the 24-hour cycle (Pedišić, 2014). Incomplete or improper consideration of other movement behaviors in the 24-hour cycle may bias findings (Dumuid, Pedišić, et al., 2019; Dumuid et al., 2018; Pedišić, 2014; Pedišić et al., 2017), and thus it is important to adjust for the full range of 24-hour sleep, sedentary time, and physical activities using suitable analytical approaches (Chaput et al., 2014; Rosenberger et al., 2019).

An appropriate balance between sleep, sedentary time, and physical activities may be needed for optimal cardiometabolic health (Chaput et al., 2014; Pedišić, 2014; Vincent et al., 2017). To investigate how the balance among these time-use components is associated with different health markers, the isotemporal substitution procedure was initially introduced (Mekary et al., 2009; Pedišić et al., 2017). Isotemporal substitution allows to explore how reallocation of time between two movement behaviors (e.g., sedentary activities to MVPA) would contribute to the theoretical changes in the health outcomes of interest (Grgic et al., 2018). More recently, research has shown that sleep and movement intensities can indeed be considered as compositional data, since they are mutually exclusive time-use components of a fixed period, such as the 24-hour day (Dumuid, Pedišić, et al., 2019; Dumuid et al., 2018). A change of time spent in one activity intensity necessitates an exchange of equal time for one or a combination of other activity intensities and/or sleep. Accordingly, there has been a recent conceptual shift in behavioral epidemiology which moves away from exploring movement behaviors as independent exposure, towards an approach which allows the influence of all movement intensities and sleep to be considered relative to each other, called time-use epidemiology approach (Pedišić et al., 2017; Rosenberger et al., 2019). This shift has been facilitated by the employment of new analytical approaches based on compositional data analysis (Dumuid, Pedišić, et al., 2019; Dumuid et al., 2018).

Compositional data analysis methods are novel statistical approaches that are able to accommodate codependent data that is constrained to a fixed amount of time, and are therefore well-suited to analyzing time budgets in a 24-hour cycle (Dumuid, Pedišić, et al., 2019; Dumuid et al., 2018). Researchers have recently started to use compositional data analysis to investigate associations of 24-hour time use with cardiometabolic health markers in adulthood (Chastin, Palarea-Albaladejo, et al., 2015; McGregor et al., 2018). These studies have found that MVPA was beneficially associated with markers of cardiometabolic health, but results for the other movement behaviors were inconsistent (Chastin, Palarea-Albaladejo, et al., 2015; McGregor et al., 2018). Additional research is needed to understand how compositions of 24-hour time-use is associated with cardiometabolic health in adults.

2.4.2 Patterns of accumulation

With the development of accelerometry techniques, researchers have also started to investigate whether and how the accumulation patterns of accelerometer-estimated physical activity across the day contribute to cardiometabolic health, beyond the total daily level of physical activity (Lee et al., 2013; Niemelä et al., 2019). For instance, although studies have often shown that accumulating the recommended amount of daily MVPA in sustained 10-minute bouts may provide additional cardiometabolic and mortality benefits, recent findings based on accelerometer data indicate that health benefits may be achieved with higher physical activity volumes from LPA upwards, regardless of accumulation patterns (Chastin et al., 2019; Glazer et al., 2013; Warburton et al., 2017).

In recent years, there have also been significant advances in understanding how the patterns of accumulation of sedentary behaviors are related to health markers. Observational studies with accelerometer-estimated sedentary time have found that accumulating sedentary time in prolonged, uninterrupted bouts is associated with higher cardiometabolic risk in adults (Carson et al., 2014; Cooper et al., 2012). Additionally, both experimental and observational studies have shown that breaking up sedentary time with short and sustained bouts of LPA and MVPA could modify those detrimental effects caused by sedentary time on cardiometabolic markers in adults (Chastin, Egerton, et al., 2015; Dunstan et al., 2012; Healy et al., 2008). However, it is still unclear when sedentary time should be interrupted before it becomes detrimental to health, and even less is known about the length and intensity of interruptions required to minimize the detrimental health effects of uninterrupted sedentary bouts (Carson et al., 2014; Chastin, Egerton, et al., 2015; Healy et al., 2011; Janssen & Cliff, 2015).

Like physical activity research, sedentary research has also generally suffered from dealing with sedentary time in isolation, independent of physical activities. Few studies have used time-use approaches for examining the cardiometabolic health associations of accumulation patterns of sedentary time after accounting for physical activities and sleep (Pedišić, 2014; Rosenberger et al., 2019). However, a main limitation of time-use approaches is that the interpretation of the results is not straightforward and becomes more complicated when accommodating a higher number of variables (Dumuid, Pedišić, et al., 2019; Dumuid et al., 2018). Those studies have therefore accommodated a limited number of variables, typically only total time spent on sedentary and physical activities (Chastin, Palarea-Albaladejo, et al., 2015; McGregor et al., 2018).

Data-driven, person-centered statistical approaches are also suitable methods for examining how patterns of accumulation of sedentary time and physical activities are related to cardiometabolic health (Migueles et al., 2021). A notable advantage of these approaches compared to other commonly-used variable-centered approaches is that a higher number of variables can be accommodated in the analyses for forming the groups (Lee et al., 2013; Niemelä et al., 2019; Verswijveren et al., 2020; Xu et al., 2005), offering the possibility for understanding how combined accumulation patterns of sedentary and activity behaviors is related to health (Migueles et al., 2021). An increasing number of studies have therefore used statistical approaches, such as latent profile analysis and ML-based clustering methods (e.g., K-means), to identify groups of individuals who share similar patterns of activity behaviors and to investigate how these distinct activity patterns are related to cardiometabolic health (Gupta et al., 2020; Niemelä et al., 2019; Verswijveren et al., 2020) and mortality risk (del Pozo Cruz et al., 2020; von Rosen et al., 2020b). However, few of these studies have been performed on adults (Gupta et al., 2020; Niemelä et al., 2019), and none of them have included variables characterizing how sedentary time was accumulated and interrupted. Further studies with data-driven approaches and variables describing how sedentary time was accumulated and interrupted may help to better understand how patterns of accumulation of sedentary time and sedentary breaks are associated with cardiometabolic health in adults.

2.5 Correlates of physical activity and sedentary behaviors in adults

Despite numerous health benefits, few adults meet the current recommendations for physical activity (Guthold et al., 2018). Correlates of activity behavior (i.e., factors associated with different activity behaviors) have therefore been well studied to understand the causes of physical inactivity and sedentary behaviors, and to inform activity promotion strategies (Bauman et al., 2012; Trost et al., 2002). To date, numerous studies have found an association between various factors of different domains such as personal, societal, and environmental factors with different indices of activity behaviors such as the daily amount of MVPA or sedentariness (Bauman et al., 2012; Choi et al., 2017; O'Donoghue et al., 2016). With advances in sensor technologies, research into correlates has shown that activity behavior could be a multidimensional behavior that is explained by a multilevel, complicated web of factors (Bauman et al., 2012; Chastin et al., 2016; Kohl 3rd et al., 2012). Accordingly, there have been calls for more research using both sophisticated statistical assessment that can capture the multilevel nature of correlates and different definitions of activity behavior that better reflect everyday life rather than unidimensional metrics (Bauman et al., 2012; Pate et al., 2018; Silva et al., 2017; Warburton et al., 2017).

Most studies to date have typically used classical statistical modeling (such as regression analyses) to examine whether and how various factors are associated with different activity metrics (Choi et al., 2017; Trost et al., 2002). In classical statistics, the analyses could remain restricted to data analysts' decisions about how the association and interaction are hypothesized (knowledge-driven), given that the factors selected for inclusion in the analyses are typically chosen subjectively according to their conceptual relevance and, in some cases, initial empirical associations (Trost et al., 2002; Venkatasubramaniam et al., 2017). This seems to be one thing that is limiting the recognition of new and innovative correlate categories, which are needed in this field for further progress (Bauman et al., 2012; Trost et al., 2002). Ecological approaches that integrate ideas from several theories are other approaches that have been used for correlates research, often to overcome the limitations of classical statistical analyses (Bauman et al., 2012). Those approaches have been used to both conceptualize the factors and their interrelationships at all levels explaining activity behavior (such as the interconnections between individuals and their social and physical environments) (Sallis et al., 2006) and guide variable selection for analyses (Bauman et al., 2012;

Trost et al., 2002). However, ecological approaches are also knowledge-driven (Choi et al., 2017) and, to some extent, rely on very well-established correlates (Chastin et al., 2016; Choi et al., 2017), which might result in missing some important factors explaining activity behavior.

We have now entered a data-intensive era, with increasing popularity for data mining approaches (Chen et al., 2014). Data mining approaches originated from statistics but are known to capture hidden and novel insights buried in large amounts of data (Bellazzi et al., 2011, 2008). They are also excellent methodologies for generating data-driven hypotheses (Bellazzi et al., 2011, 2008). These principles also apply to the field of physical activity research, in which there is a need for more complex approaches to identify the next generation of correlates of activity behaviors, understand their relative importance, and capture the complex interrelations among the factors at different levels (Bauman et al., 2012; Chastin et al., 2016; Choi et al., 2017). Studies have therefore started to use data mining approaches for creating data-driven hierarchies of correlates predicting activity behaviors (Buck et al., 2019; Lakerveld et al., 2017; Patterson et al., 2018; Yoon et al., 2015). However, those studies have generally used a limited number of already well-known correlates consisting mainly of environmental and individual factors, and defined physical activity or sedentariness according to a single unidimensional metric derived from self-reported data (Buck et al., 2019; Lakerveld et al., 2017; Patterson et al., 2018). Data mining approaches applied to a broadened list of factors and a more representative definition of activity behavior in everyday life using accelerometry data may facilitate our understanding of why some people are active but others are not.

3 Aims of the study

The present study reviewed studies calibrating and validating accelerometer-based activity monitors using machine learning approaches and developed robust models for classifying activities by intensity. Additionally, it examined how the compositions and patterns of accumulation of sedentary time and physical activities are associated with markers of cardiometabolic health and analyzed correlates of activity behaviors in a large population-based sample of middle-aged Finnish adults.

More specifically, the aims of the study were:

- To reveal the generalization and predictive ability of machine learning approaches for the calibration and validation of wearable accelerometers.
- To evaluate and enhance the generalization performance of machine learning models developed for classifying activity intensities across the entire intensity continuum (sedentary behaviors, LPA, and MVPA), particularly to examine whether activity intensity classification models developed with raw acceleration data and validated using within-sample validation are generalizable to other populations monitored with different accelerometers.
- To examine how compositions of 24-hour time-use, and time reallocations between sleep duration and accelerometer-measured sedentary time and physical activities, are associated with markers of cardiometabolic health in adults.
- To identify profiles according to distinct accumulation patterns of accelerometer-measured sedentary time and sedentary breaks in adults, and to investigate how these profiles are associated with markers of cardiometabolic health.
- To establish a multilevel, data-driven hierarchy for predicting activity behaviors, and to methodologically identify the correlates of activity behaviors in adults.

4 Materials and methods

This thesis is composed of five sub-studies, which were performed with three different data sources. The sub-studies are referred to in the text using Roman numerals I–V. Table 1 summarizes the study setup, data sources, and study design by sub-study.

Table 1. Materials, methods, and study design by sub-study.

Study	Study setup	Data source	Study design
I	Review of literature to reveal the potentials of ML approaches for calibration and validation of wearable accelerometers	62 original peer-reviewed research articles	Systematic review
II	Cross-dataset study to evaluate and enhance the generalization capability of ML models developed for activity intensity classification	Four independent studies (three open access datasets and one in-lab dataset)	Cross-dataset tests
III	Compositional data analysis to examine how compositions of 24-hour time-use is associated with cardiometabolic health markers	NFBC1966, n = 3,443	Cross-sectional
IV	Cluster analysis to investigate how patterns of accumulation sedentary time and sedentary breaks are associated with cardiometabolic health markers	NFBC1966, n = 4,439	Cross-sectional
V	Data mining using a decision tree to identify the correlates of activity behaviors	NFBC1966, n = 4,582	Cross-sectional

ML = machine learning, NFBC1966 = Northern Finland Birth Cohort 1966.

4.1 Systematic review (I)

Material (articles) for the sub-study I were retrieved from the PubMed and Scopus databases. These two databases were initially searched on July 1, 2017 for studies calibrating and validating accelerometer-based activity monitors using ML approaches. The two search strings are provided in Appendix 1. Additional articles were identified by searching the references in papers identified by the search. Data items such as the predictive accuracies and other important parameters in training ML algorithms (e.g., prediction method, windowing approach, data type, etc.) were extracted from the eligible studies, and risk of bias was assessed with a summary quality score (range: 0–1 with a higher score indicates better quality) (Kmet et al., 2004). The criteria used for scoring each study were based on earlier recommendations made for studies calibrating and validating accelerometers (Bassett Jr et al., 2012; Liu et al., 2012; Welk, 2005).

4.1.1 Eligibility criteria and study selection

We included original peer-reviewed journal articles calibrating and validating ML models based on data collected from a single body-fixed accelerometer (not smartphone-based) in order to predict the type, category, intensity, and/or energy expenditure of activities. The monitored activities had to be health-related and daily activities such as walking, cycling, sedentary activities, etc. In case of multiple accelerometers in various body locations, the study was included if calibration and validation of ML models were based on data acquired from each attachment site separately. Studies validating a previously developed ML-based model were also included. Here, we only present extracted data related to the predictive accuracy of the activity recognition models. We use the term "activity recognition" to refer to all classification models developed for classifying activities into types, classes, or intensities (i.e., using classification techniques not from MET-estimated values).

Briefly, the initial literature search produced 3,171 articles (sub-study I, Fig. 1). Additional, 13 articles were manually identified from references in the papers. Eventually, 104 articles were read in full text and checked for eligibility using a predefined form including the eligibility criteria items, of which 62 articles were finally considered eligible for inclusion in this review. Further details about how these studies were screened and selected can be found in sub-study I.

4.2 Labelled training data (II)

Data for sub-study II were from four independent studies. In the systematic review (I), we found three studies (Oregon State University (Trost et al., 2014), PAMAP2 Physical Activity Monitoring (Reiss et al., 2012), and Daily and Sports Activities (Altun et al., 2010)) that made their labeled datasets publicly available to the research community at the time of the study (II), and all comprised raw acceleration data measured by wearable activity monitors. Additionally, we had similar data from another study that was performed by our research team at the University of Oulu (Tjurin et al., 2017). A brief description of each study is provided in Table 2.

Table 2. Detailed characteristics of the four independent studies used for conducting cross-dataset tests in sub-study II.

Dataset	Collection environment	Participants	Sensors and accelerometer sensor specifics	Wear locations
University of Oulu (UOULU)	A dataset collected inside and outside a laboratory.	22 participants (11 male, 11 female), mean age: 27.5 (11.2), age range: 17–58, BMI: 25.1 (2.2) $kg.m^{-2}$	Hookie AM20 triaxial accelerometer, scale: ±16g, sampling rate: 100 Hz	Right hip
Oregon State University (OSU)	An open-access dataset collected inside a laboratory.	52 participants (28 boys, 24 girls), mean age: 13.7 (3.1), age range: 7.2–18.9, body mass: 50.6 (13.5) kg, handedness: not specified	ActiGraph GT3X+ triaxial accelerometer; scale: ±6g, sampling rate: 30 Hz	Right hip and non-dominant wrist
PAMAP2 Physical Activity Monitoring (PAMAP)	An open-access dataset collected inside and outside a laboratory.	9 participants (8 male, 1 female), mean age: 27.2 (3.3), age range: 23–32, BMI: 25.1 (2.6) $kg.m^{-2}$, handedness: 7 right, 1 left, 1 not reported	Colibri wireless inertial measurement unit containing two 3D accelerometers, 3D gyroscope sensor, 3D magnetometer sensor, temperature, orientation and heart rate monitor sensors; scales: ±16g and ±6g, sampling rate: 100 Hz	Dominant wrist, dominant ankle, chest
Daily and Sports Activities (DSA)	An open-access dataset collected inside and outside a laboratory.	9 participants (4 male, 4 female, 1 not reported), mean age: not specified, age range: 20–30, BMI: not specified, handedness: not specified	Xsens MTx wired inertial measurement unit containing 3D accelerometer, 3D gyroscope, 3D magnetometer; scales: ±5g (wrists) and ±18g (knees and chest), sampling rate: 25 Hz	Right and left wrists, right and left knees, chest

3D = three dimensional, BMI = body mass index.

4.2.1 Dataset preparation

Raw acceleration data from the four independent studies were extracted and used to create five datasets, each containing only hip or wrist triaxial raw acceleration data (Fig. 1). Throughout the text, the five datasets are referred to as UOULU (H), OSU (H), OSU (W), PAMAP (W), and DSA (W), where H refers to hip and W to wrist. The dataset PAMAP (W) includes wrist acceleration data from the accelerometer sensor with a scale of ±16 g. The dataset DSA (W) includes the right wrist acceleration data.

In all the datasets, direct observation served as the criterion for physical activity. The Compendium of Physical Activity for adults (Ainsworth et al., 2011) and youths (Butte et al., 2018) were used to assess the energy expenditure associated with each activity in the adult (UOULU, PAMAP, DSA) and youth (OSU) datasets, respectively. Based on the recommended anatomical postures and absolute MET thresholds (Tremblay et al., 2017), the performed activities within each dataset were categorized into three intensity categories: ≤1.5 MET: sedentary behavior, 1.5–3.0 MET: LPA, and ≥3.0 MET: MVPA (sub-study II, Table 2). We created two types of dataset; datasets with original down-sampled raw acceleration data and datasets with tailored data. The down-sampling from 100 Hz to 25 Hz and 30 Hz to 25 Hz was performed by taking one data point out of every four data points and five data points out of every six data points, respectively.

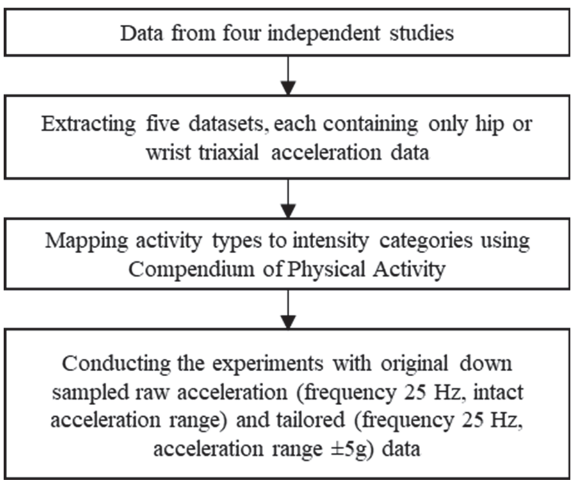

Fig. 1. The general workflow of data preparation (II).

4.3 Northern Finland Birth Cohort 1966 (III, IV, and V)

Data for sub-studies III–V were from the population-based Northern Finland Birth Cohort 1966 study (NFBC1966). NFBC1966 is a life-course study involving participants whose dates of birth was expected to be in 1966 in Northern Finland (N = 12,058). In general, eligible participants for sub-studies III–V were those members of NFBC1966 who had participated in the latest follow-up performed at the age 46 years, and wore a hip-worn (III and IV) and a wrist-worn activity monitor (V) for the device-based measurement of daily activities (Fig. 2).

Briefly, the 46-year follow-up included completion of postal questionnaires. Cohort members were also invited to attend a clinical examination after fasting overnight for 12 hours for the collection of fasting blood samples and anthropometric measurements, and an oral glucose tolerance test on a separate day. Anthropometric and behavioral information on cohort members participating in the 46-year follow-up and attending the clinical examination day (n = 5,840) is shown in Table 3.

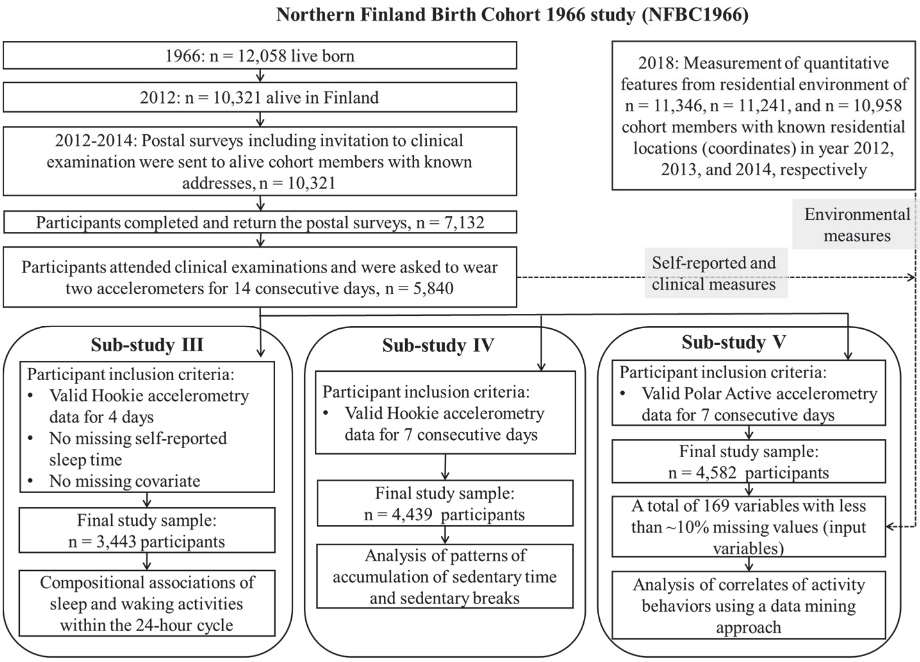

Fig. 2. The collected data in the latest follow–up of Northern Finland Birth Cohort 1966, and how they were used in sub-studies III–V.

Table 3. Anthropometric and behavioral information on cohort members participating in the 46-year follow-up, and attending the clinical examination day (n = 5,840)

Variable	Men (n = 2,565)	Women (n = 3,257)
Height, cm	178.5 (6.3)	164.8 (6.0)
Weight, kg	86.9 (14.9)	71.7 (14.8)
Body fat, %	23.4 (7.2)	33.2 (8.4)
Fat mass, kg	21.1 (9.9)	24.8 (11.1)
BMI, kg/m^2	27.3 (4.3)	26.5 (5.3)
Waist circumference, cm	97.6 (11.8)	87.2 (13.1)
Health-related quality of life score	0.93 (0.1)	0.91 (0.1)
Education		
Secondary school	102 (4.3%)	81 (2.7%)
Vocational/college level education	1673 (71.3%)	1976 (65.5%)
Polytechnic/university degree	571 (24.3%)	960 (31.8%)
Employment status		
Employed	2029 (88.2%)	2643 (88.1%)
Unemployed	158 (6.9%)	137 (4.6%)
Other (e.g., student, homemaker)	113 (4.9%)	220 (7.3%)
Smoking status		
Non-smoker	1133 (47.8%)	1808 (58.3%)
Former smoker	728 (30.7%)	757 (24.4%)
Current smoker	511 (21.5%)	534 (17.3%)

Values are mean (SD) or count (%). BMI = body mass index.

4.3.1 Measurements

Questionnaires

A postal questionnaire was sent to all living cohort members with known addresses (n = 10,321) in 2012–2014. The questionnaire included items on social background, frequency and intensity of physical activity, physical and psychological health, and socioeconomic situation. Previous diagnosis of hypertension, heart diseases, and diabetes as well as medication use for hypertension, high cholesterol, and diabetes were also measured. Health-related behaviors and quality of life were measured by a separate questionnaire (Sintonen, 2001). Temperament and personality trait scores were assessed from responses to another separate survey that was used to address opinions and experiences (Cloninger et al., 1994).

Clinical examination

On the day of clinical examination, participants abstained from smoking and drinking coffee. Trained nurses measured participants' height, weight, blood pressure, and calculated waist-hip ratio and BMI. Participants' body composition was measured with bio-impedance measurement (InBody720, InBody, Seoul, Korea) (Jensky-Squires et al., 2008). A static back muscle strength test (Biering-Sorensen trunk extension test) was performed to evaluate physical performance (Biering-Sørensen, 1984). A submaximal four-minute single-step test during which heart rate was continuously monitored was performed to assess cardiorespiratory fitness (Kiviniemi et al., 2017). In the laboratory, fasting blood samples were drawn and stored for further analyses. Participants who were not previously diagnosed with diabetes underwent a 75g oral glucose tolerance test (Alberti et al., 1998) on a second fasted examination day.

