Université de Strasbourg

UNIVERSITÉ DE STRASBOURG

ÉCOLE DOCTORALE DES SCIENCES DE LA VIE ET DE LA SANTÉ

IPHC - DEPE / UMR 7178

THÈSE présentée par :

Isaac DEBACHE

soutenue le : **10/03/2020**

pour obtenir le grade de : **Docteur de l'université de Strasbourg**

Discipline/ Spécialité : Epidémiologie

Relationships between Urban Environment, Physical Activity, and Health

THÈSE dirigée par :

Prof. Cédric SUEUR	Maître de conférences, université de Strasbourg

RAPPORTEURS :

Prof. Anne VUILLEMIN	Professeur des universités, Université Côte d'Azur
Mme Emmanuelle CADOT	Chargée de recherche, Institut de Recherche pour le Développement

AUTRES MEMBRES DU JURY :

Dr. Stéphane BLANC	Directeur de Recherche, Institut Pluridisciplinaire Hubert Curien

INVITÉS :

Dr. Audrey BERGOUIGNAN	Directeur de Recherche, Institut Pluridisciplinaire Hubert Curien

UNIVERSITÉ DE STRASBOURG

École doctorale Vie et Santé (ED414)

Thèse présentée et soutenue par

Isaac DEBACHE

Pour obtenir le grade de docteur de l'Université de Strasbourg

Disciplines : Épidémiologie

Le 10 mars 2020

Relationships between Urban Environment, Physical Activity and Health

Jury

Dr. Stéphane BLANC……………………………………………..……………………Rapporteur interne

Prof. Anne VUILLEMIN………………………………………..………………..Rapporteur externe

Dr. Emmanuelle CADOT………………………………………………………………Rapporteur externe

Dr. Cédric SUEUR…………………………………………………………………Directeur de thèse

Dr. Audrey BERGOUIGNAN……………..……..…………………………..…………………Invitée

Remerciements

Cette thèse a été financée par l'**Agence Nationale pour la Recherche** et la **Région Alsace**. Je tiens à remercier tous ceux qui ont rendu ce financement possible. J'espère avoir été à la hauteur des attentes qui ont été placées en moi.

Je remercie chaleureusement mon directeur de thèse, **Cédric Sueur**, et ma co-encadrante, **Audrey Bergouignan**, pour leurs conseils, leur gentillesse et la confiance qu'ils m'ont accordée.

Merci à **Anne Vuillemin**, **Emmanuelle Cadot** et **Stéphane Blanc** qui m'ont fait l'honneur d'être membres du jury de cette thèse.

Merci à **Basile Chaix** de l'UMR 707 de l'INSERM à Paris pour ses brillantes intuitions et son accueil, ainsi qu'à **Emiel Sneekes** du service de rééducation de l'hôpital Erasmus à Rotterdam pour avoir guidé mes premiers pas dans le monde de l'accélérométrie.

Depuis novembre 2016, le **Département d'Ecologie, Physiologie et Ethologie** (DEPE) de l'Institut Pluridisciplinaire Hubert Curien à Strasbourg a été pour moi comme un second foyer. Merci à tous les membres de ce laboratoire pour l'accueil chaleureux qu'ils m'ont réservé.

Enfin, merci à mes collègues **doctorants**, avec qui j'ai eu le bonheur de partager cette aventure, pour leur amitié et leur soutien à toute épreuve. J'aimerais remercier tout particulièrement **Lorène Jeantet** pour ses encouragements et sa précieuse contribution au troisième chapitre de cette thèse.

Contributions

Articles intégrés à cette thèse

Debache I, Bergouignan A, Chaix B, Sneekes EM, Thomas F , Sueur C. Associations of Sensor-*Derived Physical Behavior with Metabolic Health: A Compositional Analysis in the Record Multisensor Study*. International journal of environmental research and public health 16 (5), 741. 2019.

De Jong N, **Debache I**, Pan Z, Garnotel M, Lyden K, Sueur C, Simon C, Bessesen D, Bergouignan *A. Breaking up sedentary time in overweight/obese adults on work days and non-work days: results from a feasibility study*. International journal of environmental research and public health 15 (11), 2566 (2019).

Debache I, Jeantet L, Bergouignan, A., Chevallier, D. Sueur C. *A lean and performant hierarchical model for human activity recognition using body-mounted sensors.* En cours de soumission.

Debache I, Sueur C, Bergouignan A, Vallée J, Chaix B. *Effects of built and social environmental factors on distribution of physical activity and body postures during daily commutes*. En cours de soumission.

Sanjeev B, **Debache I**, Chaix B. *Physical activity and sedentary behaviour related to transport activity: Analysis of multiple body-worn accelerometer data from the RECORD MultiSensor Study*. En cours de soumission.

Communications orales et affichées

Rynders CA, Steinke S, Kealey E, Tussey E, **Debache I**, Bergouignan A, Zaman A, Thomas E. *Associations Among Sedentary Time, Feeding Duration, And Sleep Duration In Adults With Overweight And Obesity*. 23 Mai 2019. 66th annual meeting of the American college of sports medicine congress, Orlando, Florida, USA.

Debache, I: *Associations of sensor-derived physical behavior with metabolic health: a compositional analysis*. 25 Mai 2019, 14th Annual conference of the European Human Behaviour and Evolution Association, Toulouse, France.

DeJong NS, Lange AS, Mendez CS, **Debache I**, Garnotel M, Zahariev A, Simon C, Bessesen DH, Bergouignan A. *Effect of One-Month Breaking Up Sedentary Behaviors on Metabolic Profiles*. The Obesity Society, November 2019, Las Vegas, NV, USA

DeJong NS, Schreck LS, Lange AS, Mendez CS, **Debache I**, GarnotelS, Zahariev A, Simon C, Bessesen DH, Bergouignan A. *Effect of short bouts of activity spread throughout the day on the components of energy balance*. The Obesity Society, November 2019, Las Vegas, NV, USA

Table of contents

PREAMBLE

The epidemic of physical inactivity

A lifestyle unsuited to human physiology

Humans, for most of their history, evolved to fit conditions in which sources of energy were scarce and foraging required considerable physical efforts [1]. Nevertheless, in a process known as the physical activity transition, modern humans progressively adopted a physiologically less suitable sedentary lifestyle [2]. The Neolithic Revolution (about 10,000 years ago), which transformed human economy from foraging to agriculture, induced a first massive sedentarization [3]. However, for most people, farming still required intense physical labor, which could have been even greater than before [4]. More recently, developed societies have become less and less dependent on physical labor for survival. With the Industrial Revolution beginning in the late 18th century in England and spreading throughout the Western world, machines have been progressively replacing human labor, resulting in a decrease in occupational physical activity [5]. More recently, technological advances introduced home appliances, such as washers and vacuum cleaners, which induced a further reduction in physical activity associated with household chores [6]. Likewise, the popularization of private motorized vehicles has further reduced physical activity required for travelling [2]. Despite a marginal emergence of leisure physical activity (for instance, American women reported engaging in leisure physical activity 1.1 hr/week in 1965 and 2.3 hr/week in 2010), active time saved with automation is still massively reallocated to inactive leisure, such as television watching [7]. Thus, humanity has embarked on a transition from a condition in which the ability to achieve high levels of physical was paramount for its survival to a condition in which physical activity is almost engineered out of all domains of life.

Whilst technology in developed societies made most of the physical labor needed in everyday life unnecessary, food became more and more accessible to the individual living in economically developed societies. The human body, through its metabolic mechanisms, has evolved to remain functional in situations of high energy expenditure and low food intake, but it is not adapted to situations of high food intake and low levels of activity. This situation favors positive energy balance, and hence unhealthy weight, and increased risks of physiological disorders, resulting in widespread morbidity at the population level [8].

These trends have been observed in the past half century in the Western world and, more recently, in large developing countries adopting this way of life, such as China, India and Brazil [6]. Thus, increasingly larger parts of humanity access a life condition in which practically no physical activity is needed in everyday life. As this very sedentary lifestyle spread out in low/medium economy countries, the burden of health conditions associated with physical inactivity is growing. Physical inactivity has become a major public health concern, and decision makers have developed various strategies to deal with it.

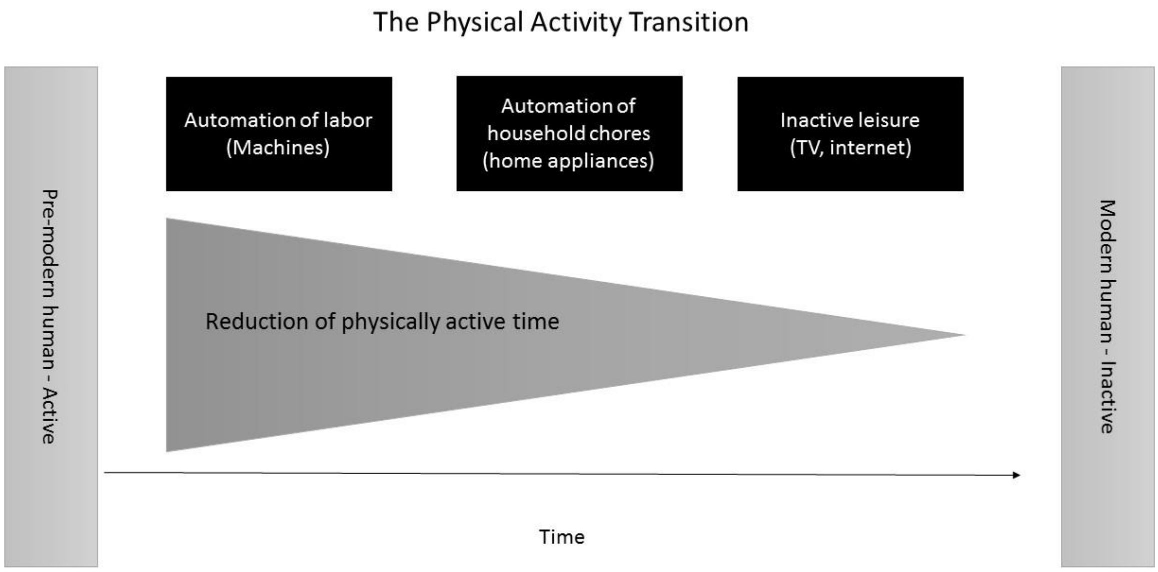

*Figure 1: **The physical activity transition***

Physical inactivity as a major cause of mortality

Physical inactivity i.e. activity level insufficient to meet the present recommendations (Scientific Report - 2018 Physical Activity Guidelines) has been recognized as one of the leading major causes of mortality over the last decade. It is associated with coronary heart disease, with an adjusted Risk Ratio (RR) of 1.16, type 2 diabetes (RR = 1.20), breast cancer (RR = 1.33 for women), colon cancer (RR = 1.32) and all-cause mortality (RR = 1.28) [10]. Worldwide, out of 57 million deaths in 2008, 5.3 million are caused by inactivity, resulting in a loss of 0.69 years of life expectancy at birth. Although this comparison has been criticized because of methodological issues, physical inactivity is thought to represent the same health burden for humanity as smoking [10]. Compared to inactive individuals, meeting the recommended levels of physical activity was found to result in about four years gain in life expectancy at age 30 years [11].