Environmental measures

Residential coordinates of all cohort members whose residences were known at the time of the 46-year follow-up data collection (2012–2014) were obtained from the Finnish Population Register Centre. A geographic information system (ArcGIS 10.3) was used to calculate a comprehensive list of built, natural, and socioeconomic environment variables that might describe the conduciveness of participants' residential environment to physical activity (Kärmeniemi et al., 2019). All the variables were calculated in the year the participant attended the 46-year data collection. Quantitative environmental features were estimated using a one-kilometer-radius circular buffer around the residential locations, and the distances (as the crow flies) to amenities were measured using road network data.

4.3.2 Monitoring and assessment of daily activities

Those cohort members attending the clinical examination were also asked to wear two accelerometers for the measurement of daily activities for 14 consecutive days. Participants were instructed to wear a wrist-worn accelerometer-based activity monitor (Polar Active, Polar Electro Oy, Kempele, Finland) on their non-dominant hand and a hip-worn accelerometer (Hookie AM20, Traxmeet Ltd, Espoo, Finland) on the right side of their hip. They were asked to wear the wrist-worn activity monitor continuously for 24 hours and the hip-worn accelerometer during all waking hours except when engaged in water-based activities. Polar Active has a uniaxial activity monitor that outputs estimated energy expenditure in MET values every 30 seconds (Kinnunen et al., 2019). Hookie is a research-based accelerometer (Aittasalo et al., 2015), and was set to measure and store raw acceleration signals at 100 Hz.

Processing raw hip-worn acceleration data (III and IV)

The hip-worn raw acceleration data from Hookie was segmented into 6-second epochs and the mean amplitude deviation (MAD) of the resultant acceleration was computed for each segment. MAD describes the mean distance of data points around the mean (MAD $= \frac{1}{n}\sum|r_i - \bar{r}|$) where n is the number of data point in the windowed segment, r_i the i^{th} resultant data point within the windowed segment, and $\bar{r}$ the mean resultant value of the windowed segment. An excellent agreement between MAD values from Hookie and the commonly-used Actigraph GTX3 accelerometer has been reported (Aittasalo et al., 2015). Non-wear time intervals were removed from the 6-second MAD values. Non-wear intervals were identified with a widely used approach for count-based data (see sub-study III for more details) (Choi et al., 2011). A valid day was defined as ≥10 hours of monitor wear time.

Time spent asleep and in sedentary and physical activities (III)

In sub-study III, activity intensities were estimated from the hip-worn accelerometry data, and eligible participants were required to provide ≥4 valid days of accelerometry. The detected wear-time intervals were cross-referenced with self-reported sleep times (captured with two questions: "At what time do you normally go to bed?" and "At what time do you normally get out of bed?"), and all accelerometer data that overlapped with a sleep interval was discarded. The remaining 6-second epochs were classified as either sedentary (sitting or lying down), standing still, light-intensity physical activity, moderate-intensity physical activity, or vigorous-intensity physical activity on the basis of MAD values (Vähä-Ypyä et al., 2018, 2015), and the duration (minutes per day (min/day)) in each activity was obtained by dividing the time spent in each activity by the number of valid days. The absolute MET cut-points used for obtaining the daily averages were as follows: sedentary and standing still: <1.5 MET, light-intensity physical activity: ≥1.5 and <3.0 MET, moderate-intensity physical activity: ≥3.0 and <6.0 MET, and vigorous-intensity physical activity: ≥6.0 MET. Further differentiation between standing still and sedentary (sitting or lying) was performed using a recently validated approach (Vähä-Ypyä et al., 2018). This approach enables posture estimation from hip-based raw acceleration data on the basis of constant Earth's gravity vector and upright walking posture, and it has shown good to excellent accuracy when compared with thigh-worn posture classification as ground truth under free-living conditions (Vähä-Ypyä et al., 2018). For the purposes of sub-study III, LPA constituted the sum of all min/day spent standing still and engaged in light-intensity physical activity, and MVPA was the sum of min/day spent in moderate- and vigorous-intensity physical activity. Sleep duration was self-reported in response to the question "How many hours do you sleep on average per day?" Responses were converted to min/day spent asleep.

Variables for characterizing sedentary time and sedentary breaks (IV)

The patterns and levels of sedentary time and physical activities may vary substantially between weekdays and weekends (Ekblom-Bak et al., 2015; Ortega et al., 2013). In sub-study IV, participants with at least seven consecutive valid days (one full week) from the hip-worn accelerometry data were included in the analyses to minimize the effects of these variations on the analyses. The detected wear-time intervals were marked according to the intensity as either sedentary behavior (≤1.5 MET), LPA (1.5–3.0 MET), or MVPA (≥3.0 MET) on the basis of MAD values (Aittasalo et al., 2015; Vähä-Ypyä et al., 2015). According to consensus definitions (Tremblay et al., 2017), we identified all the sedentary bouts lasting for ≥1 minute with no tolerance time and considered the whole time period between two consecutive sedentary bouts as sedentary breaks, if no time-epoch was marked as non-wear, starting from the first sedentary bout to the end of the second sedentary bout (Fig. 3).

Thereafter, we computed 10 variables to describe the accumulation pattern of sedentary bouts, and 55 variables to describe how these sedentary bouts of different lengths were interrupted. The variables computed for characterizing the accumulation patterns of sedentary time included duration (in minutes) and frequency (number) of 1–5-minute, 5–10-minute, 10–15-minute, 15–30-minute, and ≥30-minute sedentary bouts. These variables were averaged across seven consecutive valid days to derive per-day values. The accumulation pattern variables computed for describing the characteristics of sedentary breaks were the total duration of sedentary break and accumulated MVPA time, LPA time, and sedentary time (in bouts of <1 minute) in the sedentary breaks. We also computed the frequency (number) of <5-minute, 5–10-minute, and ≥10-minute LPA and MVPA bouts, and the frequency (number) of <1-minute sedentary bouts within the sedentary breaks. We stratified all the variables describing the characteristics of sedentary breaks based on the length of their precedent sedentary bout (i.e., 1–5 minute, 5–10 minute, 10–15 minute, 15–30 minute, ≥30-minute sedentary bouts), and averaged them over the number of corresponding sedentary bouts to derive per-sedentary-bout values. We used these per-sedentary-bout values for describing the characteristics of sedentary breaks since they would altogether be indicative of the total duration and frequency of sedentary breaks, how much (total duration) LPA and MVPA were included in the sedentary breaks on average per sedentary bout, and how often (number) these LPA and MVPA were accumulated in bouts of <5 minute, 5–10 minute, and ≥10 minute.

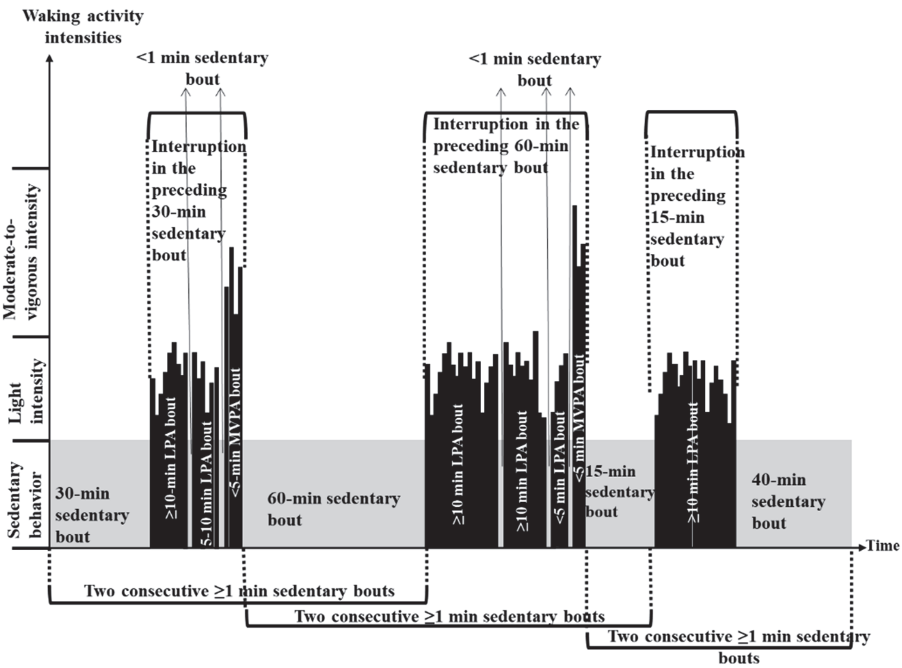

Fig. 3. Schematic representation of how sedentary bouts and sedentary breaks were defined and identified in sub-study IV.

Machine-learned activity behaviors (V)

In sub-study V, we defined participants' activity behaviors based on their machine-learned activity profiles, which we built in a previous study (Niemelä et al., 2019) using the outputs of the wrist-mounted (Polar Active) accelerometer. The recent evidence suggests that any single unidimensional metric (including the most commonly used criterion that defines physical inactivity as the insufficient activity level to meet present recommendations (Warburton et al., 2017)) might not be enough to define individuals' activity behaviors (Chaput et al., 2014; Rosenberger et al., 2019; Stamatakis, Ekelund, et al., 2019). A distinct aspect of the current approach is that continuous accelerometer-measured activity intensities in one full week across the whole intensity continuum (sedentary, LPA, and MVPA) were incorporated into an ML approach to obtain the activity profiles using a multidimensional approach (Niemelä et al., 2019). The Polar Active device was worn continuously for 24 hours during the measurement period. Briefly, accelerometer-measured activity intensities (sedentary, LPA, and MVPA) in one full week were incorporated into an ML approach to create the activity profiles (n = 4,582, valid day defined as ≥10 hours of monitor wear time). Four distinct activity groups (clusters) were established, and were named Inactive (n = 1,881), Moderately active (n = 802), Evening active (n = 1,297), and Very active (n = 602) (Niemelä et al., 2019). For the purposes of sub-study V, we defined those in the Moderately active, Evening active, or Very active clusters as active (n = 2,701), and the remaining ones in the Inactive cluster as inactive (n = 1,881). From the same seven consecutive valid measurement days used to establish the activity profiles, total time (min/day) spent in sedentary (≤2.0 MET), LPA (2.0–3.5 MET), and MVPA (≥3.5 MET) was also estimated using previously validated cut-points for Polar Active (Jauho et al., 2015). These cut-points have shown to provide comparable results to commonly used Actigraph GT3X accelerometer across different activity intensity categories (Leinonen et al., 2016).

4.3.3 Cardiometabolic health markers and confounders

Cardiometabolic health markers (sub-studies III and IV) included a range of adiposity markers including waist circumference, BMI, body fat, fat mass, and visceral fat area. Additionally, fasting blood samples from participants were analyzed for plasma glucose, serum insulin, total cholesterol, high-density lipoprotein (HDL) cholesterol, low-density lipoprotein (LDL) cholesterol, and triglycerides as previously described elsewhere (Kiviniemi et al., 2017). The ratios of total to HDL (total/HDL cholesterol ratio) and LDL to HDL (LDL/HDL cholesterol ratio) cholesterol levels were computed (Millán et al., 2009). The homeostasis model assessment of insulin resistance (HOMA-IR) was calculated from fasting plasma glucose and insulin levels (Wallace et al., 2004). For those participants who underwent a 75g oral glucose tolerance test (Alberti et al., 1998), two-hour postload plasma glucose and insulin levels were also obtained.

Potential confounders were chosen a priori based on previous research. Sex and birth weight were extracted from medical records. Other confounders were self-reported education level, employment status, marital status, household income, lifestyle (smoking status and alcohol consumption), health-related quality of life, and use of medication (for hypertension, high cholesterol, and diabetes).

4.4 Data analysis

4.4.1 Machine learning for activity intensity classification (II)

In sub-study II, we evaluated the generalization capability of ML models developed for activity intensity classification from raw acceleration data. We tested the generalization performance of models validated within one population to independent ones with different characteristics and accelerometers performing different sets of activities from the one used to develop the model (Fig. 4 (a)). For this, the placement-specific classification models were validated using leave-one-subject-out (LOSO) cross-validation within each dataset (Staudenmayer et al., 2012), and the most optimal fit with the highest accuracy for each dataset was obtained. The final models were trained with the data of all participants, and then cross-validated on independent populations with similar accelerometer placement (out-of-sample testing) to evaluate their generalization performance.

Additionally, we examined whether the generalization performance of intensity prediction models on independent populations can be improved by incorporating the information that acceleration data from different body sites (i.e., hip or wrist) acquired from various populations might contain (Fig. 4 (b)). For this, to keep the validation set independent, one dataset was used for validation at a time and was left out from model development (it was not used as training data). Then, using the remaining datasets, a merged dataset consisting of a combination of both hip- and wrist-based datasets was built and used as a training set to train an intensity classification model. The trained model using the merged training set was then validated with the left-out population. To find the most optimal model, all possible dataset combinations were tested to find the optimal training sets achieving the highest accuracy in predicting activity intensity categories in the left-out dataset. For each created merged training set, the most optimal fit was found, and the most optimal model for classifying the left-out dataset was selected as the final model. These steps were repeated when one of the five datasets was left out at a time to find the optimal merged training set for all the five left-out datasets.

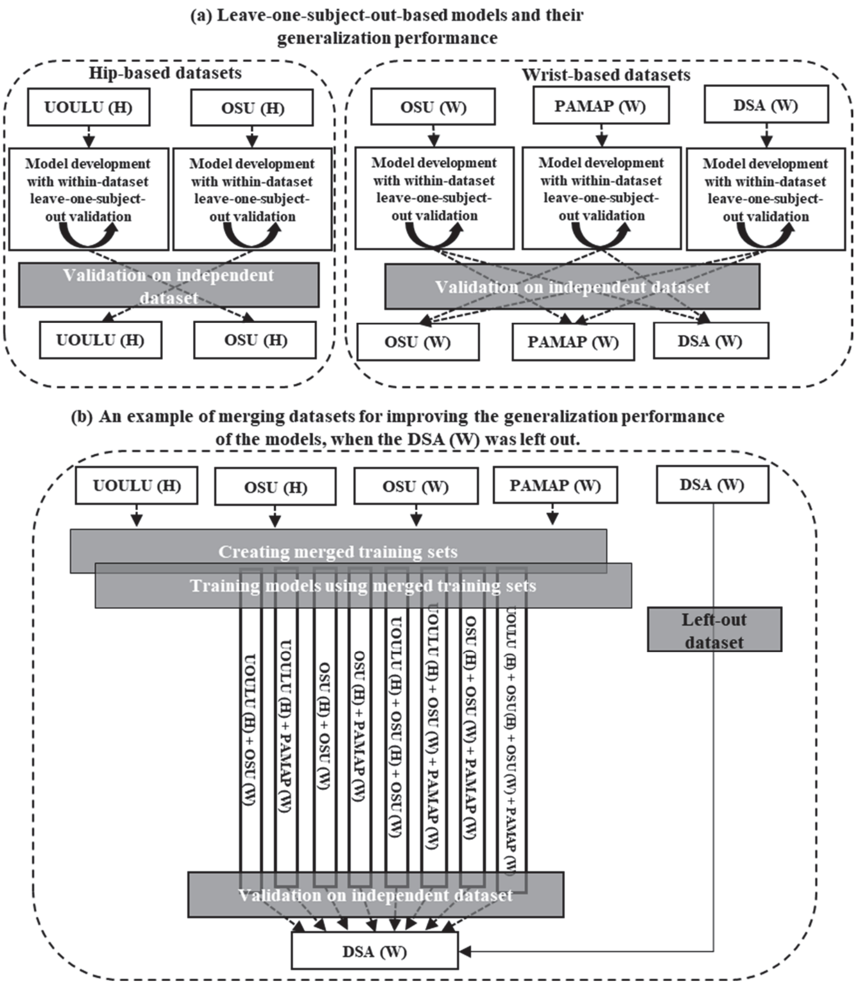

Fig. 4. General schema of the tests in sub-study II.

Feature extraction, classification algorithm, and performance evaluation

The same feature sets were extracted for all five datasets/placements. The non-overlapping 60-second window length was chosen to segment the data, which is an appropriate window length for activity classification for both youths (Trost et al., 2012) and adults (Ellis et al., 2014). For each interval, thirteen time- and frequency-domain features were extracted from the three axes of measurement (i.e., x, y, and z) and the vector magnitude (i.e., $\sqrt{x^2 + y^2 + z^2}$), resulting in a total of 52 features (see the full list of extracted features in sub-study II). Artificial neural networks were selected to classify the three different activity intensities: sedentary behavior, LPA, and MVPA. Confusion matrices were used to evaluate the performance of ANN models, and weighted Kappa statistics (K) were calculated to ensure that the overall predication accuracy was not done by chance (Cohen, 1968).

4.4.2 Sedentary time, physical activities, and cardiometabolic health in NFBC1966 (III and IV)

Compositional data analysis (III)

In study III, the 24-hour time-use composition for each participant was created by linearly rescaling the duration of all activities to sum to a total of 1,440 min/day. The 24-hour movement behavior composition was described using compositional (geometric) means (Chastin, Palarea-Albaladejo, et al., 2015). The variation matrix, which provides a proper estimation of dispersion in compositional data (Chastin, Palarea-Albaladejo, et al., 2015), was also calculated for the movement behavior composition based on the variances of the logs of all pair-wise ratios between behaviors (e.g., variance of ln (sedentary behavior/LPA)).

All cardiometabolic outcome variables were log transformed prior to compositional data analyses. Multiple linear regression was used to investigate associations of the 24-hour time-use composition with cardiometabolic health outcomes. The 24-hour movement behavior composition for each participant was expressed as ratios of its parts using isometric log ratio (ilr) transformations prior to the regression analyses (Dumuid et al., 2018). The same ilr coordinate system was used to back-transform the log-ratio coordinates into proportions for interpretation as min/day. From the results of ilr multiple linear regression models, we also estimated how time reallocations between movement behaviors were associated with differences in the cardiometabolic health markers (Dumuid et al., 2018).

Prior to regression analyses, we examined whether associations between sleep duration and cardiometabolic outcomes are U-shaped (Simonsohn, 2018). If evidence for a U-shaped association was observed for an outcome, the analysis for that outcome was stratified by sleep duration. On the basis of the existing literature (Hirshkowitz et al., 2015; Knutson, 2010) and sleep durations in this study sample, we stratified the analysis by ≤7.5 and >7.5 hours per day (hours/day) asleep for the outcomes displaying a U-shaped relationship (see sub-study III, Supplementary Material).

Profile analysis (IV)

Profile analysis was performed with the K-means clustering algorithm (Kanungo et al., 2002). K-means partitions the data into a user-defined number (K) of disjoint clusters based on the input variables (features) (Kanungo et al., 2002). All the 65 accumulation pattern variables were included in the cluster analysis. Prior to inclusion in the cluster analysis, all the input variables were standardized using the min-max method to have a range of 0–1 (Mohamad et al., 2013). The similarity of subjects was assessed using Euclidian distance, and the optimal number of clusters was obtained using the "elbow method" (Kodinariya et al., 2013).

Differences between profiles according to the variables used for describing the accumulation pattern of sedentary bouts and breaks were examined with one-way analysis of variance (ANOVA), and the p-values of the overall tests were presented. When the differences between groups were found significant ($p<0.05$), pairwise comparison was performed with Tukey post hoc tests for normally distributed variables and Kruskal–Wallis tests for skewed variables.

Linear regression models were conducted to analyze the associations (%difference) between the group/profile membership (included as categorical predictor) and each of the cardiometabolic health outcome in separate models. All the cardiometabolic outcomes were log-transformed prior to inclusion in the regression analyses. For each outcome, we tested the association with five incremental models, including an unadjusted model and four adjusted models. The unadjusted included only profile membership and cardiometabolic outcomes. Model 1 was partially adjusted for selected confounders including age, sex, education, and employment and marital status, and Model 2 was further adjusted for medication use, health-related quality of life score, smoking, alcohol consumption, and income. Model 3 was additionally adjusted for sedentary time and Model 4 for MVPA time to examine whether the associations would persist independent of these two variables. To address reverse causation (Kivimäki et al., 2019), we repeated the regression analyses after excluding those participants who had hypertension, heart diseases, and/or diabetes. For the association analyses, the group that was considered unhealthiest based on their accumulation patterns according to existing literature was selected as the referent group (Owen et al., 2010; Tremblay et al., 2017).

4.4.3 Correlates of activity behaviors in NFBC1966 (V)

In study V, we used the questionnaire and clinical and environmental measures in the 46-year follow-up (excluding those that had more than ~10% missing values) as input variables in their original form to classify the activity behavior categories (*active* and *inactive*) using the Chi-squared Automatic Interaction Detection (CHAID) decision tree algorithm (Kass, 1980). The decision tree model was created and validated with 10-fold cross-validation to ensure the robustness of the final decision tree model (Blockeel et al., 2002), and a confusion matrix was shown. In the visualization of the final tree, the percentage of *active* and *inactive* participants in each subgroup, along with the response index (RI), were presented. The RI is the percentage of inactive participants in each subgroup relative to that of inactive participants in the total sample (i.e., 41.1%), and could be utilized as an indicator of the direction and strength of the association (Lakerveld et al., 2017).

We also examined the association between factors emerging from the model and time spent in sedentary, LPA, and MVPA to determine the significance and relative importance of the methodologically identified factors. We used adjusted generalized linear mixed models, including urban–rural area as a random effect, to examine the associations between each independent variable (factor emerging in the decision tree) separately with time (min/day) spent in sedentary, LPA, and MVPA. Age and gender were used as covariates in all models. We standardized the continuous independent variables to obtain a mean of zero and a standard deviation (SD) of one before including them in regression analyses. As such, we could interpret coefficients (B) from the models encompassing a continuous independent variable as a change in the outcome (e.g., min/day of LPA) for every 1 SD change in the independent variable. We included the categorical and ordinal independent variables in the regression analyses in the form of dummy variables and set response categories at the lowest end as the reference category.

Statistical tools and implementation (II–V)

In study II, the "nnet" package in R (version 3.6.2, R Core Team, Vienna, Austria) was used to train the ANN, and each network comprised a single hidden layer with 10 nodes. In study III, all analyses were performed in R, and the compositional data analysis was performed with the packages "lmtest," "robCompositions," and "Compositions." In study IV, clustering analysis was performed in MATLAB (version R2019b, MathWorks, Natick, MA), and association analyses were performed using IBM SPSS Statistics (version 25.0, IBM Corporation, Armonk, USA). In study V, all analyses (including the decision tree analysis) were performed with IBM SPSS Statistics for Windows. For decision tree analysis, the pruning criteria were set such that groups smaller than 80 were not split any further, and no group smaller than 40 was formed. Additionally, the tree growth was limited to 10 layers, and missing values were included in the analysis as a separate category that was allowed to merge with other categories in the decision tree. A p-value of 0.05 was used to interpret significance in all statistical analyses (III–V).

5 Results

5.1 Summary of the systematic review (I)

An overview of all 62 included studies is presented in sub-study I, Table 1. The mean quality score in the included studies was high (average: 0.88, range: 0.60–1.00). The main sources of bias were small sample size and monitoring a limited set of activity types (see Appendix 1). The models were mainly developed for hip- (n = 43), wrist- (n = 31), ankle- (n = 15), and thigh-worn (n = 10) accelerometers. The predictive accuracies were therefore extracted in relation to these four attachment sites. Other accelerometer placements (e.g., chest, ear, knee, shin) were used in a total of fourteen studies, and we extracted and reported their predictive accuracies under the category "other placements." The ANN was the most used ML algorithm among the included studies (n = 32), followed by support vector machines (n = 18), random forests (n = 12), and decision tree (n = 11).

Forty-eight (77%) out of the 62 included studies calibrated and validated activity recognition models. From the studies developing activity recognition models, ten (21%) compared activity recognition models with hip and wrist, six (12%) with hip and ankle, four (8%) with hip and thigh, seven (15%) with wrist and ankle, three (6%) with wrist and thigh, and two (4%) with ankle and thigh (Fig. 5). Overall, the predictive accuracy of models developed with acceleration data from different attachment sites were highly comparable regardless of age groups. For instance, although most studies comparing hip- and wrist-based models reported a higher accuracy for the hip-based model, on average the absolute difference between accuracy of the hip- and wrist-based models was 5.5% (SD 4.6%).