The prevalence of physical activity greatly varies across countries, and so do the health consequences of it. For instance, according to age-adjusted estimates of World Health Organization for 2016, prevalence of insufficient physical activity among adults (18+ years) was 36.8 % in high-income countries, 26.0% in middle-income countries, and 16.2% in low-income countries (Global Health Observatory data repository). These figures rely on self-reported measures of activity, which were shown to be significantly higher than those obtained using objective measures. For instance, a 2008 study showed that while 65% of Americans reported meeting recommended the activity guidelines, only 5% were found to meet them when using objective methods of activity assessment [12]. We observe the same cross-regional disparities when estimating loss of life expectancy, which ranges from 0.41 years in Southeast Asia and 0.95 in the Eastern Mediterranean countries (see map in Figure 2).

This pandemic of physical inactivity comes with an important economic burden, both indirect (productivity loss) and direct cost (healthcare). According to 'conservative estimates', the annual cost of physical inactivity amounted to $ 53.8 billion (international $, i.e. adjusted for purchase power differences) in 2013, of which 9.7 billion being paid by households [13]. According to another study, the annual per capita cost (both direct and indirect) of physical inactivity in North America ranged from $150 to $420 [14].

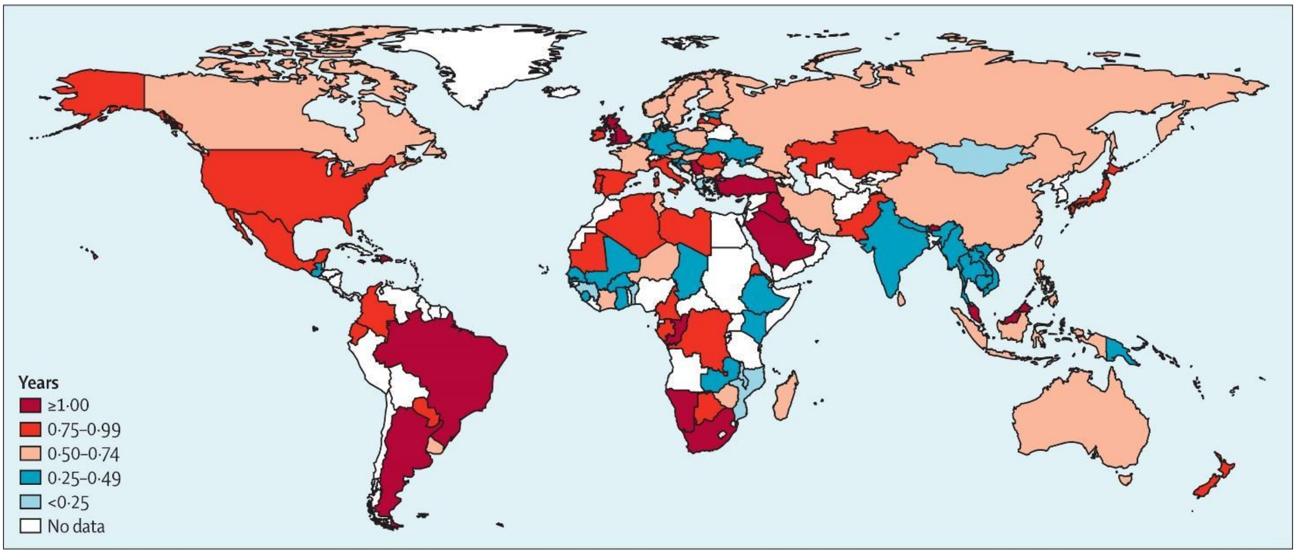

Figure 2: ***Gain in life expectancy at birth by country if physical inactivity were eliminated*** *(Lee et al. 2012)*

Physical inactivity as public health challenge

Thus, we can see that physical inactivity represents a considerable challenge for our health systems and economies. As the physical activity transition sweeps through the developing economies, the epidemic affects a growing number of people, and becomes one of the main challenges of the public health research of the 21st century.

Addressing the issue of physical inactivity, from the point of view of a researcher in public health, represents a twofold challenge. The first challenge consists in improving our understanding of the effects of different types of physical activity on various health outcomes, in order to refine the recommendations to the population, and for each individual. Thus, we must ask ourselves the following questions:

- **What are the physical activities that are relevant for health and how do they affect health? (Chapters I and II)**

As it is often the case in the medical field, research on physical activity comprises an epidemiological and a physiological facet. The first, at a larger scale, looks for specific patterns of physical activity and links them to health variables in the population. The second aspect aims at testing the existence of the observed associations in the epidemiologic and observational studies in controlled intervention settings and at understanding their underlying mechanisms. This thesis does not address the physiological question but will touch upon it incidentally, insofar as it is relevant for the epidemiological research design.

Addressing these questions using empirical research presupposes an ability to measure physical activity accurately. Consequently, we need to deal with this additional technological challenge by investigating the best way to measure activity in experimental settings.