Similar patterns of findings were observed when the predictive accuracy of activity recognition models from different wear locations were compared (Fig. 5). On average, the absolute difference between accuracy of the hip- and ankle-based models was 5.6% (SD 6.3%), between hip- and thigh-based models was 8.9% (SD 5.6%), between wrist- and ankle-based models was 7.1% (SD 5.3%), and between wrist- and thigh-based models was 6.0% (SD 7.2%). Six studies tested the validity of activity recognition models in an independent population. All those studies reported accuracy degradation when the models cross-validated on an independent population (range: ~2%–30%).

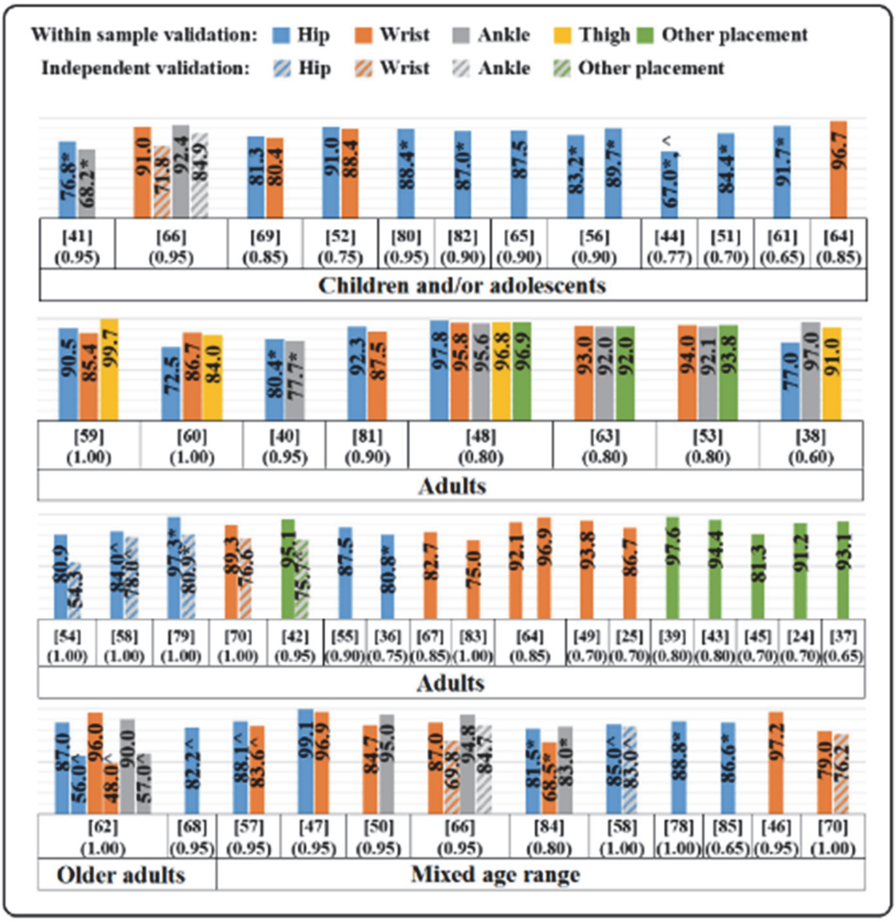

Fig. 5. The extracted prediction accuracies (overall accuracy in %, shown inside the bars) of activity recognition models in relation to accelerometer placement, categorized by the age range of the population from which the models were developed. The literature reference numbers and quality scores are presented in square brackets [ ] and parentheses ( ), respectively. The list of reference numbers is according to the reference list in sub-study I. *Indicates the data type was not raw acceleration. ^Indicates the data was acquired in free-living settings.

5.2 Generalization performance of machine-learned models for activity intensity classification (II)

In LOSO cross-validation (Fig. 6), with both raw and tailored data, the overall classification accuracy of the hip- and wrist-based models was high, achieving above 80% (range: 81.8–95.4%); the only exception was PAMAP (W), which showed a slightly lower classification accuracy (raw: 71.9%, tailored: 79.6%). Across all five datasets, the differences in classification accuracy of the three intensity categories with raw and tailored data were marginal (range: 0.2–9.2 percentage points), resulting in marginal differences in overall classification accuracy (0.7–7.7 percentage points) and Kappa values (K = 0.01–0.09).

When the models were cross-tested with another dataset (Fig. 7), all models demonstrated lower performance than those obtained by within-dataset cross-validation (Fig. 6). The reduction in classification accuracy ranged from 20.1 to 44.3 percentage points. Raw and tailored data minimally affected the overall accuracies (0.7–13.1 percentage points) and Kappa values (K = 0.01–0.11).

The performance of the ANN models trained on merged training sets in the classification of activity intensities in another population that was not part of the training data is shown in Fig. 8. All possible dataset combinations were analyzed, having one population left out at a time, but only the most optimal results are reported here. In all cases (with both raw and tailored data), the models trained on the merged datasets yielded a better generalization performance (Fig. 8) compared to those obtained by placement-specific models (Fig. 7). This was primarily attributable to the better classification of all three intensity categories. The overall classification accuracy and agreements across the datasets were lower (2.5–19.3 percentage points, K = 0.01–0.21), but approached those obtained by within-dataset cross-validation (Fig. 6), and were slightly higher for the dataset PAMAP (W).

Validation: OULU (H)

Raw				
Class	SB	LPA	MVPA	Accuracy:
SB	81.1	17.9	1.1	83.4
LPA	9.1	85.5	5.5	(78.1-88.8) Kappa, K:
MVPA	1.2	17.8	81.1	0.78

Tailored				
Class	SB	LPA	MVPA	Accuracy:
SB	83.2	15.3	1.6	85.6
LPA	9.7	86.7	3.6	(80.2-91.0) Kappa, K:
MVPA	0.4	14.7	84.9	0.81

Validation: OSU (H)

Raw				
Class	SB	LPA	MVPA	Accuracy:
SB	98.0	0.6	1.4	94.4
LPA	6.2	82.3	11.5	(93.0-95.8) Kappa, K:
MVPA	1.6	2.6	95.8	0.92

Tailored				
Class	SB	LPA	MVPA	Accuracy:
SB	98.6	0.3	1.1	95.4
LPA	3.2	85.3	11.5	(94.1-96.8) Kappa, K:
MVPA	0.7	2.7	96.6	0.94

Validation: OSU (W)

Raw				
Class	SB	LPA	MVPA	Accuracy:
SB	96.8	2.3	0.9	87.8
LPA	4.8	58.1	37.1	(83.8-91.9) Kappa, K:
MVPA	2.0	4.4	93.6	0.85

Tailored				
Class	SB	LPA	MVPA	Accuracy:
SB	96.4	0.0	3.6	87.1
LPA	8.1	57.3	34.7	(81.5-92.6) Kappa, K:
MVPA	1.2	6.1	92.7	0.84

Validation: PAMAP (W)

Raw				
Class	SB	LPA	MVPA	Accuracy:
SB	66.4	30.0	3.6	71.9
LPA	17.8	51.4	30.8	(60.0-83.9) Kappa, K:
MVPA	4.9	11.4	83.8	0.63

Tailored				
Class	SB	LPA	MVPA	Accuracy:
SB	68.2	12.7	19.1	79.6
LPA	9.3	53.3	37.4	(65.5-93.8) Kappa, K:
MVPA	2.2	4.9	93.0	0.72

Validation: DSA (W)*

Raw				
Class	SB	LPA	MVPA	Accuracy:
SB	71.7	11.7	16.7	81.8
LPA	2.5	48.1	49.4	(73.6-89.9) Kappa, K:
MVPA	1.1	3.4	95.6	0.66

Tailored				
Class	SB	LPA	MVPA	Accuracy:
SB	71.7	11.7	16.7	81.8
LPA	2.5	48.1	49.4	(73.6-89.9) Kappa, K:
MVPA	1.1	3.4	95.6	0.66

Fig. 6. Confusion matrices showing the performance of ANN models in activity intensity classification with raw data and tailored data (acceleration limited to ±5 g) validated using leave-one-subject-out cross-validation within datasets. The values across the intensity categories and overall accuracy (95% confidence interval) are presented in percentage (%). ∗: The raw data and tailored data were similar. SB = sedentary behavior, LPA = light physical activity, MVPA = moderate-to-vigorous physical activity.

Model: UOULU (H), Validation: OSU (H)

Raw				
Class	SB	LPA	MVPA	Accuracy: 46.5 (43.9-49.1) Kappa, K: 0.38
SB	49.6	46.3	4.1	
LPA	37.4	50.6	12.0	
MVPA	1.4	54.6	44.1	

Tailored				
Class	SB	LPA	MVPA	Accuracy: 52.2 (48.8-55.7) Kappa, K: 0.47
SB	66.2	32.9	0.9	
LPA	53.1	46.6	0.2	
MVPA	2.8	49.8	47.5	

Model: OSU (H), Validation: UOULU (H)

Raw				
Class	SB	LPA	MVPA	Accuracy: 59.9 (54.1-65.7) Kappa, K: 0.48
SB	64.7	1.6	33.7	
LPA	31.5	30.0	38.5	
MVPA	1.2	5.0	93.8	

Tailored				
Class	SB	LPA	MVPA	Accuracy: 59.1 (55.1-63.1) Kappa, K: 0.47
SB	74.2	3.2	22.6	
LPA	27.9	37.6	34.5	
MVPA	1.5	23.9	74.5	

Model: OSU (W), Validation: PAMAP (W)

Raw				
Class	SB	LPA	MVPA	Accuracy: 51.8 (43.2-60.5) Kappa, K: 0.39
SB	77.2	12.7	10.1	
LPA	38.4	23.9	37.7	
MVPA	3.8	32.4	63.8	

Tailored				
Class	SB	LPA	MVPA	Accuracy: 54.0 (47.4-60.6) Kappa, K: 0.44
SB	87.3	0.9	11.8	
LPA	43.0	13.1	43.9	
MVPA	8.6	35.1	56.2	

Model: OSU (W), Validation: DSA (W)

Raw				
Class	SB	LPA	MVPA	Accuracy: 44.1 (39.5-48.7) Kappa, K: 0.25
SB	92.5	0.0	7.5	
LPA	58.1	10.0	31.9	
MVPA	28.2	28.4	43.4	

Tailored				
Class	SB	LPA	MVPA	Accuracy: 42.8 (37.2-48.4) Kappa, K: 0.21
SB	94.2	1.7	4.2	
LPA	52.5	9.4	38.1	
MVPA	38.9	20.0	41.1	

Model: PAMAP (W), Validation: OSU (W)

Raw				
Class	SB	LPA	MVPA	Accuracy: 43.9 (39.0-48.8) Kappa, K: 0.30
SB	79.5	18.2	2.3	
LPA	35.5	49.2	15.3	
MVPA	24.4	57.3	18.3	

Tailored				
Class	SB	LPA	MVPA	Accuracy: 57.0 (48.9-65.1) Kappa, K: 0.41
SB	88.2	9.1	2.7	
LPA	58.1	15.3	26.6	
MVPA	35.5	14.0	50.6	

Model: PAMAP (W), Validation: DSA (W)

Raw				
Class	SB	LPA	MVPA	Accuracy: 45.8 (41.2-50.4) Kappa, K: 0.21
SB	15.0	82.5	2.5	
LPA	18.1	48.8	33.1	
MVPA	8.2	39.2	52.6	

Tailored				
Class	SB	LPA	MVPA	Accuracy: 41.2 (34.8-47.6) Kappa, K: 0.22
SB	97.5	2.5	0.0	
LPA	86.9	3.8	09.4	
MVPA	52.6	7.8	39.6	

Model: DSA (W), Validation: OSU (W)

Raw				
Class	SB	LPA	MVPA	Accuracy: 44.1 (39.5-48.7) Kappa, K: 0.25
SB	92.5	0.0	7.5	
LPA	58.1	10.0	31.9	
MVPA	28.2	28.4	43.4	

Tailored				
Class	SB	LPA	MVPA	Accuracy: 42.8 (37.2-48.4) Kappa, K: 0.21
SB	94.2	1.7	4.2	
LPA	52.5	9.4	38.1	
MVPA	38.9	20.0	41.1	

Model: DSA (W), Validation: PAMAP (W)

Raw				
Class	SB	LPA	MVPA	Accuracy: 49.1 (37.6-60.5) Kappa, K: 0.24
SB	28.2	25.5	46.4	
LPA	25.2	13.1	61.7	
MVPA	1.1	12.4	86.5	

Tailored				
Class	SB	LPA	MVPA	Accuracy: 49.8 (38.5-61.1) Kappa, K: 0.25
SB	28.2	25.5	46.4	
LPA	25.2	13.1	61.7	
MVPA	1.1	10.8	88.1	

Fig. 7. Confusion matrices showing the performance of leave-one-subject-out validated ANN models in activity intensity classification in an independent population group with raw data and tailored data (acceleration limited to ±5 g). The values across the intensity categories and overall accuracy (95% confidence interval) are presented in percentage (%). SB = sedentary behavior, LPA = light physical activity, MVPA = moderate-to-vigorous physical activity.

Model: OSU (H) + OSU (W) + PAMAP (W), Validation: UOULU (H)

Raw				
Class	SB	LPA	MVPA	Accuracy: 80.9 (77.8-84.0) Kappa, K: 0.77
SB	92.6	5.3	2.1	
LPA	25.5	66.1	8.5	
MVPA	2.3	7.7	90.0	

Tailored				
Class	SB	LPA	MVPA	Accuracy: 70.3 (64.9-75.7) Kappa, K: 0.66
SB	91.6	4.2	4.2	
LPA	35.5	43.3	21.2	
MVPA	3.1	9.3	87.6	

Model: UOULU (H) + PAMAP (W) + DSA (W), Validation: OSU (H)

Raw				
Class	SB	LPA	MVPA	Accuracy: 82.9 (81.4-84.4) Kappa, K: 0.79
SB	82.1	17.9	0.0	
LPA	14.7	81.3	4.0	
MVPA	1.7	14.7	83.6	

Tailored				
Class	SB	LPA	MVPA	Accuracy: 82.2 (79.7-84.6) Kappa, K: 0.79
SB	87.5	12.5	0.0	
LPA	25.7	70.8	3.5	
MVPA	1.3	16.0	82.7	

Model: UOULU (H) + PAMAP (W) + DSA (W), Validation: OSU (W)

Raw				
Class	SB	LPA	MVPA	Accuracy: 68.5 (65.1-71.9) Kappa, K: 0.64
SB	62.3	37.7	0.0	
LPA	0.0	33.9	66.1	
MVPA	1.2	14.5	84.3	

Tailored				
Class	SB	LPA	MVPA	Accuracy: 71.0 (67.7-74.4) Kappa, K: 0.67
SB	73.2	23.2	3.6	
LPA	0.0	50.8	49.2	
MVPA	0.0	23.5	76.5	

Model: OSU (H) + OSU (W) + UOULU (H), Validation: PAMAP (W)

Raw				
Class	SB	LPA	MVPA	Accuracy: 79.4 (72.7-86.2) Kappa, K: 0.73
SB	84.8	15.2	0.0	
LPA	23.2	67.4	9.4	
MVPA	1.1	10.3	88.6	

Tailored				
Class	SB	LPA	MVPA	Accuracy: 83.7 (76.9-90.4) Kappa, K: 0.78
SB	90.9	4.5	4.5	
LPA	19.6	68.2	12.1	
MVPA	1.1	10.3	88.6	

Model: OSU (H) + OSU (W) + UOULU (H), Validation: DSA (W)

Raw				
Class	SB	LPA	MVPA	Accuracy: 67.1 (63.9-70.4) Kappa, K: 0.69
SB	76.7	23.3	0.0	
LPA	3.1	63.8	33.1	
MVPA	0.2	33.9	65.9	

Tailored				
Class	SB	LPA	MVPA	Accuracy: 71.1 (66.3-75.9) Kappa, K: 0.68
SB	74.2	25.8	0.0	
LPA	6.9	63.1	30.0	
MVPA	3.6	23.4	73.1	

Fig. 8. Confusion matrices showing the performance of ANN models trained on merged datasets with raw data and tailored data (acceleration limited to ±5 g) in activity intensity classification in an independent population group. The values across the intensity categories and overall accuracy (95% confidence interval) are presented in percentage (%). SB = sedentary behavior, LPA = light physical activity, MVPA = moderate-to-vigorous physical activity.

5.3 Studies with NFBC1966 (III, IV, and V)

5.3.1 Twenty-four-hour time-use and cardiometabolic health (III)

A total of 3,443 cohort members provided valid hip-based acceleration data, questionnaires and other measurement data that were needed. Evidence for U-shaped relationships were seen for fasting serum insulin, 2-hour glucose, HOMA-IR, triglycerides, visceral fat area, and BMI (see sub-study III, Supplementary Material). Compared to the compositional means of participants who slept >7.5 hours/day, participants who slept ≤7.5 hours/day had a larger compositional mean for sedentary behavior, LPA, and MVPA (see sub-study III, Table 2). The variation matrix of the included sample, overall and stratified by sleep duration is described in sub-study III, Supplementary Material. Overall, the largest log-ratio variances all included MVPA, which indicates that MVPA was least dependent on the other movement behaviors. The lowest log-ratio variance was between sleep and sedentary behavior (0.063), which indicates more consistent proportionality (co-dependency) between these behaviors.

The composition of movement behaviors across the 24-hour day was significantly associated with each of the cardiometabolic outcomes (model P value <0.001 for all). Regardless of the shape of the association with sleep duration (Tables 4 and Table 5), relative to all other behaviors more daily time in both MVPA and LPA were consistently beneficially associated with cardiometabolic outcomes, for example 2-hour insulin (MVPA, $\beta = -0.28$; LPA, $\beta = -0.30$) and body fat (MVPA, $\beta = -0.11$; LPA, $\beta = -0.15$). For outcomes with a linear relationship with sleep duration (Table 4), relative to all other behaviors, more time spent asleep and in sedentary behaviors were both detrimentally associated with outcomes, for example total/HDL cholesterol ratio (sleep, $\beta = 0.13$; sedentary behavior, $\beta = 0.05$) and body fat (sleep, $\beta = 0.01$; sedentary behavior, $\beta = 0.16$). Time in sedentary behaviors was not associated with fasting plasma glucose and time in sleep was not significantly associated with waist circumference (although the association bordered significance, $p = 0.091$). For outcomes that showed a U-shaped relationship with sleep duration (Table 5), generally more daily sedentary behaviors relative to all the other behaviors was detrimentally associated with outcomes.

The results for time reallocations between 24-hour movement behaviors with all cardiometabolic health outcomes are presented in sub-study III, Supplementary Material. From the estimates (percentage change), it was apparent that more time in MVPA at the expense of all other behaviors was associated with favorable changes in outcomes. For instance, as shown in Fig. 9, 30 min/day more MVPA relative to the remaining behaviors was significantly associated with lower 2-hour insulin (-11.8%, 95% confidence interval (-13.9, -9.6)). In general, reallocating time from sedentary behaviors or sleep to LPA was favorably associated with outcomes but to a lesser extent compared to MVPA (Fig. 10). For instance, 60 min/day more LPA at the expense of sedentary behaviors and sleep was associated with lower 2-hour insulin (-6.1% (-7.7, -4.4) and -7.4% (-10.3, -4.5), respectively) (Fig. 10). Conversely, opposite time reallocations including adding time to any other behavior from MVPA, or generally adding time to sleep or sedentary behavior from LPA, was associated with unfavorable changes in outcomes.

Table 4. Compositional regression estimates for cardiometabolic outcomes that displayed a linear relationship with sleep duration.

Cardiometabolic markers	Sleep		SB		LPA		MVPA	
	β (95% CI)	p-value	β (95% CI)	p-value	β (95% CI)	p-value	β (95% CI)	p-value
2-hour insulin	**0.36 (0.27, 0.46)**	**<0.001**	**0.22 (0.15, 0.28)**	**0.001**	**-0.30 (-0.36, -0.24)**	**<0.001**	**-0.28 (-0.30, -0.20)**	**<0.001**
Fasting plasma glucose	**0.03 (0.01, 0.04)**	**0.05**	0.01 (0, 0.02)	0.178	**-0.03 (-0.04, -0.02)**	**0.001**	**-0.01 (-0.02, -0.10)**	**0.002**
Total/HDL cholesterol ratio	**0.13 (0.10, 0.16)**	**<0.001**	**0.05 (0.03, 0.07)**	**0.006**	**-0.11 (-0.13, -0.09)**	**<0.001**	**-0.08 (-0.09, -0.07)**	**<0.001**
LDL/HDL cholesterol ratio	**0.17 (0.12, 0.21)**	**<0.001**	**0.09 (0.06, 0.12)**	**0.003**	**-0.14 (-0.17, -0.11)**	**<0.001**	**-0.11 (-0.13, -0.10)**	**<0.001**
Body fat	**0.01 (0.06, 0.13)**	**0.004**	**0.16 (0.14, 0.19)**	**<0.001**	**-0.15 (-0.17, -0.13)**	**<0.001**	**-0.11 (-0.12, -0.10)**	**<0.001**
Fat mass	**0.13 (0.08, 0.18)**	**0.009**	**0.26 (0.23, 0.30)**	**<0.001**	**-0.24 (-0.27, -0.21)**	**<0.001**	**-0.15 (-0.17, -0.14)**	**<0.001**
Waist circumference	0.03 (0.01, 0.04)	0.091	**0.07 (0.06, 0.08)**	**<0.001**	**-0.05 (-0.06, -0.04)**	**<0.001**	**-0.04 (-0.05, -0.03)**	**<0.001**

Only the regression coefficients corresponding to the first ilr coordinate are shown, since the first ilr coordinates contain all the information relative to the remaining movement behaviors. All models were adjusted for age, sex, birth weight, education level, employment status, marital status, household income, health-related quality of life, lifestyle factors (smoking status and alcohol consumption), and medication (for hypertension, cholesterol, and/or diabetes). Significant associations are shown in bold. SB = sedentary behavior, LPA = light physical activity, MVPA = moderate-to-vigorous physical activity, HDL = high-density lipoprotein, LDL = low-density lipoprotein.

Table 5. Compositional regression estimates for cardiometabolic outcomes that displayed a U-shaped relationship with sleep duration.

Cardiometabolic markers	Sleep		SB		LPA		MVPA	
	β (95% CI)	p-value	β (95% CI)	p-value	β (95% CI)	p-value	β (95% CI)	p-value
Fasting serum insulin								
Sleep duration ≤7.5 hours/day	0.17 (0.06, 0.29)	0.134	**0.25 (0.17, 0.32)**	**0.001**	**-0.24 (-0.30, -0.18)**	**<0.001**	**-0.18 (-0.21, -0.15)**	**<0.001**
Sleep duration >7.5 hours/day	0.24 (0.08, 0.40)	0.131	**0.20 (0.11, 0.30)**	**0.036**	**-0.30 (-0.38, -0.21)**	**0.001**	**-0.15 (-0.18, -0.12)**	**<0.001**
2-hour glucose								
Sleep duration ≤7.5 hours/day	-0.02 (-0.08, 0.04)	0.725	**0.09 (0.06, 0.13)**	**0.006**	-0.01 (-0.04, 0.02)	0.695	**-0.06 (-0.08, -0.05)**	**<0.001**
Sleep duration >7.5 hours/day	0.12 (0.04, 0.13)	0.141	-0.02 (-0.06, 0.03)	0.721	-0.06 (-0.10, -0.02)	0.151	**-0.04 (-0.06, -0.03)**	**0.004**

Cardiometabolic markers	Sleep		SB		LPA		MVPA	
	β (95% CI)	p-value	β (95% CI)	p-value	β (95% CI)	p-value	β (95% CI)	p-value
HOMA-IR								
Sleep duration ≤7.5 hours/day	0.18 (0.06, 0.31)	0.152	**0.26 (0.18, 0.34)**	**0.001**	**-0.26 (-0.34, -0.19)**	**<0.001**	**-0.18 (-0.21, -0.15)**	**<0.001**
Sleep duration >7.5 hours/day	0.32 (0.15, 0.50)	0.065	0.2 (0.10, 0.31)	0.057	**-0.36 (-0.45, -0.27)**	**<0.001**	**-0.16 (-0.20, -0.13)**	**<0.001**
Triglycerides								
Sleep duration ≤7.5 hours/day	0.13 (0.04, 0.22)	0.152	**0.19 (0.13, 0.25)**	**0.001**	**-0.23 (-0.28, -0.18)**	**<0.001**	**-0.09 (-0.11, -0.07)**	**<0.001**
Sleep duration >7.5 hours/day	**0.41 (0.27, 0.54)**	**0.003**	-0.02 (-0.10, 0.06)	0.785	**-0.26 (-0.33, -0.19)**	**<0.001**	**-0.13 (-0.15, -0.10)**	**<0.001**
Visceral fat area								
Sleep duration ≤7.5 hours/day	0.12 (0.04, 0.21)	0.141	**0.23 (0.18, 0.28)**	**<0.001**	**-0.17 (-0.22, -0.13)**	**<0.001**	**-0.18 (-0.20, -0.16)**	**<0.001**
Sleep duration >7.5 hours/day	0.19 (0.08, 0.30)	0.094	**0.16 (0.09, 0.23)**	**0.021**	**-0.11 (-0.25, -0.14)**	**0.001**	**-0.15 (-0.17, -0.13)**	**<0.001**
BMI								
Sleep duration ≤7.5 hours/day	0.04 (0.01, 0.07)	0.188	**0.07 (0.05, 0.09)**	**<0.001**	**-0.06 (-0.08, -0.05)**	**<0.001**	**-0.05 (-0.06, -0.04)**	**<0.001**
Sleep duration >7.5 hours/day	**0.11 (0.06, 0.15)**	**0.022**	0.04 (0.01, 0.07)	0.13	**-0.10 (-0.13, -0.08)**	**<0.001**	**-0.05 (-0.06, -0.04)**	**<0.001**

Only the regression coefficients corresponding to the first ilr coordinate are shown, since the first ilr coordinates contain all the information relative to the remaining movement behaviors. All models were adjusted for age, sex, birth weight, education level, employment status, marital status, household income, health-related quality of life, lifestyle factors (smoking status and alcohol consumption), and medication (for hypertension, cholesterol, and/or diabetes). Significant associations are shown in bold. SB = sedentary behavior, LPA = light physical activity, MVPA = moderate-to-vigorous physical activity, BMI = body mass index, HOMA-IR = homeostasis model assessment of insulin resistance.

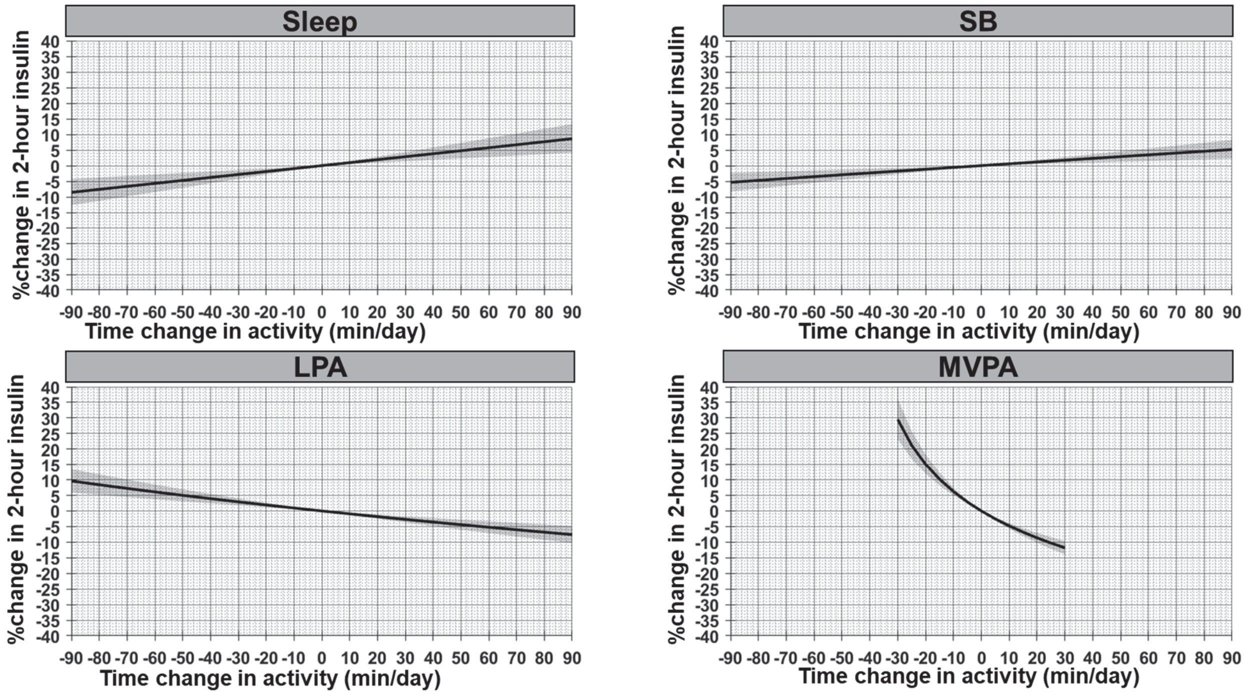

Fig. 9. **Percentage change for reallocation of time from one movement behavior relative to the remaining movement behaviors (e.g., Sixty min/day more SB is associated with a percentage change of 3.5, compared with 2-hour insulin at the mean composition). SB = sedentary behavior, LPA = light physical activity, MVPA = moderate-to-vigorous physical activity.**

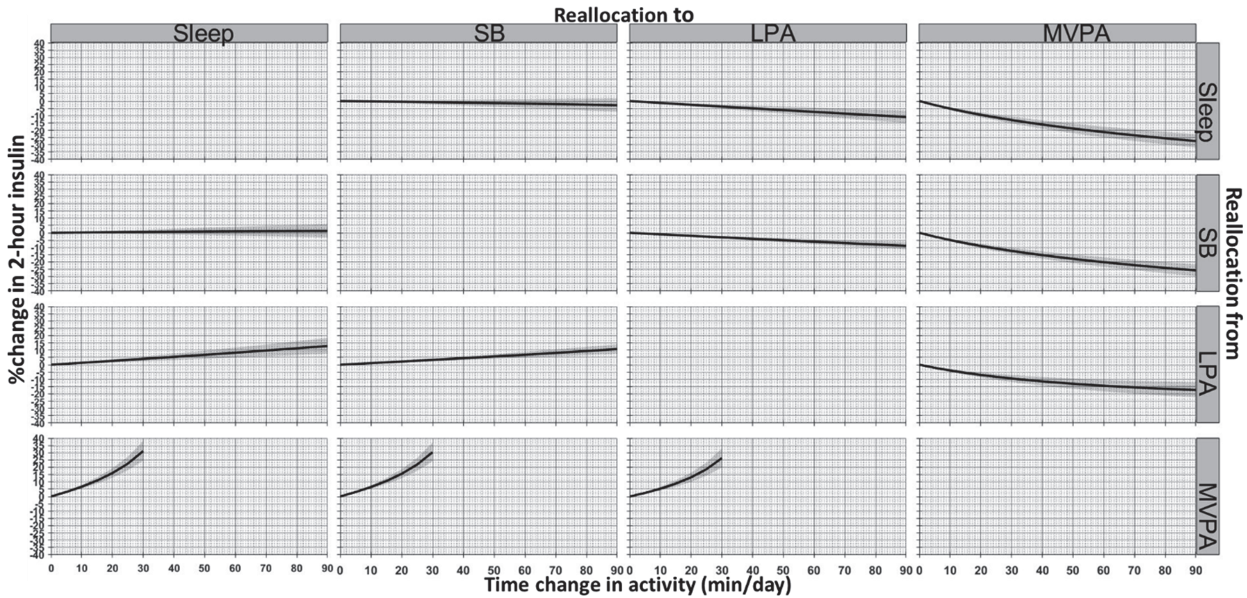

Fig. 10. Percentage change for pairwise reallocation of time from one movement behavior (rows) to another movement behaviors (columns) (e.g., row two, column three indicates the estimated difference in 2-hour insulin by reallocating time from SB to LPA, compared with 2-hour insulin at the mean composition). SB = sedentary behavior, LPA = light physical activity, MVPA = moderate-to-vigorous physical activity.

5.3.2 Sedentary time, sedentary breaks, and cardiometabolic health (IV)

The optimal number of clusters characterizing how sedentary time was accumulated and interrupted was selected as 4 (see sub-study IV, Supplementary Material). All the 65 variables describing patterns of the accumulation of sedentary time and sedentary breaks were significantly different between clusters (see sub-study IV, Table 2), indicating the relevance of all the variables used for creating the analysis (overall p-values for all variables in ANOVA tests <0.001).

Groups of participants with similar accumulation patterns were labelled according to their distinguishing accumulation patterns, as shown by high and low Z-values (see sub-study IV, Fig. 2) and means (SD) for all the 65 accumulation pattern variables (see sub-study IV, Table 2). We named these groups "Couch potatoes" (n = 1222, 28% of the sample), "Prolonged sitters" (n = 1179, 27%), "Shortened sitters" (n = 1529, 34%), and "Breakers" (n = 509, 11%). Couch potatoes had a high number of sedentary bouts of different lengths that, compared to the other three groups, were interrupted less frequently by non-sedentary bouts lasting for relatively shorter durations. The duration of interruptions of sedentary bouts of different lengths were comparable in Prolonged sitters and Shortened sitters, but Prolonged sitters accumulated most of their sedentary time in longer bouts of ≥15–30 minutes, while Shortened sitters did so in bouts of <15–30 minutes. Breakers were engaged in short sedentary bouts, which were, compared to the other three groups, more frequently interrupted by non-sedentary bouts of relatively longer duration. Since Couch potatoes spent the longest time in sedentary activities and had the shortest duration of interruptions of sedentary time, this group was considered the unhealthiest and was used as a reference for comparisons.

Overall, the average consumption of alcohol was highest among the Couch potatoes (11.3 grams per day), and they were least frequently non-smokers (49.2%) and most frequently (19.0%) on medication for diabetes, cholesterol, and/or hypertension (see sub-study IV, Table 1). The proportion of females was lowest in the Breakers (48.1%), followed by Prolonged sitters (52.9%), Couch potatoes (55.4%), and Shortened sitters (63.9%).

Table 6 shows the associations between the four distinct groups and cardiometabolic outcomes. In unadjusted regression models, Prolonged sitters had favorable differences in the markers of cardiometabolic health compared to Couch potatoes (range: 1.7%–8.1% lower values depending on the outcome), for example lower levels of 2-hour insulin (8.1%), triglycerides (5.2%), and body fat (1.7%). However, although the favorable associations for triglycerides, total/HDL cholesterol, LDL/HDL cholesterol, and visceral fat area were still significant in partially adjusted models (Model 1), none remained significant when the models were further adjusted for all potential confounders (Model 2).

When included in unadjusted models (Table 6), compared to Couch potatoes, Shortened sitters and Breakers both had favorable differences in the markers of cardiometabolic health (range: 2.1%–23.5% lower values depending on the outcome and group), for example, lower levels of 2-hour insulin (12.6% and 23.5%), fasting serum insulin (13.6% and 13.2%), triglycerides (11.1% and 12.9%), and body fat (2.2% and 5.6%). When unadjusted, Shortened sitters also had significantly lower fasting plasma glucose (2.2%) compared to Couch potatoes. These associations were all retained significantly in partially adjusted models (Model 1) and when the models were adjusted for all potential confounders (Model 2); the only exception was that Shortened sitters were not associated with a percentage difference in 2-hour glucose (in Model 1 and Model 2).

Results for adjusted models for all potential confounders and sedentary time (Model 3) and MVPA time (Model 4) can be found in sub-study IV, Table 3 and Table 4. When included in adjusted models for all potential confounders and sedentary time (see sub-study IV, Table 3 and Table 4, Model 3), compared to Couch potatoes, Shortened sitters and Breakers both had typically favorable differences in the markers of cardiometabolic health, for example, lower levels of fasting serum insulin (8.0% and 7.6%), body fat (1.8% and 2.8%), and fat mass (5.4% and 6.2%). Shortened sitters had also favorable differences in cardiometabolic health markers compared to Couch potatoes when the models were adjusted for both potential confounders and MVPA time (Model 4), such as lower levels of 2-hour insulin (6.5%), fasting serum insulin (8.8%), triglycerides (5.2%), and body fat (2.2%). However, when the models were adjusted for both potential confounders and MVPA time (see sub-study IV, Table 3 and Table 4, Model 4), compared to Couch potatoes, the differences in Breakers for 2-hour insulin, fasting serum insulin, triglycerides, 2-hour glucose, and all adiposity measures did not reach the level of significance. Similar patterns of associations were observed when the analyses were repeated after excluding those participants who had hypertension, heart diseases, and/or diabetes (see sub-study IV, Supplementary Material, Table S1and Table S2).

Table 6. Linear regression analysis of the association (percentage difference with (95% confidence intervals (CI)) between the four distinct clusters characterizing accumulation and interruption sedentary time and cardiometabolic markers among 4,439 middle-aged participants.

Cardiometabolic markers	Prolonged sitters vs. Couch potatoes		Shortened sitters vs. Couch potatoes		Breakers vs. Couch potatoes	
	%difference (95% CI)	p-value	%difference (95% CI)	p-value	%difference (95% CI)	p-value
2-hour insulin						
Unadjusted	**-8.1 (-13.8, -2.0)**	**0.011**	**-12.6 (-17.8, -7.1)**	**<0.001**	**-23.5 (-29.7, -16.8)**	**<0.001**
Model 1	-5.2 (-11.5, 1.5)	0.129	**-11.8 (-17.4, -6.0)**	**<0.001**	**-25.0 (-31.5, -18.0)**	**<0.001**
Model 2	-4.1 (-10.7, 3.0)	0.258	**-10.0 (-15.9, -3.7)**	**0.002**	**-24.3 (-31.3, -16.6)**	**<0.001**
Fasting serum insulin						
Unadjusted	-4.2 (-8.6, 0.5)	0.077	**-13.6 (-17.4, -9.7)**	**<0.001**	**-13.2 (-18.4, -7.8)**	**<0.001**
Model 1	-3.3 (-8.0, 1.5)	0.168	**-13.8 (-17.7, -9.8)**	**<0.001**	**-15.5 (-20.7, -9.9)**	**<0.001**
Model 2	-0.7 (-5.6, 4.5)	0.788	**-11.3 (-15.5, -6.9)**	**<0.001**	**-13.2 (-18.9, -7.1)**	**<0.001**
Triglycerides						
Unadjusted	**-5.2 (-8.7, -1.4)**	**0.008**	**-11.1 (-14.3, -7.8)**	**<0.001**	**-12.9 (-17.1, -8.3)**	**<0.001**
Model 1	**-5.1 (-8.8, -1.3)**	**0.009**	**-9.5 (-12.8, -6.1)**	**<0.001**	**-14.4 (-18.7, -10.0)**	**<0.001**
Model 2	-1.5 (-5.4, 2.7)	0.481	**-6.7 (-10.2, -3.0)**	**0.001**	**-11.3 (-16.1, -6.3)**	**<0.001**
Total/HDL cholesterol						
Unadjusted	**-2.8 (-4.9, -0.6)**	**0.013**	**-6.9 (-8.8, -5.0)**	**<0.001**	**-7.5 (-10.1, -4.9)**	**<0.001**
Model 1	**-3.2 (-5.3, -1.1)**	**0.003**	**-5.4 (-7.3, -3.5)**	**<0.001**	**-9.2 (-11.7, -6.7)**	**<0.001**
Model 2	-1.8 (-4.0, 0.5)	0.119	**-4.1 (-6.1, -2.1)**	**<0.001**	**-8.2 (-11.0, -5.4)**	**<0.001**
LDL/HDL cholesterol						
Unadjusted	**-3.8 (-6.9, -6.8)**	**0.018**	**-9.4 (-12.2, -6.7)**	**<0.001**	**-10.6 (-14.3, -6.9)**	**<0.001**
Model 1	**-4.1 (-7.1, -1.0)**	**0.010**	**-7.6 (-10.3, -4.8)**	**<0.001**	**-13.3 (-16.9, -9.6)**	**<0.001**
Model 2	-2.5 (-5.7, 0.9)	0.151	**-6.0 (-9.0, -3.0)**	**<0.001**	**-12.4 (-16.3, -8.3)**	**<0.001**
2-hour glucose						
Unadjusted	-1.6 (-3.7, -0.6)	0.160	**-2.1 (-4.1, -0.1)**	**0.046**	**-2.8 (-5.5, -0.1)**	**0.05**
Model 1	-1.8 (-4.1, 0.5)	0.130	-2.0 (-4.1, 0.2)	0.072	**-3.5 (-6.5, -0.5)**	**0.023**
Model 2	-2.0 (-4.3, 0.5)	0.122	-1.7 (-3.9, 0.6)	0.149	**-3.9 (-7.1, -0.7)**	**0.017**
Fasting plasma glucose						
Unadjusted	-0.5 (-1.5, 0.5)	0.299	**-2.2 (-3.1, -1.3)**	**<0.001**	0.0 (-1.3, 1.3)	0.986
Model 1	-0.4 (-1.5, 0.6)	0.342	**-1.7 (-2.6, -0.7)**	**0.001**	-0.6 (-2.0, -0.7)	0.401
Model 2	0.2 (-0.9, 1.3)	0.711	**-1.1 (-2.1, -0.1)**	**0.032**	0.4 (-1.1, 1.8)	0.617
Body fat						
Unadjusted	**-1.7 (-2.9, -0.5)**	**0.005**	**-2.2 (-3.2, -1.1)**	**<0.001**	**-5.6 (-7.0, -4.2)**	**<0.001**
Model 1	-1.1 (-2.1, 0.0)	0.053	**-3.5 (-4.5, -2.6)**	**<0.001**	**-5.3 (-6.6, -3.9)**	**<0.001**
Model 2	-0.6 (-1.7, 0.5)	0.301	**-3.0 (-3.9, -1.9)**	**<0.001**	**-4.7 (-6.0, -3.2)**	**<0.001**

Cardiometabolic markers	Prolonged sitters vs. Couch potatoes		Shortened sitters vs. Couch potatoes		Breakers vs. Couch potatoes	
	%difference (95% CI)	p-value	%difference (95% CI)	p-value	%difference (95% CI)	p-value
Fat mass						
Unadjusted	**-3.9 (-7.3, -0.5)**	**<0.001**	**-10.1 (-13.2, -7.1)**	**<0.001**	**-14.4 (-18.2, -10.3)**	**0.025**
Model 1	-2.5 (-6.0, 1.1)	0.176	**-11.5 (-14.4, -8.3)**	**<0.001**	**-15.0 (-19.0, -10.9)**	**<0.001**
Model 2	-0.9 (-4.5, 2.9)	0.656	**-9.4 (-12.6, -6.1)**	**<0.001**	**-13.1 (-17.3, -8.5)**	**<0.001**
Visceral fat area						
Unadjusted	**-5.2 (-8.2, -1.9)**	**0.002**	**-8.9 (-11.7, -5.9)**	**<0.001**	**-11.4 (-15.1, -7.5)**	**<0.001**
Model 1	**-3.8 (-7.1, -0.4)**	**0.029**	**-10.0 (-12.9, -6.9)**	**<0.001**	**-13.2 (-17.1, -9.1)**	**<0.001**
Model 2	-2.0 (-5.5, 1.6)	0.274	**-8.1 (-11.1, -4.8)**	**<0.001**	**-11.4 (-15.5, -6.9)**	**<0.001**
BMI						
Unadjusted	-1.3 (-2.6, 0.1)	0.07	**-3.7 (-5.0, -2.5)**	**<0.001**	**-2.4 (-4.0, -0.6)**	**0.007**
Model 1	-0.9 (-2.4, -0.5)	0.197	**-3.6 (-4.9, -2.4)**	**<0.001**	**-3.5 (-5.4, -1.8)**	**<0.001**
Model 2	-0.3 (-1.7, 1.2)	0.699	**-2.8 (-4.1, -1.5)**	**<0.001**	**-2.8 (-4.6, -0.9)**	**0.005**
Waist circumference						
Unadjusted	-0.8 (-2.0, 0.3)	0.170	**-3.8 (-4.9, -2.9)**	**<0.001**	**-2.4 (-3.7, -0.9)**	**0.002**
Model 1	-0.7 (-1.9, -0.4)	0.190	**-3.0 (-4.0, -2.0)**	**<0.001**	**-3.9 (-5.3, -2.6)**	**<0.001**
Model 2	-0.1 (-1.3, 1.0)	0.809	**-2.3 (-3.3, -1.2)**	**<0.001**	**-3.1 (-4.6, -2.7)**	**<0.001**

Couch potatoes was considered as the unhealthiest profile and selected as the referent group. Unadjusted models included only group membership. Model 1 was partially adjusted for age, sex, education, employment, and marital status, and Model 2 was further adjusted for medication use (for hypertension, cholesterol, and/or diabetes), health-related quality of life score, smoking, alcohol consumption, and income. Model 3 was additionally adjusted for total sedentary time, and Model 4 for total MVPA time. Results for Model 3 and Model 4 can be found in sub-study IV, Table 3 and Table 4. Significant associations are shown in bold. BMI = body mass index, HDL = high-density lipoprotein, LDL = low-density lipoprotein.

5.3.3 Data-driven correlates of activity behaviors (V)

Overall, the variables related to medication use and diseases had the highest number of missing values (~20%–50%) while the number of missing values in environmental and adiposity-related variables were lowest (~1%–5%). We used a total of 168 factors as input variables after eliminating those with over ~10% missing values (see the full list of input variables in sub-study V, Supplementary Material, Tables S1–S3).

The overall classification accuracy of the decision tree was 69.7% (Table 7). The final decision tree is shown in sub-study V, Fig. 2. The decision tree algorithm selected a total of 36 different factors of different domains, by which 54 subgroups of participants were formed, with 26 predicted as active and 28 as inactive. Overall, participants with higher body fat percentage (>31%) were more likely to be inactive (RI: 1.16 and 1.49) compared with those with lower body fat (<28%). The largest subgroup of inactive participants (n = 193, RI = 1.55) included those with the highest body fat who reported their physical activity frequency through gardening more than once a month, and were with a normalized heart rate recovery slope of <55% per second. The largest active subgroup (n = 335, RI = 0.39) was composed of participants with the lowest body fat in the study population and with a normalized heart rate recovery 60 seconds after exercise of >25 beats per minute. Participants who lived in city/rural centers and had a physically demanding occupation (i.e., process and transport workers, forestry workers and farmers, and other manual workers) had the least risk of being inactive (RI = 0.11).

Table 7. Confusion matrix showing the performance of decision tree model with 10-fold cross validation.

	Predicted outcome		
Actual outcome	Active, n	Inactive, n	Percent correct
Active, n	2014	687	74.6%
Inactive, n	705	1176	62.5%

Most continuous factors in the relatively high layers of the decision tree model and larger subgroups significantly explained min/day in all three activity intensities (Table 8). For example, body fat was positively associated with sedentary level (B = 26.5) and inversely associated with LPA (B = -16.1) and MVPA (B = -11.7) levels. Categorical and ordinal factors were also associated with min/day in sedentariness, LPA, and/or MVPA (sub-study V, Table 4). Overall, from the regression coefficients (B values, indicative of changes in min/day of sedentary, LPA, and MVPA for every 1 SD change in the predictor and of changes from the reference response categories, respectively), the associations seemed generally stronger for those factors that emerged in the higher layer and larger subgroups. For instance, a higher body fat percentage and a lower normalized heart rate recovery slope were associated with lower and higher min/day in MVPA, respectively, but the former, which appeared in the higher level of the decision tree, was associated with MVPA to a greater extent (B = -11.7 vs. 9.5).

Table 8. Associations with the whole study population (n=4,582) between the continuous factors emerged in the decision tree model and time spent in sedentary activity, light physical activity (LPA), and moderate-to-vigorous physical activity (MVPA).