- **How can we measure different aspects of physical activity in observational studies? (Chapter III)**

The second main challenge consists in looking for ways to promote compliance with these recommendations in the population. There exist psychological, cultural, social and environmental determinants of physical activity. By understanding the ways by which they determine physical activity, we can offer better strategies to encourage physical activity in the population. As it is often the case with human behavior, comprehending the entire system of determinants of physical activity is a daunting task; however, we can aim at understanding the effects of some major factors of the urban built environment on the physical activity of those exposed to them, with a particular emphasis on the public transportation infrastructure

- **Can we identify characteristics of the built environment that affect physical activity? How do these characteristics affect the different aspects of physical activity? (Chapters IV and V)**

Objectives and organization of the present thesis

The present thesis aims to address the main public health challenges posed by the physical activity epidemic, and it is divided in three main parts. After an introduction, the first part aims at shedding more light on the complex relationships between different aspects of physical activity and key health outcomes (Chapter I and II). The second part proposes an innovative methodology to measure physical activity in real-life experiments by means of accelerometers (Chapter III). The third part identifies possible fields of intervention for promoting physical activity by investigating the causal link between various characteristics of the built environment and physical activity (Chapter IV and V). The thesis concludes with discussing the findings and perspectives for future research.

INTRODUCTION

The aim of this introduction is to go over the concepts and terms that are used in this thesis, to position the research done within the existing state of knowledge, and to highlight gaps of knowledge that are important to address. Its structure reflects the layout of the thesis. The first section will address the different aspects of physical activity, in a broad sense, and their relationship with health. The second section discusses different methods for measuring physical activity and introduces methodological elements in physical activity recognition by means of body-mounted sensors (accelerometers and gyroscopes), which are the main tool used to objectively assess human physical activity in large epidemiological surveys. The third section will discuss, from a public health perspective, the main challenges in measuring the effect or the relationship between urban built environments and physical activity. The introduction concludes by briefly summarizing the ideas developed so far and by linking them, as objectives, to the chapters of the thesis.

Physical activity: New paradigms

Physical inactivity, as a health hazard, is not limited to the lack of physical exercise. Numerous studies, which I will discuss below, showed that sedentary behaviors could represent a hazard distinct from the lack of moderate or intense physical activity. In addition, recent findings suggest that the segmentation patterns of a given amount of activity time could also be relevant for health. This section discusses the concept of sedentary behaviors, their definition and their relationship to health. Likewise, evidence to the effects of activity segmentation are reviewed, and some conceptual and methodological challenges are addressed.

Definition of the terms of physical activity

This section uses a variety of terms referring to different types and aspects of physical activity. Here, I explain the terms used in this section, based on, unless otherwise specified, the definitions of the Advisory Committee Report of the 2018 Physical Activity Guideline.

Physical activity refers to “bodily movement produced by skeletal muscles that results in energy expenditure”. Physical activity can be categorized by its **intensity**. The most common way to quantify intensity is by using units of Metabolic Equivalents of Tasks (**MET**), one unit representing the energy expenditure while sitting at rest. For instance, walking at a 5 km/h requires about 3.3 METs. It is common to divide physical activity into four categories, based on intensity:

- **vigorous-intensity activity**, for 6 METs or higher (e.g. walking very fast, running, aerobic classes),
- **moderate-intensity activity**, for energy expenditure between 3 METs and 6 METs (e.g. walking at 5 km/h, vacuuming),
- **light-intensity activity** for energy expenditure between 1.6 and 3 METs (e.g. slow walking for leisure, cooking, standing while scanning groceries as a cashier).
- In the past, energy expenditure of 1.5 METs or less was often referred to as **sedentary behavior** or **sedentary activity,** but it is now accepted that sedentary behavior refers to energy expenditure of 1.5 METs or lower *while sitting, lying or reclining*. In this thesis, I therefore refer to low energy expenditure (≤1.5METs) in any body position, including standing as **low-energy activities.** As the fraction of time spent performing vigorous-intensity activity is very small, it is common in the literature to aggregate moderate- and vigorous-intensity activities into one category called **moderate-to-vigorous physical activity** (MVPA).

Derived from measures of physical activity by accelerometers, a **count** is recorded when the body motion observed exceeds a certain threshold fixed by the manufacturer of the device or the researcher. The number of counts over a period is a proxy to the energy expenditure during this period.

To refer to any of these activities or behaviors, I use the term **physical behavior**. The literature refers to the time proportion devoted to a certain physical behavior as **volume.** I refer to the distribution of the total time volume under consideration amongst the different behaviors as the **time budget** of physical behaviors.

The categories of behaviors presented here are often used in the literature, but it is needless to point out that any possible categorization of human behaviors that the scientist deems as relevant is possible. A very simple categorization would be active/inactive, while a very detailed one could comprise lying prone, lying supine, sitting, walking at different paces, jumping, bicycling etc.

The emergence of the sedentary behavior paradigm and

There is a broad consensus about the health benefits of regular MVPA and the health risks of the lack thereof [10, 15]. The findings of the very large body of evidence as to these effects are regularly summarized and translated into largely accepted practical guidelines to the population (e.g. 2018 Physical Activity Guidelines for Americans). However, there is need for distinction between the effects of different physical behaviors. While traditional research focused on MVPA, newer research, starting in the 2000's, stated that low-energy activity and too much sitting (as distinct from the lack of MVPA) could represent an independent risk factor with its own physiological mechanisms [1, 16, 17]. Based on evidence from large surveys, Healy and colleagues showed that total volumes of low-energy activities, as measured by a body-mounted accelerometer (here less than 100 "counts" per minute) was detrimentally associated with various metabolic risk biomarkers, independent of levels of MVPA, sex and adiposity [17, 18]. Likewise, Koster et al. found that large volumes of low-energy activities associated with higher mortality levels [19]. In 2007, the physical activity recommendation for adults from the American College of Sports Medicine and the American Heart Rate Association was updated, highlighting that the recommended MVPA levels had to come in addition to routine light-intensity tasks, such as casual walking and household chores [20]. In other terms, periods of physical inactivity should be avoided by remaining slightly active all along the day, as much as possible.