Continuous factors emerged in the decision tree model	Sedentary (min/day)	LPA (min/day)	MVPA (min/day)
	B (95% CI)	B (95% CI)	B (95% CI)
Body fat percentage	26.5 (23.5, 29.6)**	-16.1 (-18.5, -13.6)**	-11.7 (-12.9, -10.6)**
Normalized heart rate recovery 60 seconds after exercise	-16.1 (-18.1, -13.4)**	9.9 (7.7, 12.1)**	9.6 (8.6, 10.6)**
Extravagance score	6.3 (3.5, 9.0)**	-3.7 (-5.9, -1.5)**	-0.6 (-1.6, 0.5)
Average weekday total sitting time	34.1 (31.5, 36.7)**	-25.3 (-27.4, -23.3)**	-5.8 (-6.8, -4.7)**
Number of workplaces	2.9 (-0.1, 6.0)	-3.2 (-5.6, -0.7)*	0.6 (-0.5, 1.7)
Normalized heart rate recovery 30 seconds after exercise	-16.9 (-19.6, -14.2)**	11.0 (8.8, 13.2)**	9.1 (8.0 to 10.1)**
Fear of uncertainty score	-1.8 (-4.6, 0.9)	0.7 (-1.6, 2.9)	-0.6 (-1.7, 0.4)
Weight	13.3 (10.3, 16.3)**	-8.4 (-10.7, -6.0)**	-3.4 (-4.5, -2.2)**
Skeletal muscle mass	-8.4 (-13.1, -3.8)**	3.5 (-0.1, 7.2)	9.6 (7.8, 11.3)**
Normalized heart rate recovery slope	-17.7 (-15.0, -20.4)**	10.6 (12.8, 8.4)**	9.5 (10.5, 8.4)**
Fitness score	-23.4 (-26.2, -20.7)**	15.1 (12.9, 17.3)**	10.7 (9.7, 11.8)**
Population density	7.5 (3.7, 11.4)**	-7.0 (-10.1, -3.1)**	0.8 (-0.5, 2.2)
Number of housing unit in row houses	3.2 (0.2, 6.2)*	-2.9 (-5.3, -0.5)*	-0.2 (-1.3, 0.9)
Average weekday sitting time at the office or other such place	27.9 (25.2, 30.6)**	-22.4 (-24.6, -20.3)**	-2.6 (-3.6, -1.5)**
Number of public transportations stops	4.1 (1.8, 8.1)**	-3.6 (-6.1, -1.0)**	0.3 (-0.8, 1.5)
Number of road accidents	5.9 (2.5, 9.4)**	-3.8 (-6.6, -1.1)**	0.1 (-1.13, 1.3)
Explorative excitability score	3.1 (0.4, 5.9)*	-2.2 (-4.4, 0.03)	0.6 (-0.4, 1.7)
Overall health-related quality of life score	-8.4 (-11.1, -5.7)**	5.81 (3.6, 7.9)**	4.1 (3.1, 5.1)**
Lean body mass	-8.8 (-13.4, -4.3)**	3.1 (0.3, 7.6)*	9.4 (7.8, 11.2)**
Disorderliness score	6.4 (3.6, 9.1)**	-4.8 (-6.1, -2.6)**	-0.7 (-1.8, 0.3)
Impulsiveness score	2.2 (-0.5, 4.1)	-1.5 (-3.7, 0.7)	-0.4 (-1.4, 0.6)
Average weekday computer use time	14.7 (11.9, 17.5)**	-11.4 (-13.7, -9.2)**	-4.5 (-5.5, -3.4)**

The regression coefficients (B) with (95% confidence interval) from generalized linear mixed model controlling for gender and age with urban-rural area as a random effect. *p<0.05; **p<0.01. min/day = minutes per day.

6 Discussion

The present study reviewed studies calibrating and validating wearable accelerometers using machine learning approaches, and evaluated the generalization capability of machine learning models developed for classifying activity intensities across the full intensity continuum in terms of sedentary behavior, LPA, and MVPA. Additionally, this study examined the associations between levels and patterns of activity across the entire intensity continuum with cardiometabolic health markers in a large population-based sample of Finnish adults, and a data-driven hierarchy for predicting their machine-learned activity behaviors was created. Based on the systematic review, overall predictive accuracies of ML-based models developed for activity recognition could be acceptable irrespective of accelerometer placement. Our cross-dataset tests revealed that ML models (i.e., ANN) developed with raw acceleration data and a within-population cross-validation technique are not generalizable to other populations monitored with different accelerometers. Overall, the findings in this study suggest that adults may gain cardiometabolic health benefits not only by MVPA, but also through LPA. Finally, the data-driven hierarchy of correlates created here consisted of factors of relative importance, and can be used to target and tailor interventions for promoting physical activity among middle-aged adults.

6.1.1 Accelerometers and machine leaning approaches (I and II)

From the systematic review (sub-study I), it was apparent that some ML approaches including ANN, support vector machines, and random forests were more commonly used than others. Although studies comparing ML approaches to cut-points and regression equations have generally reported better predictive abilities for ML models (Ellis et al., 2016; Lyden, Kozey-Keadle, et al., 2014; Trost et al., 2017), it is still difficult to identify a single ML technique that is universally better than others for activity recognition. Hence, comparative studies examining the predictive ability of different ML techniques developed for activity recognition have reported that the predictive accuracy of different ML approaches are comparable, (e.g., Trost et al., 2017; Zhang, Rowlands, et al., 2012). This suggests that certain ML techniques could be adopted in order to harmonize data processing methods, and to avoid the proliferation of data processing methods and further methodological discrepancies.

Currently, it seems that there is no agreement in the existing literature on a single optimal accelerometer placement that would provide the highest overall predictive accuracy. Previous studies have suggested that accelerometry results from different wear locations may be incomparable (Ellis et al., 2016). To narrow down the gap between differences in accelerometry results, the existing literature has demanded consensus on both data measurement and processing protocols to enhance the interstudy comparison among future accelerometry-based studies (Strath et al., 2012; Wijndaele et al., 2015). We found that the variation among the reported overall predictive accuracy of ML models is minimally related to accelerometer placement, as opposed to cut-points and regression equations (Kim et al., 2012; Migueles et al., 2019). However, one important observation in the systematic review was that a limited number of studies tested the validity of their models outside of the development dataset. Therefore, it remains unclear whether these minimal differences in the predictive accuracies of ML models would remain when applied to another population and free-living data. Nevertheless, it is probable that standardized data measurement and processing protocols could be easier to achieve and adopt, considering that advanced ML modeling approaches may provide the opportunity of processing accelerometer data with minimal accuracy loss according to the wear location.

Based on the findings of systematic review, moving towards standardization of accelerometry data processing methods may require further clarification of the optimal values for different parameters required to calibrate and validate accelerometer-based activity monitors using ML approaches, such as sampling rate and window length. In particular, it is still elusive which input features from which measurement axes (x, y, z, and/or VM) should be selected in relation to accelerometer placement for prediction of activities in terms of type, category, intensity, or energy expenditure. An alternative method that does not require feature selection is the deep learning approach. Novel deep learning approach may eventually facilitate the standardization of accelerometry data processing techniques considering that they can automatically identify and learn the representations of raw accelerometry data that are needed for activity prediction (Wang et al., 2019). Nevertheless, it appears that to date most researchers have continued to use cut-points to interpret accelerometer data (Migueles et al., 2017). This is partly because cut-point-based methods are relatively easier to deploy and understand (Migueles et al., 2017). To facilitate the employment of ML approaches, further collaboration may therefore be required between the measurement experts, statisticians, developers, and end-users in order to develop easy-to-deploy and point-and-click platforms for ML techniques, and eventually to move towards standardized data processing approaches (Bassett Jr et al., 2012; van Hees et al., 2016).

Our cross-dataset tests (sub-study II) highlighted that the high accuracy of within-dataset-validated models is not transferable to another population, indicating that within-dataset cross-validation alone may not be sufficient to understand how ML models would perform in another population. In line with our findings, few studies examining the generalization capability of activity classification models have also shown that there could be a significant reduction in the performance of within-sample-validated ML models when applied to another dataset (Bastian et al., 2015; Freedson et al., 2011; Montoye, Westgate, et al., 2018). Our results extend this finding by signifying that raw accelerometry and advanced modeling techniques do not necessarily warrant the generalizability of models to different populations, whose activities are monitored by different accelerometers.

In sub-study II, our ANN activity intensity prediction models trained on a merged training set classified the activity intensities with acceptable performance in another population group with markedly different characteristics, which was not part of the training phase and was monitored using a different accelerometer. These results suggest that integrating multiple independent accelerometry datasets into the training set might be a viable approach to augment the generalization performance of the models. This finding is promising given that there has been a lack of methodologies for enhancing the robustness and generalization capability of ML models (de Almeida Mendes et al., 2018; Montoye, Nelson, et al., 2018; Montoye, Westgate, et al., 2018). The marginal differences between raw and tailored data are also encouraging because they imply that enhancing the generalization performance of intensity prediction models can be done by combining original raw acceleration data even without data preprocessing (e.g., data filtering or conversion).

6.1.2 Sedentary time, physical activities, and cardiometabolic health (III and IV)

The main finding of sub-study III was that, relative to all other behaviors, more daily time in both LPA and MVPA was beneficially associated with multiple cardiometabolic health markers. The health benefits of MVPA have been well documented in both compositional (Chastin, Palarea-Albaladejo, et al., 2015; Dumuid, Wake, et al., 2019; McGregor et al., 2018; McGregor, Palarea-Albaladejo, Dall, Stamatakis, et al., 2019) and non-compositional studies (Ekelund et al., 2016; Lee et al., 2012; Strain et al., 2020), but the health-enhancing potential of LPA has received less research interest (Chastin et al., 2019). LPA is a more feasible intensity of movement that is accessible regardless of physical fitness, inclination, and opportunity to be more physically active. Evidence has already started to emerge that more time spent in LPA (Chastin et al., 2019; Füzéki et al., 2017), even after accounting for MVPA (Amagasa et al., 2018), could improve cardiometabolic health outcomes. Furthermore, compositional studies have also reported that more LPA could be beneficial for reduced mortality risk, even after accounting for other activity intensities (von Rosen et al., 2020a) and sleep (McGregor, Palarea-Albaladejo, Dall, del Pozo Cruz, et al., 2019). Our findings, while supporting the established physical activity recommendations that encourage MVPA for better health (Bull et al., 2020; Rosenberger et al., 2019), also confirm the findings of recent studies suggesting that LPA may also confer meaningful cardiometabolic health benefits in adults (Chastin et al., 2019; Füzéki et al., 2017), particularly when it replaces sedentary time.

More daily time spent in sedentary and sleeping beyond 7.5 hours/day (both relative to all other behaviors) were both unfavorably associated with cardiometabolic health markers. Our results for sedentary time were in line with previous studies that reported that more sedentary time is associated with poorer cardiometabolic health, although most of those studies failed to account for sleep (Thorp et al., 2011). It is noteworthy that, for the outcomes that displayed U-shaped relationships with sleep duration, more daily time in sedentary was detrimentally associated with outcomes, irrespective of whether participants slept more or less than 7.5 hours/day. This suggests that reduced sedentary time may be beneficial for cardiometabolic health regardless of sleep duration, which is in line with the findings of a previous study using an isotemporal substitution approach (Buman et al., 2013).

In sub-study IV, both Breakers and Shortened sitters were associated with favorable differences in cardiometabolic health markers compared to Couch potatoes, and Breakers had larger favorable differences in cardiometabolic health markers than Shortened sitters. Additionally, Breakers and Shortened sitters were associated with favorable differences in cardiometabolic health markers compared to Couch potatoes even after accounting for potential confounders and sedentary time. These results are collectively in agreement with the existing studies, indicating that, in addition to the total volume of sedentary time, patterns of accumulation of sedentary time may also be related to cardiometabolic health markers and mortality risk in adults (Carson et al., 2014; Diaz et al., 2017). Nevertheless, after accounting for potential confounders and MVPA time, compared with Couch potatoes, Shortened sitters were associated with favorable differences in cardiometabolic health markers, but the associations between Breakers and the same cardiometabolic health markers did not generally reach the significance level. This was not surprising considering that MVPA was a major part of Breakers' physical activity profile, while Shortened sitters were more active through LPA. These results of sub-study IV collectively suggest that more frequent interruptions in sedentary time with any activity intensity from LPA upwards for any duration might be beneficial for adults' cardiometabolic health, which complements the findings of sub-study III indicating that, in addition to MVPA, LPA may also confer meaningful cardiometabolic health benefits in adults.

There were two distinguishable differences in the underlying accumulation patterns of sedentary time and breaks of Breakers and Shortened sitters compared to other groups. First, these two groups were both engaged in relatively fewer uninterrupted sedentary bouts of ≥15–30 minutes and simultaneously included more LPA bouts of different lengths in their sedentary breaks. Second, Breakers also had a relatively lower number of shorter sedentary bouts lasting <15 minutes and, in addition to LPA bouts of different lengths, included more spontaneous MVPA bouts in their sedentary breaks. Currently, little is known about the underlying mechanisms by which prolonged sedentary time may cause detrimental changes to cardiometabolic outcomes (Powell et al., 2018). Hence, epidemiological evidence is continuing to accumulate that frequent sedentary breaks could be beneficial for counteracting such detrimental changes to cardiometabolic health markers in adults caused by sedentary time (Carson et al., 2014; Chastin, Egerton, et al., 2015; Cooper et al., 2012; Healy et al., 2011; Henson et al., 2013). Experimental studies have generally supported this evidence and shown that avoiding prolonged sedentary bouts with light-intensity activities (e.g., walking) could be beneficial for cardiometabolic health in adults (Chastin, Egerton, et al., 2015). For instance, consistent with our results, a recent study in a sample of adults with type-2 diabetes showed that interrupting sedentary time every 15 minutes with light-intensity walking could be beneficial for glucose control (Paing et al., 2019). Our results, while supporting the evolving evidence suggesting that sedentary breaks of at least LPA may improve cardiometabolic health in adults (Chastin, Egerton, et al., 2015), further indicate that in addition to more sedentary breaks it is also important to keep the sedentary bouts shorter than 15–30 minutes for better cardiometabolic health.

6.1.3 Hierarchy of correlates of activity behaviors (V)

In line with the findings of existing studies (Bauman et al., 2012; Chastin et al., 2016; Garcia et al., 2017; Sallis et al., 2006), our data-driven model indicated that activity behavior could be explained by a multilevel hierarchy composed of various factors. Previous studies focusing on understanding the causation of activity behaviors have typically conceptualized the influence of activity behaviors by theoretically combining common sense and well-established evidence, and therefore provided a broad view of activity behavior and its causation for general populations (Bauman et al., 2012; Chastin et al., 2016; Garcia et al., 2017). On the contrary, our data-driven model specified that activity behavior correlates at different levels in each subgroup and may better inform tailored, multilevel intervention allocation and design for our study population. Additionally, the results of association analyses indicated the relative importance of the identified factors, supporting the suggestion that our results can be used to highlight the factors associating with activity behavior in terms of priority.

Most emerged factors in the decision tree model have already been recognized as factors associated with activity behavior in the existing literature, including education level, profession, overall health status, fitness status, and population density (Bauman et al., 2012; Choi et al., 2017; Trost et al., 2002). However, the decision tree model also included some factors that were relatively less studied within the current literature. Such factors include those that were related to personality and temperament, body composition (i.e., lean body mass and skeletal muscle mass), and heart rate recovery, as well as a few psychological and environmental factors (e.g., enjoyment of daily activities and number of road accidents) (Carnethon et al., 2005; Chastin et al., 2016; Choi et al., 2017; O'Donoghue et al., 2016). It seems likely that these factors could have remained underreported (or unexamined) because of the subjective tendency in existing studies toward examining only those factors for which evidence of significant associations (positive or negative) with different activity behaviors has been well understood (Trost et al., 2002).

The less established and previously undiscovered factors found in our study could be candidates for the next generation of correlates (Bauman et al., 2012). These factors were selected by the decision tree to create the final model from a broad list of input (independent) variables. This suggests that the emergent factors in the decision tree model might be relatively more important correlates and likely surrogates for the other previously less established or well-established factors that the decision tree excluded in creating the model, such as behavioral attributes (e.g., alcohol consumption, smoking, etc.) or socioeconomic status (Choi et al., 2017). Yet, one must infer the relative importance of the emergent factors with caution. Due to the data-driven nature of sub-study V, we cannot completely rule out the importance of some of the well-known activity behavior correlates that did not appear in the final model. For instance, the study participants had a narrow age range (46–48 years). This might explain why some of already well-established activity behavior correlates, for example age and gender, did not emerge in the final model (Bauman et al., 2012; Choi et al., 2017; Trost et al., 2002).

6.1.4 Strengths and limitations

The strength of this study includes identifying and reviewing a comprehensive list of studies calibrating and validating accelerometer-based activity monitors using machine learning approaches. Another strength is independent validation of machine learning models developed for classifying activities in terms of intensity across different heterogenous datasets. Using a wide population-based sample with wide background information available for the participants, and device-based assessment of daily activities is also a strength.

This study has some limitations. In the systematic review, we did not focus on the predictive accuracy of ML-based modeling approaches in relation to activity types (e.g., sitting, standing, walking, etc.) together with accelerometer placement, since in many applications (e.g., observational studies) a wide range of activities under free-living conditions is of interest to provide good estimates of total sedentary time and physical activities (Bassett Jr et al., 2012; Wijndaele et al., 2015). If measuring certain types of activities is of interest, a specific accelerometer placement might provide substantially better results. Another limitation was that the comparison between the predictive accuracy of ML modeling approaches and regression- and cut-point-based methods remained outside the scope of our study. Instead, we assumed that ML models are superior to the traditional statistical procedures. This is a legitimate assumption, considering that previous review studies have emphasized the superiority of ML models (Bassett Jr et al., 2012; Strath et al., 2012), but future studies with direct comparison between different techniques are required to better understand the magnitude of the superiority of ML models.

In sub-study II, limited meta-data from the open-access datasets was preventive for providing conclusive information about which sources of heterogeneities among the dataset played a more important role in the overall performance reduction of within-dataset-validated models. Additionally, none of the datasets were collected under truly free-living conditions, and therefore further studies are needed to test how the models would perform with free-living data. Under free-living conditions MVPA generally make up approximately 3%–5% of the 24-hour movement behaviors (Chaput et al., 2014), but in our datasets MVPA consisted approximately 30%–60% of the datasets. This might also limit the ability of our models for free-living data. Using direct observation and the Compendium of Physical Activity as criterion measures for defining activity intensities rather than a gold standard (i.e., indirect calorimetry) may also be a limitation. Hence, direct observation appear to be feasible and valid (Cox et al., 2020; Lyden, Petruski, et al., 2014), and a commonly used criterion in many analytical studies calibrating and validating accelerometer-based activity monitors for estimation of activity intensities and types (de Almeida Mendes et al., 2018). Given the findings of sub-study I, we decided to select ANN as the modeling approach in sub-study II because it has been used frequently in previous studies, and it has been reported to be highly accurate in predicting both activity type and energy expenditure from accelerometer data in different age groups (de Almeida Mendes et al., 2018; Montoye et al., 2016b; Trost et al., 2012). Future studies should consider exploring other commonly used ML approaches as well as testing their generalization capabilities outside the datasets used for model development.

The studies (III–V) conducted in the NFBC1966 cohort data were cross-sectional. Inferences about the temporality of associations are therefore limited, and causality cannot be determined. Due to the birth cohort settings, the study population was homogenous in terms of age and ethnicity. Although beneficial with respect to reducing the potential for confounding the observed associations, this may limit the generalizability of the results of these three sub-studies to more diverse populations. Also, more than 85% of the original cohort members were alive in Finland at the time of the latest follow-up, but less than 50% participated and wore the accelerometers. It is possible that those were the healthier and more active ones. This might have induced selection bias and further limited the generalizability of the results. Although in sub-studies III–V accelerometer-based measurement of daily activities was captured over a relatively long timeframe, from which movement behaviors were estimated using previously validated methods (Jauho et al., 2015; Leinonen et al., 2016; Vähä-Ypyä et al., 2018, 2015), there were some unavoidable discrepancies among these three sub-studies for processing the accelerometry data and defining the intensity categories. This further highlights the urgent need for a standardized data processing method and studies to consider the agreement between different accelerometers with respect to the wear location until consensus and standardized definitions are achieved.

In sub-study III, sleep duration was self-reported, and therefore was probably measured with less accuracy than the activity intensities. However, it is unlikely that our results would be different with device-based measurement of sleep duration, considering that the differences between self-reported and device-based sleep duration in middle-aged adults could be small (Lauderdale et al., 2006). In sub-study IV, we noted that the concept of Breakers, Prolonged sitters, and Couch potatoes has been theorized and investigated previously (Bakrania et al., 2015; Owen et al., 2010; Tremblay et al., 2017), but Shortened sitters is rather a novel activity profile that was found in our study. Similar studies should be performed in other populations to determine whether similar profiles and associations to those identified in our study exist in other populations. Inclusion of only activity intensities during awake time in sub-study IV is also a limitation. Further studies with 24-hour accelerometry and characterization of sleep patterns are needed to warrant our findings. In sub-study V, the binary categorization of participants (active or inactive) was the main limitation. In sub-studies II–V, we categorized moderate-intensity and vigorous-intensity physical activities together as MVPA. In sub-study III, we further differentiated sitting/lying from standing still and categorized standing still as LPA. This could be a limitation because standing still as a posture as well as moderate-intensity physical activity and vigorous-intensity physical activity could potentially have distinct health effects (Powell et al., 2019; Shiroma et al., 2014).

7 Conclusions

The present study indicated that the variation among the reported overall predictive accuracy of machine learning models developed for activity recognition is minimally related to accelerometer placement, but the generalization capabilities of machine learning models remains a concern. Our proposed method to integrate various data sources in training sets was found to be a viable approach for training more robust models capable of classifying activities by intensity in another population monitored with a different accelerometer. The study also suggests that adults may gain cardiometabolic health benefits through LPA, particularly when it is replaced by time spent in sedentary activity. Based on the aims of the study, it can be concluded that:

- Machine learning approaches offer opportunities for predicting activity types, categories, and intensities with comparable overall predictive accuracies irrespective of accelerometer placement, although it remains unknown whether these minimized variations would remain under free-living conditions.
- Integrating heterogeneous datasets using hip or wrist data in training sets is a viable approach for enhancing the generalization performance of the ANN models developed for predicting activity intensities from raw acceleration data.
- More daily time in MVPA at the expense of any other movement intensity or sleep could be the most time-efficient change in the 24-hour movement behavior composition for improving cardiometabolic health in mid-adulthood. Alternatively, more daily time in LPA at the expense of sedentary time (or sleep) could also be beneficial for cardiometabolic health, but to a lesser extent compared to more time in MVPA.
- Avoiding uninterrupted sedentary bouts of longer than 15–30 minutes by breaking them frequently with short LPA bouts may be beneficial for cardiometabolic health in middle-aged adults. In addition to LPA bouts, further inclusion of spontaneous MVPA bouts in sedentary breaks may confer additional cardiometabolic health benefits for adults.
- The created data-driven hierarchy consisted of factors of relative importance from different domains, and could be used for multilevel intervention allocation and design for inactive people in the study population. Additionally, the novel and less-established set of factors that were methodologically discovered can be a basis for additional hypothesis testing in activity behavior correlates research.

References

Ainsworth, B. E., Haskell, W. L., Herrmann, S. D., Meckes, N., Bassett, D. R., Tudor-Locke, C., … Leon, A. S. (2011). 2011 Compendium of Physical Activities. *Medicine & Science in Sports & Exercise*, *43*(8), 1575–1581. doi: 10.1249/MSS.0b013e31821ece12

Aittasalo, M., Vähä-Ypyä, H., Vasankari, T., Husu, P., Jussila, A.-M., & Sievänen, H. (2015). 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 Science, Medicine and Rehabilitation*, *7*(18). doi: 10.1186/s13102-015-0010-0

Alberti, K. G. M. M., & Zimmet, P. Z. (1998). Definition, diagnosis and classification of diabetes mellitus and its complications. Part 1: diagnosis and classification of diabetes mellitus. Provisional report of a WHO consultation. *Diabetic Medicine*, *15*(7), 539–553. doi: 10.1002/(SICI)1096-9136(199807)15:7<539::AID-DIA668>3.0.CO;2-S

Alberto, F.-P., Nathanael, M., Mathew, B., & Ainsworth, B. E. (2017). Wearable monitors criterion validity for energy expenditure in sedentary and light activities. *Journal of Sport and Health Science*, *6*(1), 103–110. doi: 10.1016/j.jshs.2016.10.005

Altun, K., Barshan, B., & Tunçel, O. (2010). Comparative study on classifying human activities with miniature inertial and magnetic sensors. *Pattern Recognition*, *43*(10), 3605–3620. doi: 10.1016/j.patcog.2010.04.019

Amagasa, S., Machida, M., Fukushima, N., Kikuchi, H., Takamiya, T., Odagiri, Y., & Inoue, S. (2018). Is objectively measured light-intensity physical activity associated with health outcomes after adjustment for moderate-to-vigorous physical activity in adults? a systematic review. *International Journal of Behavioral Nutrition and Physical Activity*, *15*(65). doi: 10.1186/s12966-018-0695-z

Atallah, L., Lo, B., King, R., & Yang, G.-Z. (2011). Sensor positioning for activity recognition using wearable accelerometers. *IEEE Transactions on Biomedical Circuits and Systems*, *5*(4), 320–329. doi: 10.1109/TBCAS.2011.2160540

Atkinson, G., & Davenne, D. (2007). Relationships between sleep, physical activity and human health. *Physiology & Behavior*, *90*(2–3), 229–235. doi: 10.1016/j.physbeh.2006.09.015

Bakrania, K., Edwardson, C. L., Bodicoat, D. H., Esliger, D. W., Gill, J. M. R., Kazi, A., … Davies, M. (2015). Associations of mutually exclusive categories of physical activity and sedentary time with markers of cardiometabolic health in English adults: a cross-sectional analysis of the Health Survey for England. *BMC Public Health*, *16*(25). doi: 10.1186/s12889-016-2694-9

Basner, M., Fomberstein, K. M., Razavi, F. M., Banks, S., William, J. H., Rosa, R. R., & Dinges, D. F. (2007). American time use survey: sleep time and its relationship to waking activities. *Sleep*, *30*(9), 1085–1095. doi: 10.1093/sleep/30.9.1085

Bassett Jr, D. R., Rowlands, A. V, & Trost, S. G. (2012). Calibration and validation of wearable monitors. *Medicine and Science in Sports and Exercise*, *44*(1 Suppl 1), S32–S38. doi: 10.1249/MSS.0b013e3182399cf7

Bastian, T., Maire, A., Dugas, J., Ataya, A., Villars, C., Gris, F., … Simon, C. (2015). Automatic identification of physical activity types and sedentary behaviors from triaxial accelerometer: laboratory-based calibrations are not enough. *Journal of Applied Physiology*, *118*(6), 716–722. doi: 10.1152/japplphysiol.01189.2013

Bauman, A. E., Reis, R. S., Sallis, J. F., Wells, J. C., Loos, R. J. F., & Martin, B. W. (2012). Correlates of physical activity: why are some people physically active and others not? *The Lancet*, *380*(9838), 258–271. doi: 10.1016/S0140-6736(12)60735-1

Beard, J. R., Officer, A., De Carvalho, I. A., Sadana, R., Pot, A. M., Michel, J.-P., … Chatterji, S. (2016). The World report on ageing and health: a policy framework for healthy ageing. *The Lancet*, *387*(10033), 2145–2154. doi: 10.1016/S0140-6736(15)00516-4

Bellazzi, R., Diomidous, M., Sarkar, I. N., Takabayashi, K., Ziegler, A., & McCray, A. T. (2011). Data analysis and data mining: current issues in biomedical informatics. *Methods of Information in Medicine*, *50*(6), 536–544. doi: 10.3414/ME11-06-0002

Bellazzi, R., & Zupan, B. (2008). Predictive data mining in clinical medicine: current issues and guidelines. *International Journal of Medical Informatics*, *77*(2), 81–97. doi: 10.1016/j.ijmedinf.2006.11.006

Berendsen, B. A., Hendriks, M. R., Meijer, K., Plasqui, G., Schaper, N. C., & Savelberg, H. H. (2014). Which activity monitor to use? validity, reproducibility and user friendliness of three activity monitors. *BMC Public Health*, *14*(749). doi: 10.1186/1471-2458-14-749

Biering-Sørensen, F. (1984). Physical measurements as risk indicators for low-back trouble over a one-year period. *Spine*, *9*(2), 106–119. doi: 10.1097/00007632-198403000-00002

Blockeel, H., & Struyf, J. (2002). Efficient algorithms for decision tree cross-validation. *Journal of Machine Learning Research*, *3*, 621–650.

Brocklebank, L. A., Falconer, C. L., Page, A. S., Perry, R., & Cooper, A. R. (2015). Accelerometer-measured sedentary time and cardiometabolic biomarkers: a systematic review. *Preventive Medicine*, *76*, 92–102. doi: 10.1016/j.ypmed.2015.04.013

Bromley, L. E., Booth 3rd, J. N., Kilkus, J. M., Imperial, J. G., & Penev, P. D. (2012). Sleep restriction decreases the physical activity of adults at risk for type 2 diabetes. *Sleep*, *35*(7), 977–984. doi: 10.5665/sleep.1964

Buck, C., Loyen, A., Foraita, R., Van Cauwenberg, J., De Craemer, M., Mac Donncha, C., … Chastin, S. F. M. (2019). Factors influencing sedentary behaviour: a system based analysis using Bayesian networks within DEDIPAC. *PloS One*, *14*, e0211546. doi: 10.1371/journal.pone.0211546

Bull, F. C., Al-Ansari, S. S., Biddle, S., Borodulin, K., Buman, M. P., Cardon, G., … Willumsen, J. F. (2020). World Health Organization 2020 guidelines on physical activity and sedentary behaviour. *British Journal of Sports Medicine*, *54*(24), 1451–1462. doi: 10.1136/bjsports-2020-102955

Buman, M. P., Mullane, S. L., Toledo, M. J., Rydell, S. A., Gaesser, G. A., Crespo, N. C., … Pereira, M. A. (2017). An intervention to reduce sitting and increase light-intensity physical activity at work: design and rationale of the 'stand & move at work'group randomized trial. *Contemporary Clinical Trials*, *53*, 11–19. doi: 10.1016/j.cct.2016.12.008

Buman, M. P., Winkler, E. A. H., Kurka, J. M., Hekler, E. B., Baldwin, C. M., Owen, N., … Gardiner, P. A. (2013). Reallocating time to sleep, sedentary behaviors, or active behaviors: associations with cardiovascular disease risk biomarkers, NHANES 2005-2006. *American Journal of Epidemiology*, *179*(3), 323–334. doi: 10.1093/aje/kwt292

Buman, M. P., & Youngstedt, S. D. (2015). Physical activity, sleep, and biobehavioral synergies for health. In *Sleep and Affect* (pp. 321–337). doi: 10.1016/B978-0-12-417188-6.00015-3

Butte, N. F., Watson, K. B., Ridley, K., Zakeri, I. F., Mcmurray, R. G., Pfeiffer, K. A., … Fulton, J. E. (2018). A youth compendium of physical activities: activity codes and metabolic intensities. *Medicine & Science in Sports & Exercise*, *50*(2), 246–256. doi: 10.1249/MSS.0000000000001430

Carnethon, M. R., Jacobs, J. D. R., Sidney, S., Sternfeld, B., Gidding, S. S., Shoushtari, C., & Liu, K. (2005). A longitudinal study of physical activity and heart rate recovery: CARDIA, 1987-1993. *Medicine and Science in Sports and Exercise*, *37*(4), 606–612. doi: 10.1249/01.mss.0000158190.56061.32

Carr, L. J., & Mahar, M. T. (2011). Accuracy of intensity and inclinometer output of three activity monitors for identification of sedentary behavior and light-intensity activity. *Journal of Obesity*, *2012*(460271). doi: 10.1155/2012/460271

Carson, V., Wong, S. L., Winkler, E., Healy, G. N., Colley, R. C., & Tremblay, M. S. (2014). Patterns of sedentary time and cardiometabolic risk among Canadian adults. *Preventive Medicine*, *65*, 23–27. doi: 10.1016/j.ypmed.2014.04.005

Celis-Morales, C. A., Perez-Bravo, F., Ibanez, L., Salas, C., Bailey, M. E. S., & Gill, J. M. R. (2012). Objective vs. self-reported physical activity and sedentary time: effects of measurement method on relationships with risk biomarkers. *PloS One*, *7*(5), e36345. doi: 10.1371/journal.pone.0036345

Chaput, J.-P., Carson, V., Gray, C., & Tremblay, M. (2014). Importance of all movement behaviors in a 24 hour period for overall health. *International Journal of Environmental Research and Public Health*, *11*(12), 12575–12581. doi: 10.3390/ijerph111212575

Chastin, S. F. M., De Craemer, M., De Cocker, K., Powell, L., Van Cauwenberg, J., Dall, P., … Stamatakis, E. (2019). How does light-intensity physical activity associate with adult cardiometabolic health and mortality? systematic review with meta-analysis of experimental and observational studies. *British Journal of Sports Medicine*, *53*(6), 370–376. doi: 10.1136/bjsports-2017-097563

Chastin, S. F. M., De Craemer, M., Lien, N., Bernaards, C., Buck, C., Oppert, J.-M., … Cardon, G. (2016). The SOS-framework (Systems of Sedentary behaviours): an international transdisciplinary consensus framework for the study of determinants, research priorities and policy on sedentary behaviour across the life course: a DEDIPAC-study. *International Journal of Behavioral Nutrition and Physical Activity*, *13*(83). doi: 10.1186/s12966-016-0409-3

Chastin, S. F. M., Egerton, T., Leask, C., & Stamatakis, E. (2015). Meta-analysis of the relationship between breaks in sedentary behavior and cardiometabolic health. *Obesity*, *23*(9), 1800–1810. doi: 10.1002/oby.21180

Chastin, S. F. M., Palarea-Albaladejo, J., Dontje, M. L., & Skelton, D. A. (2015). 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*, *10*(10), e0139984. doi: 10.1371/journal.pone.0139984

Chatterji, S., Byles, J., Cutler, D., Seeman, T., & Verdes, E. (2015). Health, functioning, and disability in older adults—present status and future implications. *The Lancet*, *385*(9967), 563–575. doi: 10.1016/S0140-6736(14)61462-8

Chen, C. L. P., & Zhang, C.-Y. (2014). Data-intensive applications, challenges, techniques and technologies: a survey on Big Data. *Information Sciences*, *275*, 314–347. doi: 10.1016/j.ins.2014.01.015

Choi, J., Lee, M., Lee, J., Kang, D., & Choi, J.-Y. (2017). Correlates associated with participation in physical activity among adults: a systematic review of reviews and update. *BMC Public Health*, *17*(356). doi: 10.1186/s12889-017-4255-2

Choi, L., Liu, Z., Matthews, C. E., & Buchowski, M. S. (2011). Validation of accelerometer wear and nonwear time classification algorithm. *Medicine and Science in Sports and Exercise*, *43*(2), 357–364. doi: 10.1249/MSS.0b013e3181ed61a3

Cleland, I., Kikhia, B., Nugent, C., Boytsov, A., Hallberg, J., Synnes, K., … Finlay, D. (2013). Optimal placement of accelerometers for the detection of everyday activities. *Sensors*, *13*(7), 9183–9200. doi: 10.3390/s130709183

Cleland, V. J., Schmidt, M. D., Dwyer, T., & Venn, A. J. (2008). Television viewing and abdominal obesity in young adults: is the association mediated by food and beverage consumption during viewing time or reduced leisure-time physical activity? *The American Journal of Clinical Nutrition*, *87*(5), 1148–1155. doi: 10.1093/ajcn/87.5.1148

Cloninger, C. R., Przybeck, T. R., Svrakic, D. M., & Wetzel, R. D. (1994). *The Temperament and Character Inventory (TCI): A guide to its development and use*. Center for Psychobiology of Personality, Washington University.

Cohen, J. (1968). Weighted kappa: Nominal scale agreement provision for scaled disagreement or partial credit. *Psychological Bulletin*, *70*(4), 213–220. doi: 10.1037/h0026256

Cooper, A. R., Sebire, S., Montgomery, A. A., Peters, T. J., Sharp, D. J., Jackson, N., … Andrews, R. C. (2012). Sedentary time, breaks in sedentary time and metabolic variables in people with newly diagnosed type 2 diabetes. *Diabetologia*, *55*(3), 589–599. doi: 10.1007/s00125-011-2408-x

Cox, M. F., Petrucci, G. J., Marcotte, R. T., Masteller, B. R., Staudenmayer, J., Freedson, P. S., & Sirard, J. R. (2020). A novel video-based direct observation system for assessing physical activity and sedentary behavior in children and young adults. *Journal for the Measurement of Physical Behaviour*, *3*(1), 50–57. doi: 10.1123/jmpb.2019-0015

Crouter, S. E., DellaValle, D. M., Haas, J. D., Frongillo, E. A., & Bassett, D. R. (2013). Validity of ActiGraph 2-regression model, Matthews cut-points, and NHANES cut-points for assessing free-living physical activity. *Journal of Physical Activity and Health*, *10*(4), 504–514. doi: 10.1123/jpah.10.4.504

Daskalopoulou, C., Stubbs, B., Kralj, C., Koukounari, A., Prince, M., & Prina, A. M. (2017). Physical activity and healthy ageing: a systematic review and meta-analysis of longitudinal cohort studies. *Ageing Research Reviews*, *38*, 6–17. doi: 10.1016/j.arr.2017.06.003

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

de Rezende, L. F. M., Lopes, M. R., Rey-López, J. P., Matsudo, V. K. R., & do Carmo Luiz, O. (2014). Sedentary behavior and health outcomes: an overview of systematic reviews. *PloS One*, *9*(8), e105620. doi: 10.1371/journal.pone.0105620

del Pozo Cruz, B., McGregor, D. E., del Pozo Cruz, J., Buman, M. P., Palarea-Albaladejo, J., Alfonso-Rosa, R. M., & Chastin, S. F. M. (2020). Integrating sleep, physical activity, and diet quality to estimate all-cause mortality risk: a combined compositional clustering and survival analysis of the NHANES 2005-2006 cycle. *American Journal of Epidemiology*, *189*(10), 1057–1064. doi: 10.1093/aje/kwaa057

Diaz, K. M., Howard, V. J., Hutto, B., Colabianchi, N., Vena, J. E., Safford, M. M., … Hooker, S. P. (2017). Patterns of sedentary behavior and mortality in US middle-aged and older adults: a national cohort study. *Annals of Internal Medicine*, *167*(7), 465–475. doi: 10.7326/M17-0212

Ding, D., Varela, A. R., Bauman, A. E., Ekelund, U., Lee, I.-M., Heath, G., … Pratt, M. (2020). Towards better evidence-informed global action: lessons learnt from the Lancet series and recent developments in physical activity and public health. *British Journal of Sports Medicine*, *54*(8), 462–468. doi: 10.1136/bjsports-2019-101001

DiPietro, L., Al-Ansari, S. S., Biddle, S. J. H., Borodulin, K., Bull, F. C., Buman, M. P., … Willumsen, J. F. (2020). Advancing the global physical activity agenda: recommendations for future research by the 2020 WHO physical activity and sedentary behavior guidelines development group. *International Journal of Behavioral Nutrition and Physical Activity*, *17*(143). doi: 10.1186/s12966-020-01042-2

Dumuid, D., Pedišić, Ž., Stanford, T. E., Martín-Fernández, J.-A., Hron, K., Maher, C. A., … Olds, T. (2019). The compositional isotemporal substitution model: a method for estimating changes in a health outcome for reallocation of time between sleep, physical activity and sedentary behaviour. *Statistical Methods in Medical Research*, *28*(3), 846–857. doi: 10.1177/0962280217737805

Dumuid, D., Stanford, T. E., Martin-Fernández, J.-A., Pedišić, Ž., Maher, C. A., Lewis, L. K., … Olds, T. (2018). Compositional data analysis for physical activity, sedentary time and sleep research. *Statistical Methods in Medical Research*, *27*(12), 3726–3738. doi: 10.1177/0962280217710835

Dumuid, D., Wake, M., Clifford, S., Burgner, D., Carlin, J. B., Mensah, F. K., … Olds, T. (2019). The association of the body composition of children with 24-hour activity composition. *The Journal of Pediatrics*, *208*, 43–49. doi: 10.1016/j.jpeds.2018.12.030

Dunstan, D. W., Kingwell, B. A., Larsen, R., Healy, G. N., Cerin, E., Hamilton, M. T., … Owen, N. (2012). Breaking up prolonged sitting reduces postprandial glucose and insulin responses. *Diabetes Care*, *35*(5), 976–983. doi: 10.2337/dc11-1931

Dyrstad, S. M., Hansen, B. H., Holme, I. M., & Anderssen, S. A. (2014). Comparison of self-reported versus accelerometer-measured physical activity. *Medicine & Science in Sports & Exercise*, *46*(1), 99–106. doi: 10.1249/MSS.0b013e3182a0595f

Edwardson, C. L., Winkler, E. A. H., Bodicoat, D. H., Yates, T., Davies, M. J., Dunstan, D. W., & Healy, G. N. (2017). Considerations when using the activPAL monitor in field-based research with adult populations. *Journal of Sport and Health Science*, *6*(2), 162–178. doi: 10.1016/j.jshs.2016.02.002

Ekblom-Bak, E., Ekblom, Ö., Bergström, G., & Börjesson, M. (2016). Isotemporal substitution of sedentary time by physical activity of different intensities and bout lengths, and its associations with metabolic risk. *European Journal of Preventive Cardiology*, *23*(9), 967–974. doi: 10.1177/2047487315619734

Ekblom-Bak, E., Olsson, G., Ekblom, Ö., Ekblom, B., Bergström, G., & Börjesson, M. (2015). The daily movement pattern and fulfilment of physical activity recommendations in Swedish middle-aged adults: the SCAPIS pilot study. *PloS One*, *10*(5), e0126336. doi: 10.1371/journal.pone.0126336

Ekelund, U., Brage, S., Griffin, S. J., & Wareham, N. J. (2009). Objectively measured moderate-and vigorous-intensity physical activity but not sedentary time predicts insulin resistance in high-risk individuals. *Diabetes Care*, *32*(6), 1081–1086. doi: 10.2337/dc08-1895

Ekelund, U., Steene-Johannessen, J., Brown, W. J., Fagerland, M. W., Owen, N., Powell, K. E., … Lee, I.-M. (2016). 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. *The Lancet*, *388*(10051), 1302–1310. doi: 10.1016/S0140-6736(16)30370-1

Ekelund, U., Tarp, J., Steene-Johannessen, J., Hansen, B. H., Jefferis, B., Fagerland, M. W., … Shiroma, E. (2019). Dose-response associations between accelerometry measured physical activity and sedentary time and all cause mortality: systematic review and harmonised meta-analysis. *BMJ*, *366*, l4570. doi: 10.1136/bmj.l4570

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

Ellis, K., Kerr, J., Godbole, S., Staudenmayer, J., & Lanckriet, G. (2016). Hip and wrist accelerometer algorithms for free-living behavior classification. *Medicine and Science in Sports and Exercise*, *48*(5), 933–940. doi: 10.1249/MSS.0000000000000840

Fishman, E. I., Steeves, J. A., Zipunnikov, V., Koster, A., Berrigan, D., Harris, T. A., & Murphy, R. (2016). Association between objectively measured physical activity and mortality in NHANES. *Medicine and Science in Sports and Exercise*, *48*(7), 1303–

1311. doi: 10.1249/MSS.0000000000000885

Freedson, P. S., Bowles, H. R., Troiano, R., & Haskell, W. (2012). Assessment of physical activity using wearable monitors: recommendations for monitor calibration and use in the field. *Medicine and Science in Sports and Exercise*, *44*(1 Suppl 1), S1–S4. doi: 10.1249/MSS.0b013e3182399b7e

Freedson, P. S., Lyden, K., Kozey-Keadle, S., & Staudenmayer, J. (2011). Evaluation of artificial neural network algorithms for predicting METs and activity type from accelerometer data: validation on an independent sample. *Journal of Applied Physiology*, *111*(6), 1804–1812. doi: 10.1152/japplphysiol.00309.2011

Füzéki, E., Engeroff, T., & Banzer, W. (2017). Health benefits of light-intensity physical activity: a systematic review of accelerometer data of the National Health and Nutrition Examination Survey (NHANES). *Sports Medicine*, *47*(9), 1769–1793. doi: 10.1007/s40279-017-0724-0

Garcia, L. M. T., Roux, A. V. D., Martins, A. C. R., Yang, Y., & Florindo, A. A. (2017). Development of a dynamic framework to explain population patterns of leisure-time physical activity through agent-based modeling. *International Journal of Behavioral Nutrition and Physical Activity*, *14*(111). doi: 10.1186/s12966-017-0553-4

Glazer, N. L., Lyass, A., Esliger, D. W., Blease, S. J., Freedson, P. S., Massaro, J. M., … Vasan, R. S. (2013). Sustained and shorter bouts of physical activity are related to cardiovascular health. *Medicine and Science in Sports and Exercise*, *45*(1), 109–115. doi: 10.1249/MSS.0b013e31826beae5

Grgic, J., Dumuid, D., Bengoechea, E. G., Shrestha, N., Bauman, A., Olds, T., & Pedisic, Z. (2018). Health outcomes associated with reallocations of time between sleep, sedentary behaviour, and physical activity: a systematic scoping review of isotemporal substitution studies. *International Journal of Behavioral Nutrition and Physical Activity*, *15*(69). doi: 10.1186/s12966-018-0691-3

Gupta, N., Hallman, D. M., Dumuid, D., Vij, A., Rasmussen, C. L., Jørgensen, M. B., & Holtermann, A. (2020). Movement behavior profiles and obesity: a latent profile analysis of 24-h time-use composition among Danish workers. *International Journal of Obesity*, *44*(2), 409–417. doi: 10.1038/s41366-019-0419-8

Guthold, R., Stevens, G. A., Riley, L. M., & Bull, F. C. (2018). Worldwide trends in insufficient physical activity from 2001 to 2016: a pooled analysis of 358 population-based surveys with 1·9 million participants. *The Lancet Global Health*, *6*(10), e1077–e1086. doi: 10.1016/S2214-109X(18)30357-7

Hallal, P. C., Andersen, L. B., Bull, F. C., Guthold, R., Haskell, W., & Ekelund, U. (2012). Global physical activity levels: surveillance progress, pitfalls, and prospects. *The Lancet*, *380*(9838), 247–257. doi: 10.1016/S0140-6736(12)60646-1

Hamer, M., Stamatakis, E., Chastin, S., Pearson, N., Brown, M., Gilbert, E., & Sullivan, A. (2020). Feasibility of measuring sedentary time using data from a thigh-worn accelerometer: the 1970 British Cohort Study. *American Journal of Epidemiology*, *189*(9), 963–971. doi: 10.1093/aje/kwaa047

Hart, T. L., Ainsworth, B. E., & Tudor-Locke, C. (2011). Objective and subjective measures of sedentary behavior and physical activity. *Medicine and Science in Sports and*

Exercise, *43*(3), 449–456. doi: 10.1249/MSS.0b013e3181ef5a93

Healy, G. N., Dunstan, D. W., Salmon, J., Cerin, E., Shaw, J. E., Zimmet, P. Z., & Owen, N. (2008). Breaks in sedentary time: beneficial associations with metabolic risk. *Diabetes Care*, *31*(4), 661–666. doi: 10.2337/dc07-2046

Healy, G. N., Matthews, C. E., Dunstan, D. W., Winkler, E. A. H., & Owen, N. (2011). Sedentary time and cardio-metabolic biomarkers in US adults: NHANES 2003-06. *European Heart Journal*, *32*(5), 590–597. doi: 10.1093/eurheartj/ehq451