While these studies investigated the effects of low-energy activities (no matter the body posture), other investigated the effects (or correlations) of a more specific behavior, the sedentary behavior, on different health outcomes. The importance of posture allocation was clearly illustrated in an early study by Levine and colleagues, which showed that posture time allocation between quiet sitting and standing, measured by an inclinometer, was different across obese and lean groups, thus proposing an explanation to inter-individual variation in obesity [21]. Another study showed attenuated blood glucose excursions following an afternoon of standing at work compared to sitting at work [22]. Using iso-temporal substitution analysis on large amount of data in free-living conditions, Healy and colleagues showed that replacing sitting volume with standing volume was associated with lower fasting plasma glucose and triglycerides levels and higher high-density lipoproteins (HDL) concentrations [23]. From a physiological point of view, this definition accounts for the varying levels of isometric muscle contraction required for maintaining different postures in a motionless state. However, it should be noted that, in a recent study, the energy expenditure associated with maintenance of standing was found to be, on average, marginal compared with maintenance of sitting posture [24]. Likewise, some standing activities do not exceed 1.5 MET in the Compendium of Physical Activity, thus being undistinguishable from quiet sitting, as far as energy expenditure is considered [25]. Thus, the beneficial effects of standing versus sitting, insofar as they are real, may be independent of the energy balance and rather be related to contraction of muscles involved in weight bearing.

Thus, although MVPA volumes are still viewed as a chief determinant of health, the focus of research has recently shifted from MVPA to the avoidance of low-energy behavior. In the new paradigm, a healthy physical behavior should not only comprise sufficiently large volumes of MVPA, but also a reallocation of time from low-energy behaviors or sedentary behaviors) to light-intensity activity (or at least to quiet standing). As Maher and colleagues stated it, when considering sedentary behavior, light-intensity physical activity and MVPA, and keeping MVPA constant, saying that volumes of sedentary behaviors must be reduced is equivalent to saying the volume of light-intensity physical activity must be increased [26]. In fact, light-intensity physical activity has recently started to trigger scientific interest as a component of the time budget of physical behaviors [27]. Consequently, the concept of sedentary behavior can be meaningful only when considering *the entire time budget* (Figure 3) and the interdependency between its components, or when considering the *segmentation patterns* of the sedentary time. Both concepts are developed in the following sections.

Discrimination between different sedentary behaviors

Before discussing the ideas behind the concepts of time budgets and segmentation patterns, an important issue needs to be pointed out. Investigating sedentary behavior as a stand-alone behavior allowed an important advance in our understanding of the etiology of various diseases. However, sedentary behaviors are regarded in the literature as a single component. An interesting question would be to discriminate further between different sedentary behaviors. Notably, lying and sitting are distinct behaviors considered as sedentary, but the literature hardly addresses the question as to whether their effect on specific health outcomes is different. One of the objectives of Chapter I will be to break down sedentary behavior into two smaller components – sitting and lying – and investigate their separate relationships with some key health outcomes.

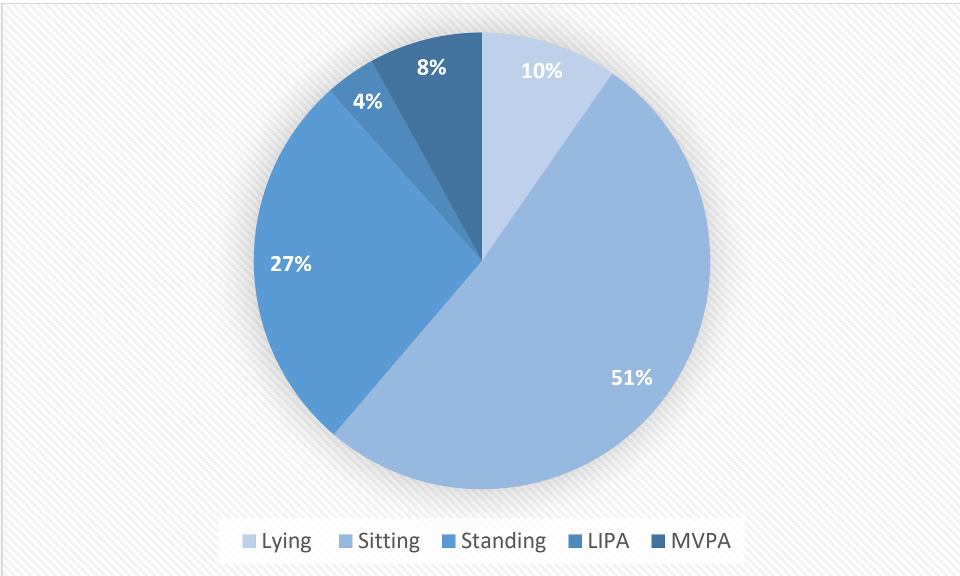

Figure 3: ***Daily time spent in different physical behaviors*** *as measured in a population of adults in the region of Paris, France [28]*