Healy, G. N., Winkler, E. A. H., Owen, N., Anuradha, S., & Dunstan, D. W. (2015). Replacing sitting time with standing or stepping: associations with cardiometabolic risk biomarkers. *European Heart Journal*, *36*(39), 2643–2649. doi: 10.1093/eurheartj/ehv308

Helmerhorst, H. J. F., Brage, S., Warren, J., Besson, H., & Ekelund, U. (2012). A systematic review of reliability and objective criterion-related validity of physical activity questionnaires. *International Journal of Behavioral Nutrition and Physical Activity*, *9*(103). doi: 10.1186/1479-5868-9-103

Henson, J., Yates, T., Biddle, S. J. H., Edwardson, C. L., Khunti, K., Wilmot, E. G., … Davies, M. J. (2013). Associations of objectively measured sedentary behaviour and physical activity with markers of cardiometabolic health. *Diabetologia*, *56*(5), 1012–1020. doi: 10.1007/s00125-013-2845-9

Hirshkowitz, M., Whiton, K., Albert, S. M., Alessi, C., Bruni, O., DonCarlos, L., … Hillard, P. J. A. (2015). National Sleep Foundation's sleep time duration recommendations: methodology and results summary. *Sleep Health*, *1*(1), 40–43. doi: 10.1016/j.sleh.2014.12.010

Howard, B., Winkler, E. A., Sethi, P., Carson, V., Ridgers, N. D., Salmon, J. O., … Dunstan, D. W. (2015). Associations of low- and high-intensity light activity with cardiometabolic biomarkers. *Medicine and Science in Sports and Exercise*, *47*(10), 2093–2101. doi: 10.1249/MSS.0000000000000631

Janssen, X., Basterfield, L., Parkinson, K. N., Pearce, M. S., Reilly, J. K., Adamson, A. J., & Reilly, J. J. (2015). Objective measurement of sedentary behavior: impact of non-wear time rules on changes in sedentary time. *BMC Public Health*, *15*(504). doi: 10.1186/s12889-015-1847-6

Janssen, X., & Cliff, D. P. (2015). Issues related to measuring and interpreting objectively measured sedentary behavior data. *Measurement in Physical Education and Exercise Science*, *19*(3), 116–124. doi: 10.1080/1091367X.2015.1045908

Jauho, A.-M., Pyky, R., Ahola, R., Kangas, M., Virtanen, P., Korpelainen, R., & Jämsä, T. (2015). Effect of wrist-worn activity monitor feedback on physical activity behavior: a randomized controlled trial in Finnish young men. *Preventive Medicine Reports*, *2*, 628–634. doi: 10.1016/j.pmedr.2015.07.005

Jensky-Squires, N. E., Dieli-Conwright, C. M., Rossuello, A., Erceg, D. N., McCauley, S., & Schroeder, E. T. (2008). Validity and reliability of body composition analysers in children and adults. *British Journal of Nutrition*, *100*(4), 859–865. doi: 10.1017/S0007114508925460

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

Kanagasabai, T., & Chaput, J.-P. (2017). Sleep duration and the associated cardiometabolic risk scores in adults. *Sleep Health*, *3*(3), 195–203. doi: 10.1016/j.sleh.2017.03.006

Kanungo, T., Mount, D. M., Netanyahu, N. S., Piatko, C. D., Silverman, R., & Wu, A. Y. (2002). An efficient k-means clustering algorithm: analysis and implementation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *24*(7), 881–892. doi: 10.1109/TPAMI.2002.1017616

Kärmeniemi, M., Lankila, T., Ikäheimo, T., Puhakka, S., Niemelä, M., Jämsä, T., … Korpelainen, R. (2019). Residential relocation trajectories and neighborhood density, mixed land use and access networks as predictors of walking and bicycling in the Northern Finland Birth Cohort 1966. *International Journal of Behavioral Nutrition and Physical Activity*, *16*(88). doi: 10.1186/s12966-019-0856-8

Kass, G. V. (1980). An exploratory technique for investigating large quantities of categorical data. *Journal of the Royal Statistical Society Series C, Royal Statistical Society*, *29*(2), 119–127. doi: 10.2307/2986296

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

Khaw, K.-T., Wareham, N., Bingham, S., Welch, A., Luben, R., & Day, N. (2008). Combined impact of health behaviours and mortality in men and women: the EPIC-Norfolk prospective population study. *PLoS Med*, *5*(1), e12. doi: journal.pmed.0050012

Kim, Y., Beets, M. W., & Welk, G. J. (2012). Everything you wanted to know about selecting the "right" Actigraph accelerometer cut-points for youth, but…: a systematic review. *Journal of Science and Medicine in Sport*, *15*(4), 311–321. doi: 10.1016/j.jsams.2011.12.001

Kinnunen, H., Häkkinen, K., Schumann, M., Karavirta, L., Westerterp, K. R., & Kyröläinen, H. (2019). Training-induced changes in daily energy expenditure: methodological evaluation using wrist-worn accelerometer, heart rate monitor, and doubly labeled water technique. *PloS One*, *14*(7), e0219563. doi: 10.1371/journal.pone.0219563

Kivimäki, M., Singh-Manoux, A., Pentti, J., Sabia, S., Nyberg, S. T., Alfredsson, L., … Jokela, M. (2019). Physical inactivity, cardiometabolic disease, and risk of dementia: an individual-participant meta-analysis. *BMJ*, *365*(l1495). doi: 10.1136/bmj.l1495

Kiviniemi, A. M., Perkiömäki, N., Auvinen, J., Niemelä, M., Tammelin, T., Puukka, K., … Korpelainen, R. (2017). Fitness, fatness, physical activity, and autonomic function in midlife. *Medicine & Science in Sports & Exercise*, *49*(12), 2459–2468. doi: 10.1249/MSS.0000000000001387

Kmet, L. M., Lee, R. C., & Cook, L. S. (2004). Standard quality assessment criteria for evaluating primary research papers from a variety of fields. Edmonton: Alberta Heritage Foundation for Medical Research (AHFMR). *AHFMR - HTA Initiative #13*. doi: 10.7939/R37M04F16

Knutson, K. L. (2010). Sleep duration and cardiometabolic risk: a review of the epidemiologic evidence. *Best Practice & Research Clinical Endocrinology & Metabolism*, *24*(5), 731–743. doi: 10.1016/j.beem.2010.07.001

Kodinariya, T. M., & Makwana, P. R. (2013). Review on determining number of cluster in k-means clustering. *International Journal of Advance Research in Computer Science and Management Studies*, *1*(6), 90–95.

Kohl 3rd, H. W., Craig, C. L., Lambert, E. V., Inoue, S., Alkandari, J. R., Leetongin, G., & Kahlmeier, S. (2012). The pandemic of physical inactivity: global action for public health. *The Lancet*, *380*(9838), 294–305. doi: 10.1016/S0140-6736(12)60898-8

Kozey-Keadle, S., Libertine, A., Lyden, K., Staudenmayer, J., & Freedson, P. S. (2011). Validation of wearable monitors for assessing sedentary behavior. *Medicine & Science in Sports & Exercise*, *43*(8), 1561–1567. doi: 10.1249/MSS. 0b013e31820ce174

Kozey-Keadle, S., Shiroma, E. J., Freedson, P. S., & Lee, I.-M. (2014). Impact of accelerometer data processing decisions on the sample size, wear time and physical activity level of a large cohort study. *BMC Public Health*, *14*(1210). doi: 10.1186/1471-2458-14-1210

Kujala, U. M. (2018). Is physical activity a cause of longevity? It is not as straightforward as some would believe. A critical analysis. *British Journal of Sports Medicine*, *52*(14), 914–918. doi: 10.1136/bjsports-2017-098639

Kujala, U., Pietilä, J., Myllymäki, T., Mutikainen, S., Föhr, T., Korhonen, I., & Helander, E. (2017). Physical activity: absolute intensity vs. relative-to-fitness-level volumes. *Medicine and Science in Sports and Exercise*, *49*(3), 474–481. doi: 10.1249/MSS.0000000000001134

Kyu, H. H., Bachman, V. F., Alexander, L. T., Mumford, J. E., Afshin, A., Estep, K., … Forouzanfar, M. (2016). Physical activity and risk of breast cancer, colon cancer, diabetes, ischemic heart disease, and ischemic stroke events: systematic review and dose-response meta-analysis for the Global Burden of Disease Study 2013. *BMJ*, *354*, i3857. doi: 10.1136/bmj.i3857

Lagersted-Olsen, J., Korshøj, M., Skotte, J., Carneiro, I. G., Søgaard, K., & Holtermann, A. (2014). Comparison of objectively measured and self-reported time spent sitting. *International Journal of Sports Medicine*, *35*(06), 534–540. doi: 10.1055/s-0033-1358467

Lakerveld, J., Loyen, A., Schotman, N., Peeters, C. F. W., Cardon, G., van der Ploeg, H. P., … Brug, J. (2017). Sitting too much: a hierarchy of socio-demographic correlates. *Preventive Medicine*, *101*, 77–83. doi: 10.1016/j.ypmed.2017.05.015

Lakerveld, J., Mackenbach, J. D., Horvath, E., Rutters, F., Compernolle, S., Bárdos, H., … Brug, J. (2016). The relation between sleep duration and sedentary behaviours in European adults. *Obesity Reviews*, *17*(Suppl 1), 62–67. doi: 10.1111/obr.12381

Lauderdale, D. S., Knutson, K. L., Yan, L. L., Rathouz, P. J., Hulley, S. B., Sidney, S., & Liu, K. (2006). Objectively measured sleep characteristics among early-middle-aged adults: the CARDIA study. *American Journal of Epidemiology*, *164*(1), 5–16. doi: 10.1093/aje/kwj199

Lee, I.-M., & Shiroma, E. J. (2014). Using accelerometers to measure physical activity in large-scale epidemiological studies: issues and challenges. *British Journal of Sports Medicine*, *48*(3), 197–201. doi: 10.1136/bjsports-2013-093154

Lee, I.-M., Shiroma, E. J., Lobelo, F., Puska, P., Blair, S. N., & Katzmarzyk, P. T. (2012). Effect of physical inactivity on major non-communicable diseases worldwide: an analysis of burden of disease and life expectancy. *The Lancet*, *380*(9838), 219–229. doi: 10.1016/S0140-6736(12)61031-9

Lee, P. H., Macfarlane, D. J., Lam, T. H., & Stewart, S. M. (2011). Validity of the international physical activity questionnaire short form (IPAQ-SF): a systematic review. *International Journal of Behavioral Nutrition and Physical Activity*, *8*(115). doi: 10.1186/1479-5868-8-115

Lee, P. H., Yu, Y.-Y., McDowell, I., Leung, G. M., & Lam, T. H. (2013). A cluster analysis of patterns of objectively measured physical activity in Hong Kong. *Public Health Nutrition*, *16*(8), 1436–1444. doi: 10.1017/S1368980012003631

Leinonen, A.-M., Ahola, R., Kulmala, J., Hakonen, H., Vähä-Ypyä, H., Herzig, K.-H., … Jämsä, T. (2016). Measuring physical activity in free-living conditions—comparison of three accelerometry-based methods. *Frontiers in Physiology*, *7*(681). doi: 10.3389/fphys.2016.00681

Liu, S., Gao, R., & Freedson, P. S. (2012). Computational methods for estimating energy expenditure in human physical activities. *Medicine and Science in Sports and Exercise*, *44*(11), 2138–2146. doi: 10.1249/MSS.0b013e31825e825a

Lyden, K., Kozey-Keadle, S., Staudenmayer, J., & Freedson, P. S. (2014). A method to estimate free-living active and sedentary behavior from an accelerometer. *Medicine and Science in Sports and Exercise*, *46*(2), 386–397. doi: 10.1249/MSS.0b013e3182a42a2d

Lyden, K., Petruski, N., Mix, S., Staudenmayer, J., & Freedson, P. S. (2014). Direct observation is a valid criterion for estimating physical activity and sedentary behavior. *Journal of Physical Activity and Health*, *11*(4), 860–863. doi: 10.1123/jpah.2012-0290

Maher, C. A., Mire, E., Harrington, D. M., Staiano, A. E., & Katzmarzyk, P. T. (2013). The independent and combined associations of physical activity and sedentary behavior with obesity in adults: NHANES 2003-06. *Obesity*, *21*(12), E730–E737. doi: 10.1002/oby.20430

Mannini, A., Rosenberger, M., Haskell, W. L., Sabatini, A. M., & Intille, S. S. (2017). Activity recognition in youth using single accelerometer placed at wrist or ankle. *Medicine and Science in Sports and Exercise*, *49*(4), 801–812. doi: 10.1249/MSS.0000000000001144

Mansoubi, M., Pearson, N., Biddle, S. J. H., & Clemes, S. (2014). The relationship between sedentary behaviour and physical activity in adults: a systematic review. *Preventive Medicine*, *69*, 28–35. doi: 10.1016/j.ypmed.2014.08.028

Mathie, M. J., Coster, A. C. F., Lovell, N. H., & Celler, B. G. (2004). Accelerometry: providing an integrated, practical method for long-term, ambulatory monitoring of human movement. *Physiological Measurement*, *25*(2), R1–R20. doi: 10.1088/0967-3334/25/2/r01

Matthews, C. E., Chen, K. Y., Freedson, P. S., Buchowski, M. S., Beech, B. M., Pate, R. R., & Troiano, R. P. (2008). Amount of time spent in sedentary behaviors in the United States, 2003-2004. *American Journal of Epidemiology*, *167*(7), 875–881. doi: 10.1093/aje/kwm390

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

McGregor, D. E., Carson, V., Palarea-Albaladejo, J., Dall, P., Tremblay, M., & Chastin, S. F. M. (2018). Compositional analysis of the associations between 24-h movement behaviours and health indicators among adults and older adults from the Canadian health measure survey. *International Journal of Environmental Research and Public Health*, *15*(8), 1779. doi: 10.3390/ijerph15081779

McGregor, D. E., Palarea-Albaladejo, J., Dall, P. M., del Pozo Cruz, B., & Chastin, S. F. M. (2019). Compositional analysis of the association between mortality and 24-hour movement behaviour from NHANES. *European Journal of Preventive Cardiology*. doi: 10.1177/2047487319867783

McGregor, D. E., Palarea-Albaladejo, J., Dall, P. M., Stamatakis, E., & Chastin, S. F. M. (2019). Differences in physical activity time-use composition associated with cardiometabolic risks. *Preventive Medicine Reports*, *13*, 23–29. doi: 10.1016/j.pmedr.2018.11.006

McVeigh, J. A., Winkler, E. A. H., Healy, G. N., Slater, J., Eastwood, P. R., & Straker, L. M. (2016). Validity of an automated algorithm to identify waking and in-bed wear time in hip-worn accelerometer data collected with a 24 h wear protocol in young adults. *Physiological Measurement*, *37*(10), 1636–1652. doi: 10.1088/0967-3334/37/10/1636

Mekary, R. A., Willett, W. C., Hu, F. B., & Ding, E. L. (2009). Isotemporal substitution paradigm for physical activity epidemiology and weight change. *American Journal of Epidemiology*, *170*(4), 519–527. doi: 10.1093/aje/kwp163

Mezick, E. J., Hall, M., & Matthews, K. A. (2011). Are sleep and depression independent or overlapping risk factors for cardiometabolic disease? *Sleep Medicine Reviews*, *15*(1), 51–63. doi: 10.1016/j.smrv.2010.03.001

Migueles, J. H., Aadland, E., Andersen, L. B., Brønd, J. C., Chastin, S. F., Hansen, B. H., … Ortega, F. B. (2021). GRANADA consensus on analytical approaches to assess associations with accelerometer-determined physical behaviours (physical activity, sedentary behaviour and sleep) in epidemiological studies. *British Journal of Sports Medicine*, (Publication ahead of print). doi: 10.1136/bjsports-2020-103604.

Migueles, J. H., Cadenas-Sanchez, C., Ekelund, U., Nyström, C. D., Mora-Gonzalez, J., Löf, M., … Ortega, F. B. (2017). Accelerometer data collection and processing criteria to assess physical activity and other outcomes: a systematic review and practical considerations. *Sports Medicine*, *47*(9), 1821–1845. doi: 10.1007/s40279-017-0716-0

Migueles, J. H., Cadenas-Sanchez, C., Tudor-Locke, C., Löf, M., Esteban-Cornejo, I., Molina-Garcia, P., … Ortega, F. B. (2019). Comparability of published cut-points for the assessment of physical activity: implications for data harmonization. *Scandinavian Journal of Medicine & Science in Sports*, *29*(4), 566–574. doi: 10.1111/sms.13356

Millán, J., Pintó, X., Muñoz, A., Zúñiga, M., Rubiés-Prat, J., Pallardo, L. F., … Pedro-Botet, J. (2009). Lipoprotein ratios: physiological significance and clinical usefulness in cardiovascular prevention. *Vascular Health and Risk Management*, *5*, 757–765. doi: PMC2747394

Mohamad, I., & Usman, D. (2013). Standardization and its effects on k-means clustering algorithm. *Research Journal of Applied Sciences, Engineering and Technology*, *6*(17), 3299–3303. doi: 10.19026/rjaset.6.3638

Montoye, A. H. K., Mudd, L. M., Biswas, S., & Pfeiffer, K. A. (2015). Energy expenditure prediction using raw accelerometer data in simulated free living. *Medicine & Science in Sports & Exercise*, *47*(8), 1735–1746. doi: 10.1249/MSS.0000000000000597

Montoye, A. H. K., Nelson, M. B., Bock, J. M., Imboden, M. T., Kaminsky, L. A., Mackintosh, K. A., … Pfeiffer, K. A. (2018). Raw and count data comparability of hip-worn ActiGraph GT3X+ and Link accelerometers. *Medicine and Science in Sports and Exercise*, *50*(5), 1103–1112. doi: 10.1249/MSS.0000000000001534

Montoye, A. H. K., Pivarnik, J. M., Mudd, L. M., Biswas, S., & Pfeiffer, K. A. (2016a). Comparison of activity type classification accuracy from accelerometers worn on the hip, wrists, and thigh in young, apparently healthy adults. *Measurement in Physical Education and Exercise Science*, *20*(3), 173–183. doi: 10.1080/1091367X.2016.1192038

Montoye, A. H. K., Pivarnik, J. M., Mudd, L. M., Biswas, S., & Pfeiffer, K. A. (2016b). Validation and comparison of accelerometers worn on the hip, thigh, and wrists for measuring physical activity and sedentary behavior. *AIMS Public Health*, *3*(2), 298–312. doi: 10.3934/publichealth.2016.2.298

Montoye, A. H. K., Westgate, B. S., Fonley, M. R., & Pfeiffer, K. A. (2018). Cross-validation and out-of-sample testing of physical activity intensity predictions using a wrist-worn accelerometer. *Journal of Applied Physiology*, *124*(5), 1284–1293. doi: 10.1152/japplphysiol.00760.2017

Niemelä, M., Kangas, M., Farrahi, V., Kiviniemi, A., Leinonen, A.-M., Ahola, R., … Jämsä, T. (2019). Intensity and temporal patterns of physical activity and cardiovascular disease risk in midlife. *Preventive Medicine*, *124*, 33–41. doi: 10.1016/j.ypmed.2019.04.023

Nweke, H. F., Teh, Y. W., Mujtaba, G., & Al-Garadi, M. A. (2019). Data fusion and multiple classifier systems for human activity detection and health monitoring: review and open research directions. *Information Fusion*, *46*, 147–170. doi: 10.1016/j.inffus.2018.06.002

O'Donoghue, G., Perchoux, C., Mensah, K., Lakerveld, J., Van Der Ploeg, H., Bernaards, C., … Nazare, J.-A. (2016). A systematic review of correlates of sedentary behaviour in adults aged 18-65 years: a socio-ecological approach. *BMC Public Health*, *16*(163). doi: 10.1186/s12889-016-2841-3

O'Donovan, G., Lee, I.-M., Hamer, M., & Stamatakis, E. (2017). Association of "weekend warrior" and other leisure time physical activity patterns with risks for all-cause, cardiovascular disease, and cancer mortality. *JAMA Internal Medicine*, *177*(3), 335–342. doi: 10.1001/jamainternmed.2016.8014

Ortega, F. B., Konstabel, K., Pasquali, E., Ruiz, J. R., Hurtig-Wennlöf, A., Mäestu, J., … Sjöström, M. (2013). Objectively measured physical activity and sedentary time during childhood, adolescence and young adulthood: a cohort study. *PloS One*, *8*(4), e60871. doi: 10.1371/journal.pone.0060871

Owen, N., Healy, G. N., Matthews, C. E., & Dunstan, D. W. (2010). Too much sitting: the population-health science of sedentary behavior. *Exercise and Sport Sciences Reviews*, *38*(3), 105–113. doi: 10.1097/JES.0b013e3181e373a2

Owen, N., Sugiyama, T., Eakin, E. E., Gardiner, P. A., Tremblay, M. S., & Sallis, J. F. (2011). Adults' sedentary behavior: determinants and interventions. *American Journal of Preventive Medicine*, *41*(2), 189–196. doi: 10.1016/j.amepre.2011.05.013

Paing, A. C., McMillan, K. A., Kirk, A. F., Collier, A., Hewitt, A., & Chastin, S. F. M. (2019). Dose-response between frequency of breaks in sedentary time and glucose control in type 2 diabetes: a proof of concept study. *Journal of Science and Medicine in Sport*, *22*(7), 808–813. doi: 10.1016/j.jsams.2019.01.017

Pate, R. R., Berrigan, D., Buchner, D. M., Carlson, S. A., Dunton, G., Fulton, J. E., … Whitsel, L. P. (2018). Actions to improve physical activity surveillance in the United States. *NAM Perspectives*. Discussion Paper, National Academy of Medicine, Washington, DC. doi: 10.31478/201809f

Patterson, F., Lozano, A., Huang, L., Perkett, M., Beeson, J., & Hanlon, A. (2018). Towards a demographic risk profile for sedentary behaviours in middle-aged British adults: a cross-sectional population study. *BMJ Open*, *8*(7), e019639. doi: 10.1136/bmjopen-2017-019639

Pedišić, Ž. (2014). 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*, *46*(1), 135–146.

Pedišić, Ž., Dumuid, D., & S Olds, T. (2017). Integrating sleep, sedentary behaviour, and physical activity research in the emerging field of time-use epidemiology: definitions, concepts, statistical methods, theoretical framework, and future directions. *Kinesiology*, *49*(2), 252–269.