The idea of a time budget of behaviors

The previous paragraphs emphasized the importance of accounting for the volumes of different types of behaviors when investigating the relationship between physical activity and health. Yet, the importance of a behavior volume is always related to the volumes of other behaviors. Increasing the volume of one behavior comes at the expenses of one or several other behaviors, since all volumes always add up to the total time under study. Thus, rather than talking about the relevance of the volume of such or such behavior, one should view the set of behaviors under study as a whole, that is as a *composition* or *budget.* The challenge and the necessity of such an approach becomes clear within the analytical framework of linear regression models, which are the common way to analyze relations (causal or not) between variables in epidemiological studies. Estimating the effect of each volume by adding all volume components to the model as regressors is impossible, since these are linearly dependent (i.e. perfectly collinear). Analyzing each volume in a separate linear regression model is possible but also problematic, since the effect associated with a volume can actually be attributed to another volume absent from the model (i.e. the error is correlated with the regressor). For instance, integrating sedentary time alone might results in an effect estimate on a health outcome that is due to the absence of light-intensity physical activity time, as one comes often at the expense of the other. Hence, to deal with the compositional character of volumes in an analytical framework, two methods have been proposed in the literature: iso-temporal substitution and compositional analysis.

Iso-temporal substitution analysis consists in estimating the effect of a component S_b in a budget (volume devoted to a behavior b) by integrating all other components $S_{\bar{b}}$ (any volume devoted to a behavior other than b) as regressors into the regression model. S_b being the complementary part, it is implicitly present in the model. The coefficient for every term $S_{\bar{b}}$ in the model corresponds to the effect of substituting a time unit of S_b by a unit of $S_{\bar{b}}$ on the response variable of interest [29]. Compositional analysis, instead of considering every volume as a data-point in the real domain, views the different volumes as a vector belonging to the compositional simplex, known as the Aitchison simplex (see Figure 4). It maps the vector in the simplex to the real domain by applying an appropriate transformation (e.g. the isometric log-ratio transformation in [30]) in order to integrate it as a predictor to the regression model, and converts the coefficient back to the compositional domain. Thus, changes in the health outcome can be predicted for any change in volume composition [31].

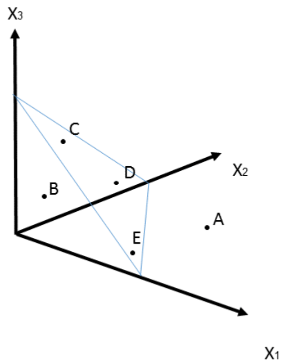

Figure 4: ***Representation of a 2-D simplex in the real 3-D space****. All three-part compositions (x1, x2, x3) lie on the triangle (simplex), because of the constraint of adding up to 1. Points C, D, and E are therefore possible compositions, but not A and B. Compositional analysis acknowledges the constrained sample space of the compositions (and thus their interdependency). In contrast, in classical regression models view each part belongs to the unconstrained real space.*

Iso-temporal substitution and compositional approaches are both addressing the methodological problems raised by the compositional character of behavior volumes. Using compositional regression is arguably more appropriate from a mathematical perspective, as it respects the mathematical properties of compositional data [31]. Nevertheless, iso-temporal substitution analysis is often preferred in the literature, because its coefficients are easily interpretable, and the analytical framework remains closer to the classical regression analysis with which most researchers are familiar.

In this thesis, we adopt a compositional approach in our analysis, accounting for the interdependency between the components of the time budget. Answering the question "Is there an effect of volume of sedentary behaviors on health that is not the effect of a reduction of other behaviors (e.g. light-intensity physical activity or MVPA)?" is clearly a logical fallacy, as all volumes are parts of the same whole. Instead, we ask the questions: "How does a specific time budget affect a certain health outcome?" or "How does reallocating time from a behavior A to a behavior B affect a certain health outcome?". Analyses are done using both methods introduced here, in order to have both a better understanding of the phenomena observed and a better comparability with other studies.

The importance of behavior segmentation for health

In the previous section, we introduced the relevance of a refined typology of physical behaviors in health research. As said above, a physical behavior refers to an ensemble of particular movements and postures of the body performed at a certain time. For practical reasons, observational studies consider behavior as discrete over time; a behavior is imputed to short time intervals, whose length typically span from 0.1 second to one minute. Thus, an epoch is the base time unit for which the researcher has a defined behavior. Whichever interval length is chosen, research mentioned above considers solely the volume associated with a certain behavior over a certain period, such as a day or a week. In other words, each time interval is viewed as independent from the preceding and following ones and contributes equally to the total volume. Although this approach based exclusively on volumes is easy to use, it does not necessarily reflect the underlying physiological processes. Indeed, the physiological significance of a behavior at a certain time most likely depends on the behaviors performed before and after it. To take a trivial example, a 1-minute interval of MVPA is arguably not the same when following 60 intervals of rest than when following 60 intervals of MVPA. Thus, as we will see, researchers are also interested, in addition to volume, in the *accumulation patterns* (or *segmentation patterns)* of a behavior over time and its relationship with health outcomes.