Powell, C., Browne, L. D., Carson, B. P., Dowd, K. P., Perry, I. J., Kearney, P. M., … Donnelly, A. E. (2019). Use of compositional data analysis to show estimated changes in cardiometabolic health by reallocating time to light-intensity physical activity in older adults. *Sports Medicine*, *50*(1), 205–217. doi: 10.1007/s40279-019-01153-2

Powell, C., Herring, M. P., Dowd, K. P., Donnelly, A. E., & Carson, B. P. (2018). The cross-sectional associations between objectively measured sedentary time and cardiometabolic health markers in adults - a systematic review with meta-analysis component. *Obesity Reviews*, *19*(3), 381–395. doi: 10.1111/obr.12642

Prince, S. A., Adamo, K. B., Hamel, M. E., Hardt, J., Connor-Gorber, S., & Tremblay, M. (2008). A comparison of direct versus self-report measures for assessing physical activity in adults: a systematic review. *International Journal of Behavioral Nutrition* *and Physical Activity*, *5*(56). doi: 10.1186/1479-5868-5-56

Proper, K. I., Singh, A. S., Van Mechelen, W., & Chinapaw, M. J. M. (2011). Sedentary behaviors and health outcomes among adults: a systematic review of prospective studies. *American Journal of Preventive Medicine*, *40*(2), 174–182. doi: 10.1016/j.amepre.2010.10.015

Pulsford, R. M., Blackwell, J., Hillsdon, M., & Kos, K. (2017). Intermittent walking, but not standing, improves postprandial insulin and glucose relative to sustained sitting: a randomised cross-over study in inactive middle-aged men. *Journal of Science and Medicine in Sport*, *20*(3), 278–283. doi: 10.1016/j.jsams.2016.08.012

Ramakrishnan, R., Doherty, A., Smith-Byrne, K., Rahimi, K., Bennett, D., Woodward, M., … Dwyer, T. (2021). Accelerometer measured physical activity and the incidence of cardiovascular disease: evidence from the UK Biobank cohort study. *PLoS Medicine*, *18*(1), e1003487. doi: 10.1371/journal.pmed.1003487

Reiss, A., & Stricker, D. (2012). Introducing a new benchmarked dataset for activity monitoring. *16th International Symposium on Wearable Computers, ISWC*, 108–109. doi: 10.1109/ISWC.2012.13

Rezende, L. F. M., Sá, T. H., Mielke, G. I., Viscondi, J. Y. K., Rey-López, J. P., & Garcia, L. M. T. (2016). All-cause mortality attributable to sitting time: analysis of 54 countries worldwide. *American Journal of Preventive Medicine*, *51*(2), 253–263. doi: 10.1016/j.amepre.2016.01.022

Rhodes, R. E., Mark, R. S., & Temmel, C. P. (2012). Adult sedentary behavior: a systematic review. *American Journal of Preventive Medicine*, *42*(3), e3–e28. doi: 10.1016/j.amepre.2011.10.020

Rosenberger, M. E., Buman, M. P., Haskell, W. L., McConnell, M. V, & Carstensen, L. L. (2016). Twenty-four hours of sleep, sedentary behavior, and physical activity with nine wearable devices. *Medicine and Science in Sports and Exercise*, *48*(3), 457–465. doi: 10.1249/MSS.0000000000000778

Rosenberger, M. E., Fulton, J. E., Buman, M. P., Troiano, R. P., Grandner, M. A., Buchner, D. M., & Haskell, W. L. (2019). The 24-hour activity cycle: a new paradigm for physical activity. *Medicine and Science in Sports and Exercise*, *51*(3), 454–464. doi: 10.1249/MSS.0000000000001811

Rowlands, A. V, Mirkes, E. M., Yates, T. E., Clemes, S. A., Davies, M. J., Khunti, K., & Edwardson, C. L. (2017). Accelerometer-assessed physical activity in epidemiology: are monitors equivalent? *Medicine & Science in Sports & Exercise*, *50*(2), 257–265. doi: 10.1249/MSS.0000000000001435

Saeidifard, F., Medina-Inojosa, J. R., Supervia, M., Olson, T. P., Somers, V. K., Erwin, P. J., & Lopez-Jimenez, F. (2018). Differences of energy expenditure while sitting versus standing: a systematic review and meta-analysis. *European Journal of Preventive Cardiology*, *25*(5), 522–538. doi: 10.1177/2047487317752186

Sallis, J. F., Cervero, R. B., Ascher, W., Henderson, K. A., Kraft, M. K., & Kerr, J. (2006). An ecological approach to creating active living communities. *Annual Review of Public Health*, *27*, 297–322. doi: 10.1146/annurev.publhealth.27.021405.102100

Samitz, G., Egger, M., & Zwahlen, M. (2011). Domains of physical activity and all-cause mortality: systematic review and dose-response meta-analysis of cohort studies. *International Journal of Epidemiology*, *40*(5), 1382–1400. doi: 10.1093/ije/dyr112

Shephard, R. J. (2003). Limits to the measurement of habitual physical activity by questionnaires. *British Journal of Sports Medicine*, *37*(3), 197–206. doi: 10.1136/bjsm.37.3.197

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

Silva, K. S., Garcia, L. M. T., Rabacow, F. M., de Rezende, L. F. M., & de Sá, T. H. (2017). Physical activity as part of daily living: moving beyond quantitative recommendations. *Preventive Medicine*, *96*, 160–162. doi: 10.1016/j.ypmed.2016.11.004

Simonsohn, U. (2018). Two lines: a valid alternative to the invalid testing of U-shaped relationships with quadratic regressions. *Advances in Methods and Practices in Psychological Science*, *1*(4), 538–555. doi: 10.1177/2515245918805755

Sintonen, H. (2001). The 15D instrument of health-related quality of life: properties and applications. *Annals of Medicine*, *33*(5), 328–336. doi: 10.3109/07853890109002086

Spittaels, H., Van Cauwenberghe, E., Verbestel, V., De Meester, F., Van Dyck, D., Verloigne, M., … De Bourdeaudhuij, I. (2012). Objectively measured sedentary time and physical activity time across the lifespan: a cross-sectional study in four age groups. *International Journal of Behavioral Nutrition and Physical Activity*, *9*(149). doi: 10.1186/1479-5868-9-149

St-Onge, M.-P., Grandner, M. A., Brown, D., Conroy, M. B., Jean-Louis, G., Coons, M., & Bhatt, D. L. (2016). Sleep duration and quality: impact on lifestyle behaviors and cardiometabolic health: a scientific statement from the American Heart Association. *Circulation*, *134*(18), e367–e386. doi: 10.1161/CIR.0000000000000444

Stamatakis, E., Ekelund, U., Ding, D., Hamer, M., Bauman, A. E., & Lee, I.-M. (2019). Is the time right for quantitative public health guidelines on sitting? a narrative review of sedentary behaviour research paradigms and findings. *British Journal of Sports Medicine*, *53*(6), 377–382. doi: 10.1136/bjsports-2018-099131

Stamatakis, E., Gale, J., Bauman, A., Ekelund, U., Hamer, M., & Ding, D. (2019). Sitting time, physical activity, and risk of mortality in adults. *Journal of the American College of Cardiology*, *73*(16), 2062–2072. doi: 10.1016/j.jacc.2019.02.031

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

Staudenmayer, J., Zhu, W., & Catellier, D. J. (2012). Statistical considerations in the analysis of accelerometry-based activity monitor data. *Medicine and Science in Sports and Exercise*, *44*(1 Suppl 1), S61–S67. doi: 10.1249/MSS.0b013e3182399e0f

Štefan, L., Horvatin, M., & Baić, M. (2019). Are sedentary behaviors associated with sleep duration? a cross-sectional case from Croatia. *International Journal of Environmental Research and Public Health*, *16*(2), 200. doi: 10.3390/ijerph16020200

Strain, T., Wijndaele, K., Dempsey, P. C., Sharp, S. J., Pearce, M., Jeon, J., … Brage, S. (2020). Wearable-device-measured physical activity and future health risk. *Nature Medicine*, *26*, 1385–1391. doi: 10.1038/s41591-020-1012-3

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

Telama, R., Yang, X., Leskinen, E., Kankaanpää, A., Hirvensalo, M., Tammelin, T., … Raitakari, O. T. (2014). Tracking of physical activity from early childhood through youth into adulthood. *Medicine & Science in Sports & Exercise*, *46*(5), 955–962. doi: 10.1249/MSS.0000000000000181

Thorp, A. A., McNaughton, S. A., Owen, N., & Dunstan, D. W. (2013). Independent and joint associations of TV viewing time and snack food consumption with the metabolic syndrome and its components; a cross-sectional study in Australian adults. *International Journal of Behavioral Nutrition and Physical Activity*, *10*(96). doi: 10.1186/1479-5868-10-96

Thorp, A. A., Owen, N., Neuhaus, M., & Dunstan, D. W. (2011). Sedentary behaviors and subsequent health outcomes in adults: a systematic review of longitudinal studies, 1996-2011. *American Journal of Preventive Medicine*, *41*(2), 207–215. doi: 10.1016/j.amepre.2011.05.004

Tjurin, P., Niemelä, M., Huusko, M., Ahola, R., Kangas, M., & Jämsä, T. (2017). Classification of physical activities and sedentary behavior using raw data of 3D hip acceleration. *Nordic-Baltic Conference on Biomedical Engineering and Medical Physics*, 872–875. doi: 10.1007/978-981-10-5122-7_218

Tremblay, M. S., Aubert, S., Barnes, J. D., Saunders, T. J., Carson, V., Latimer-Cheung, A. E., … Chinapaw, M. J. M. (2017). Sedentary behavior research network (SBRN) - terminology consensus project process and outcome. *International Journal of Behavioral Nutrition and Physical Activity*, *14*(75). doi: 10.1186/s12966-017-0525-8

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

Trost, S. G., Cliff, D., Ahmadi, M., Van Tuc, N., & Hagenbuchner, M. (2017). Sensor-enabled activity class recognition in preschoolers: hip versus wrist data. *Medicine and Science in Sports and Exercise*, *50*(3), 634–641. doi: 10.1249/MSS.0000000000001460

Trost, S. G., Mciver, K. L., & Pate, R. R. (2005). Conducting accelerometer-based activity assessments in field-based research. *Medicine & Science in Sports & Exercise*, *37*(11), S531–S543. doi: 10.1249/01.mss.0000185657.86065.98

Trost, S. G., Owen, N., Bauman, A. E., Sallis, J. F., & Brown, W. (2002). Correlates of adults' participation in physical activity: review and update. *Medicine & Science in Sports & Exercise*, *34*(12), 1996–2001. doi: 10.1097/00005768-200212000-00020

Trost, S. G., Wong, W.-K., Pfeiffer, K. A., & Zheng, Y. (2012). Artificial neural networks to predict activity type and energy expenditure in youth. *Medicine and Science in Sports and Exercise*, *44*(9), 1801–1809. doi: 10.1249/MSS.0b013e318258ac11

Trost, S. G., Zheng, Y., & Wong, W.-K. (2014). Machine learning for activity recognition: hip versus wrist data. *Physiological Measurement*, *35*(11), 2183–2189. doi: 10.1088/0967-3334/35/11/2183

Tsunoda, K., Kitano, N., Kai, Y., Uchida, K., Kuchiki, T., Okura, T., & Nagamatsu, T. (2015). Prospective study of physical activity and sleep in middle-aged and older adults. *American Journal of Preventive Medicine*, *48*(6), 662–673. doi: 10.1016/j.amepre.2014.12.006

Vähä-Ypyä, H., Husu, P., Suni, J., Vasankari, T., & Sievänen, H. (2018). Reliable recognition of lying, sitting, and standing with a hip-worn accelerometer. *Scandinavian Journal of Medicine & Science in Sports*, *28*(3), 1092–1102. doi: 10.1111/sms.13017

Vähä-Ypyä, H., Vasankari, T., Husu, P., Mänttäri, A., Vuorimaa, T., Suni, J., & Sievänen, H. (2015). Validation of cut-points for evaluating the intensity of physical activity with accelerometry-based mean amplitude deviation (MAD). *PLoS One*, *10*(8), e0134813. doi: 10.1371/journal.pone.0134813

van Hees, V. T., Thaler-Kall, K., Wolf, K.-H., Brønd, J. C., Bonomi, A., Schulze, M., … Horsch, A. (2016). Challenges and opportunities for harmonizing research methodology: raw accelerometry. *Methods of Information in Medicine*, *55*(06), 525–532. doi: 10.3414/ME15-05-0013

Venkatasubramaniam, A., Wolfson, J., Mitchell, N., Barnes, T., JaKa, M., & French, S. (2017). Decision trees in epidemiological research. *Emerging Themes in Epidemiology*, *14*(11). doi: 10.1186/s12982-017-0064-4

Verswijveren, S. J. J., Lamb, K. E., Leech, R. M., Salmon, J., Timperio, A., Telford, R. M., … Ridgers, N. D. (2020). Activity accumulation and cardiometabolic risk in youth: a latent profile approach. *Medicine & Science in Sports & Exercise*, *52*(7), 1502–1510. doi: 10.1249/MSS.0000000000002275

Vincent, G. E., Jay, S. M., Sargent, C., Vandelanotte, C., Ridgers, N. D., & Ferguson, S. A. (2017). Improving cardiometabolic health with diet, physical activity, and breaking up sitting: what about sleep? *Frontiers in Physiology*, *8*(865). doi: 10.3389/fphys.2017.00865

von Rosen, P., Dohrn, I.-M., & Hagströmer, M. (2020a). Association between physical activity and all-cause mortality: a 15-year follow-up using a compositional data analysis. *Scandinavian Journal of Medicine & Science in Sports*, *30*, 100–107. doi: 10.1111/sms.13561

von Rosen, P., Dohrn, I.-M., & Hagströmer, M. (2020b). Latent profiles analysis of physical activity and sedentary behaviour with mortality risk: a 15-year follow-up. *Scandinavian Journal of Medicine & Science in Sports*, *30*(10), 1949–1956. doi: 10.1111/sms.13761

Wallace, T. M., Levy, J. C., & Matthews, D. R. (2004). Use and abuse of HOMA modeling. *Diabetes Care*, *27*(6), 1487–1495. doi: 10.2337/diacare.27.6.1487

Wang, J., Chen, Y., Hao, S., Peng, X., & Hu, L. (2019). Deep learning for sensor-based activity recognition: a survey. *Pattern Recognition Letters*, *119*, 3–11. doi: 10.1016/j.patrec.2018.02.010

Warburton, D. E. R., & Bredin, S. S. D. (2016). Reflections on physical activity and health: what should we recommend? *Canadian Journal of Cardiology*, *32*(4), 495–504. doi: 10.1016/j.cjca.2016.01.024

Warburton, D. E. R., & Bredin, S. S. D. (2017). Health benefits of physical activity: a systematic review of current systematic reviews. *Current Opinion in Cardiology*, *32*(5), 541–556. doi: 10.1097/HCO.0000000000000437

Warren, J. M., Ekelund, U., Besson, H., Mezzani, A., Geladas, N., & Vanhees, L. (2010). 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. *European Journal of Cardiovascular Prevention & Rehabilitation*, *17*(2), 127–139. doi: 10.1097/HJR.0b013e32832ed875

Welk, G. J. (2005). Principles of design and analyses for the calibration of accelerometry-based activity monitors. *Medicine and Science in Sports and Exercise*, *37*(11 Suppl), S501–S511. doi: 10.1249/01.mss.0000185660.38335.de

Welk, G. J., McClain, J., & Ainsworth, B. E. (2012). Protocols for evaluating equivalency of accelerometry-based activity monitors. *Medicine and Science in Sports and Exercise*, *44*(1 Suppl 1), S39–S49. doi: 10.1249/MSS.0b013e3182399d8f

Wen, C. P., Wai, J. P. M., Tsai, M. K., Yang, Y. C., Cheng, T. Y. D., Lee, M.-C., … Wu, X. (2011). Minimum amount of physical activity for reduced mortality and extended life expectancy: a prospective cohort study. *The Lancet*, *378*(9798), 1244–1253. doi: 10.1016/S0140-6736(11)60749-6

WHO. (2010). *Global recommendations on physical activity for health*. Retrieved from https://www.who.int/publications/i/item/9789241599979

WHO. (2013). *Global action plan for the prevention and control of noncommunicable diseases 2013–2020.* Retrieved from https://www.who.int/nmh/publications/ncd-action-plan/en/

Wijndaele, K., Westgate, K., Stephens, S. K., Blair, S. N., Bull, F. C., Chastin, S. F. M., … Healy, G. N. (2015). Utilization and harmonization of adult accelerometry data: review and expert consensus. *Medicine and Science in Sports and Exercise*, *47*(10), 2129–2139. doi: 10.1249/MSS.0000000000000661

Wilmot, E. G., Edwardson, C. L., Achana, F. A., Davies, M. J., Gorely, T., Gray, L. J., … Biddle, S. J. H. (2012). Sedentary time in adults and the association with diabetes, cardiovascular disease and death: systematic review and meta-analysis. *Diabetologia*, *55*(11), 2895–2905. doi: 10.1007/s00125-012-2677-z

Xi, B., He, D., Zhang, M., Xue, J., & Zhou, D. (2014). Short sleep duration predicts risk of metabolic syndrome: a systematic review and meta-analysis. *Sleep Medicine Reviews*, *18*(4), 293–297. doi: 10.1016/j.smrv.2013.06.001

Xu, R., & Wunsch, D. (2005). Survey of clustering algorithms. *IEEE Transactions on Neural Networks*, *16*(3), 645–678. doi: 10.1109/TNN.2005.845141

Yang, C.-C., & Hsu, Y.-L. (2010). A review of accelerometry-based wearable motion detectors for physical activity monitoring. *Sensors*, *10*(8), 7772–7788. doi: 10.3390/s100807772

Yoon, S., Suero-Tejeda, N., & Bakken, S. (2015). A data mining approach for examining predictors of physical activity among urban older adults. *Journal of Gerontological Nursing*, *41*(17), 14–20. doi: 10.3928/00989134-20150420-01

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

Zhang, S., Rowlands, A. V, Murray, P., & Hurst, T. L. (2012). Physical activity classification using the GENEA wrist-worn accelerometer. *Medicine & Science in Sports & Exercise*, *44*(4), 742–748. doi: 10.1249/MSS.0b013e31823bf95c

Appendix 1

PubMed search string:
(Prediction OR unsupervised OR supervised OR energy expenditure OR energy cost OR physical activit* classif* OR pattern recogni* OR activit* recogni* OR machine learning) AND (acceleromet*)
Scopus search string:
(Prediction OR unsupervised OR supervised OR energy expenditure OR energy cost OR physical activit* classif* OR pattern recogni* OR activit* recogni* OR machine learning) AND (acceleromet*)

Quality scores of the included studies developing activity recognition models, listed by first author's last name (publication year).

Author (year)	Crit. 1	Crit. 2	Crit. 3	Crit. 4	Crit. 5	Crit. 6	Crit. 7	Crit. 8	Crit. 9	Crit. 10	Quality Score
Pober (2006) [36]	2	2	1	0	2	2	2	0	2	2	0.75
Bonomi (2009) [37]	1	1	2	0	2	2	2	1	1	1	0.65
Preece (2009) [38]	1	1	2	0	2	2	2	0	1	1	0.60
Khan (2010) [39]	2	2	1	1	2	2	2	0	2	2	0.80
Atallah (2011) [34]	2	1	0	0	2	2	2	0	1	1	0.65
De Vries (2011) [40]	2	2	2	1	2	2	2	2	2	2	0.95
De Vries (2011) [41]	2	2	2	1	2	2	2	2	2	2	0.95
Gyllensten (2011) [42]	2	1	2	2	2	2	2	2	2	2	0.95
Lee (2011) [43]	2	2	0	1	2	2	2	1	2	2	0.80
Ruch (2011) [44]	1	1	2	NA	2	2	2	2	1	1	0.77
Schmid (2011) [45]	1	1	1	1	2	2	1	1	2	2	0.70
Oudre (2012) [24]	1	1	2	0	2	2	2	2	1	1	0.70
Zhang (2012) [46]	2	2	2	1	2	2	2	2	2	2	0.95
Zhang (2012) [47]	2	2	2	1	2	2	2	2	2	2	0.95
Cleland (2013) [48]	2	2	2	0	2	2	2	0	2	2	0.80
Hees (2013) [26]	2	2	2	2	2	2	2	2	2	2	1
John (2013) [49]	2	1	1	0	2	2	2	0	2	2	0.70
Mannini (2013) [50]	2	2	1	2	2	2	2	2	2	2	0.95
Zhao (2013) [51]	1	1	1	1	2	2	2	2	1	1	0.70
He (2014) [28]	1	0	2	2	2	2	2	1	1	2	0.75
Trost (2014) [52]	2	2	2	1	2	0	2	2	1	1	0.75
Arif (2015) [53]	1	2	2	1	2	2	2	0	2	2	0.80
Bastian (2015) [54]	2	2	2	2	2	2	2	2	2	2	1
Fida (2015) [55]	2	2	2	1	2	2	2	1	2	2	0.90
Hagenbuchner (2015) [56]	2	2	2	0	2	2	2	2	2	2	0.90
Ellis (2016) [57]	2	2	1	2	2	2	2	2	2	2	0.95

Author (year)	Crit. 1	Crit. 2	Crit. 3	Crit. 4	Crit. 5	Crit. 6	Crit. 7	Crit. 8	Crit. 9	Crit. 10	Quality Score
Kerr (2016) [58]	2	2	2	NA	2	2	2	2	2	2	1
Margarito (2016) [25]	1	1	2	0	2	2	2	2	1	1	0.70
Montoye (2016) [59]	2	2	2	2	2	2	2	2	2	2	1
Montoye (2016) [60]	2	2	2	2	2	2	2	2	2	2	1
Ren (2016) [61]	1	1	1	1	2	0	2	2	2	1	0.65
Sasaki (2016) [62]	2	2	1	NA	2	2	2	2	2	2	1
Arif (2017) [63]	1	2	2	1	2	2	2	0	2	2	0.80
Chowdhury (2017) [64]	2	2	1	1	2	2	2	1	2	2	0.85
Kühnhausen (2017) [65]	2	2	2	0	2	2	2	2	2	2	0.90
Mannini (2017) [66]	2	2	1	2	2	2	2	2	2	2	0.95
Paveya (2017) [67]	2	1	2	0	2	2	2	2	2	2	0.85
Rosenberg (2017) [68]	2	2	1	NA	2	2	2	2	2	2	0.95
Trost (2017) [69]	2	2	2	1	2	2	2	0	2	2	0.85
Montoye (2018) [70]	2	2	2	2	2	2	2	2	2	2	1
Staudenmayer (2009) [78]	2	2	2	2	2	2	2	2	2	2	1
Freedson (2011) [79]	2	2	2	2	2	2	2	2	2	2	1
Trost (2012) [80]	2	2	2	1	2	2	2	2	2	2	0.95
Ellis (2014) [81]	2	2	2	0	2	2	2	2	2	2	0.90
Mu (2014) [82]	1	1	2	2	2	2	2	2	2	2	0.90
Staudenmayer (2015) [83]	2	2	2	2	2	2	2	2	2	2	1
Strath (2015) [84]	2	2	2	0	2	1	2	2	2	1	0.80
Kate (2016) [85]	2	1	2	0	2	1	0	2	2	1	0.65

2: indicates “yes”, 1: indicates “partial”, 0: indicates “no”, N/A: indicates “not applicable”. Crit. = criteria. Please note that the numbers presented in square brackets [ ] are reference numbers, and they are according to the list of references in sub-study I.

Criteria 1: Question/objective sufficiently described? Criteria 2: Study design evident and appropriate? Criteria 3: Subject characteristics sufficiently described and representative? Criteria 4: If performed under controlled conditions, a wide variety of activities performed with enough duration? Criteria 5: Input features clearly mentioned? Criteria 6: Signal axes used for feature extraction clearly mentioned? Criteria 7: Window length clearly mentioned? Criteria 8: Sample size appropriate? Criteria 9: Results reported in sufficient detail? Criteria 10: Conclusions supported by the results?

Original publications

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

II Farrahi, V., Niemelä, M., Tjurin, P., Kangas, M., Korpelainen, R., & Jämsä, T. (2020). Evaluating and enhancing the generalization performance of machine learning models for physical activity intensity prediction from raw acceleration data. *IEEE Journal of Biomedical and Health Informatics*, *24*(1), 27–38. doi: 10.1109/JBHI.2019.2917565

III Farrahi, V., Kangas, M., Walmsley, R., Niemelä, M., Kiviniemi, A., Puukka, K., Collings, P., Korpelainen, R., & Jämsä, T. (2021). Compositional associations of sleep and activities within the 24-h cycle with cardiometabolic health markers in adults. *Medicine and Science in Sports and Exercise*, *53*(2), 324–332. doi: 10.1249/MSS.0000000000002481

IV Farrahi, V., Kangas, M., Kiviniemi, A., Puukka, K., Korpelainen, R., & Jämsä, T. (2021). Accumulation patterns of sedentary time and breaks and their association with cardiometabolic health markers in adults. *Scandinavian Journal of Medicine & Science in Sports*, Online ahead of print. doi: 10.1111/sms.13958

V Farrahi, V., Niemelä, M., Kärmeniemi, M., Puhakka, S., Kangas, M., Korpelainen, R., & Jämsä, T. (2020). Correlates of physical activity behavior in adults: a data mining approach. *International Journal of Behavioral Nutrition and Physical Activity*, *17*(94). doi: 10.1186/s12966-020-00996-7

Original publications are not included in the electronic version of the dissertation.

ACTA UNIVERSITATIS OULUENSIS
SERIES D MEDICA

UNIVERSITY OF OULU P.O. Box 8000 FI-90014 UNIVERSITY OF OULU FINLAND

ACTA UNIVERSITATIS OULUENSIS

ISBN 978-952-62-2945-4 (Paperback)
ISBN 978-952-62-2946-1 (PDF)
ISSN 0355-3221 (Print)
ISSN 1796-2234 (Online)