Regarding MVPA, the idea that segmentation patterns are relevant for health has long existed in the physical activity guidelines. Until 2006, physical activity guidelines recommended performing exercise in continuous bouts[1] of at least 20 or 30 minutes [20, 32]. However, using objective accelerometer measurement of physical activity, Ekelund et al. investigated whether performing the same MVPA volume in continuous bouts (of at least 5 or 10 minutes) had a specific impact on different metabolic health variables [33]. Results could not verify the hypothesis of an effect of period length on adiposity and risk factors associated with metabolic syndrome. Jefferis and colleagues reached the

[1] Bouts, as a central notion in segmentation analysis, will be presented later. A separate discussion found in the annex to this thesis elaborates in detail on this widely used yet poorly defined measure. At this stage, it is sufficient to say that a sedentary bout is a period in which the most frequent observed activities were sedentary.

same conclusion studying a sample of 1009 British men [34]. Large-scale data from American and Canadian surveys yielded conflicting results. Some studies suggested that sporadic MVPA was beneficial for health, but MVPA performed in bouts of 10 minutes or longer had an additional value [32, 35, 36]. Other studies suggested that any second of MVPA had the same effect [37, 38]. Some of the discrepancy in the evidence could be attributed to different methodological approaches [34]. As we will see, methodological considerations in measuring segmentation have a considerable impact on the conclusions inferred from the data. Nevertheless, the 2018 Physical Activity Guidelines for Americans omits the recommendation to perform MVPA bouts in long bouts and mentions only total activity volumes over the week (Scientific Report - 2018 Physical Activity Guidelines). Thus, it seems that a large part of the scientific community agrees as to the fact that individuals can meet the recommended MVPA following any pattern of accumulation of their choice.

The concept of segmentation was also applied to sedentary behaviors. Healy and colleagues popularized the idea that breaking up total sedentary time into many short periods in everyday life could be beneficial for health [39]. Using real-life data, they found that, independent of the total time spent in sedentary behaviors and MVPA, the number of interruptions of sedentary behavior (mean = 600/week, standard deviation (SD) = 155/week) was positively associated with adiposity, and negatively with fasting plasma triglycerides and plasma glucose levels 2 hours after a meal (an index of glucose control). The effect magnitude for both variables was between 0.16 and 0.18 SD for one SD increase in the number of interruptions.

Several experimental studies in *controlled settings* found that interrupting sitting time with light-intensity physical activity periods significantly reduced postprandial glucose and insulin levels in both lean and overweight/obese adults [40–43]. For instance, Dunstan and colleagues measured the 5-hour incremental area under the curve (i.e. integrated levels over time relative to the base level) for plasma glucose and insulin after an oral dose of carbohydrate in overweight/obese male and female adults in three conditions in a crossover design study: (a) uninterrupted sitting, (b) sitting with 2-minute interruptions of light-intensity activity every 20 minutes, (c) sitting with 2-minute interruptions of moderate-intensity activity every 20 minutes. Breaking up sedentary behaviors with frequent light and moderate active bouts respectively reduced glucose area under the curve by 24.1% and 29.6% compared to the sedentary control condition. Post-prandial insulin concentration was reduced by 23% following both the light and moderate breaks. The effect of breaking up prolonged sitting time with standing breaks is still unclear. While a study could not elicit a reduction in postprandial glycaemia in *non-obese* of both sexes with standing breaks [44], another one on *obese* adults of both sexes observed a 11.1% reduction in post-prandial glycemia when alternating 30-minutes periods of sitting and standing, but no effect on serum insulin and triglycerides [45].

Results on segmentation of sedentary behavior raise questions as to the relevance of sedentary volume itself. Results presented in the previous sections on volumes concluded that avoidance of sedentary behaviors through increase in light-intensity was beneficial for health. However, in light of the findings on segmentation, can we say that volume of SB matters at all? Contrarily, can we say that segmentation matters at all provided that SB volume is reduced to smaller volume? And if both volume and segmentation matter, what is the effect magnitude of each?

To elucidate this question, research must systematically consider both volume and segmentation. Yet, on one hand, the major controlled studies focusing on segmentation mentioned above have *deliberately* maintained fixed sedentary volumes across groups with different segmentation patterns in order to isolate the effect of segmentation [40–43]. On the other hand, the empirical strong negative correlation between sedentary volume and segmentation (active people tend to exhibit both strong segmentation patterns and small volumes of sedentary behavior) can make it hard to accurately measure both effects [23]. Thus, to date, there is still a lack of clear evidence from heterogeneous populations on the share of each of these dimensions on key health variables.

With volume and segmentation being relevant for key health variables, analyses must integrate measures of both dimensions into their models in order to assess their distinct effects clearly. How can we quantify and characterize segmentation patterns independently from behavior volumes? The following section addresses this issue.

How to measure segmentation?

Previously, several studies emphasizing the importance of segmentation of physical behavior for health were cited. However, measuring segmentation can be tricky. A study comparing cross-sectional associations between sedentary bouts and various health indicators showed how different measures of segmentation used by different authors, which I will expose below, could yield different results [46]. As pointed out by Tremblay and colleagues., the lack of consensus on the complex derivation procedures of temporal segmentation patterns in sedentary behaviors was reported by researchers as one of the main obstacles to progress in the field of sedentary behavior research [47].

Counting breaks and transitions

The simplest way to measure segmentation of sedentary time is by counting the number of breaks. Of course, as the number of breaks is usually correlated to the total sedentary time, the break count should be standardized by the total sedentary time, yielding a relative measure such as breaks per sedentary hour. The influential study by Healy and colleagues. used number of breaks as a fragmentation measure but did not standardize it by sedentary time [39], hence causing potential bias to the estimated effects.

However, the major shortcoming of break count as measure of fragmentation, even when standardized, lies in the fact that it ignores the *duration of breaks* and of the episodes *between the breaks*. Let us think of a person who performs one hour of sedentary behavior and one hour of light-intensity physical activity. In scenario (a) the individual alternates periods of 10 minutes of each behavior. In scenario (b), he/she alternates 2-minute episodes of each five times and performs then 50 minutes of sedentary behavior and 50 minutes of light-intensity physical activity (Figure 5). In these two cases, the person has the same volume of SB and light-intensity physical activity (an hour of each), and the same number of breaks (6 breaks). However, one could easily argue that the two cases reflect two different physiological realities: in scenario (a) there exists no *prolonged* sedentary bout (>10 minutes), and so, the physiological mechanisms associated with sedentary behavior may not be ‘switched on’, while we do observe 50 minutes of continuous sedentary behavior in scenario (b). The importance of duration is all the more obvious considering that most of the evidence to segmentation concludes specifically that *prolonged* sedentary episodes represent a health hazard [39–41].

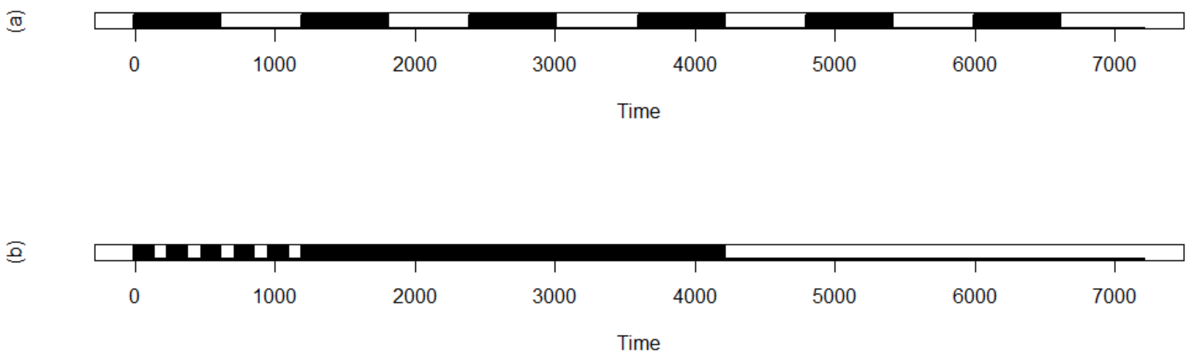

Figure 5: ***Two scenarios of successive episodes of sedentary behavior (black) and light-intensity physical activity (white)****. The same volumes follow different segmentation patterns.*

Breaks from behavior are in themselves *transitions* from a behavior/state to another. Considering sedentary behavior, any break from this behavior is a transition between sedentary behavior and another behavior, e.g. to standing, light-intensity physical activity or MVPA. Whereas breaks ignore the behavior to which the subject is changing, transitions account for both the origin behavior and the target behavior. As the number and variety of behavioral categories increase, accounting for the target behavior is important. Transitioning from sitting to quiet standing has not necessarily the same physiological significance than the transition to light-intensity physical activity. Counting all possible transitions can prove useful in phenotyping an activity profile. When we divide all possible transitions from a state to another by the total number of transitions, we obtain a Markov matrix for the observed *probabilities* of transitioning from a state to another. However, the *number* of transitions can bear a physiological meaning, independent from the individual's segmentation profile, as it points to the total muscular force needed to transition between different behaviors over the monitoring time. Here again, we see the importance of controlling for volumes of all behaviors when performing segmentation analysis.

In short, breaks and more generally transitions are an interesting and important metrics for physical activity profiling. However, since they ignore the duration of the behavior episodes and the breaks, they are insufficient to capture the effects of segmentation patterns in a satisfactory manner; in fact, by definition, duration is an essential aspect of segmentation.

Capturing distribution of bout lengths

In order to account not only for the number of breaks but also for the duration of the behavior episodes and the duration of breaks between them, the concept of bout is largely used in the literature [47]. There is no consensus on the definition of a bout and its parameters, and, although it heavily affects the conclusions drawn from observational data, it is still largely overlooked in the literature [47–49]. To be as general as possible, all definitions agree on the following: a bout of behavior is a period of a certain minimum length during which no or only negligible interruptions (typically smaller than one or two minutes) by other behaviors is observed. Hence, identifying bouts is equivalent to looking at the behavior sequence at hand in a lower resolution, considering relatively long, roughly homogenous periods of behavior. Considering behavior bouts instead of considering any behavior episode is already a first step in segmentation analysis: we retain long, nearly homogenous behavior sequences, while discarding episodes that are too segmented or too short to be of any relevance.

The bout detection algorithm outputs a set of valid bouts, of a minimum but variable length, for the behavior under consideration. For instance, for a minimum bout length of 5 minutes, we can have a set of bouts with lengths (in minutes) {5, 8, 5, 20, 110,..}. The lengths of these bouts are the information that is exploited in order to characterize the segmentation profile of an individual over a certain period of time. This section discusses different metrics mentioned in the literature; they correspond to different approaches to segmentation and their relevance for health strongly depends on the parameters of the algorithm used to detect the bouts.

A first approach consists in considering the time spent in bouts as the only 'valid' time of the behavior in question, while sequences that do not belong to a bout are disregarded as incidental. Consequently, tenants of this approach will look at the sum of bout durations as the only valid behavior time [50]. This approach was implicitly adopted by major works on the importance of segmentation of MVPA, e.g. [32, 35]. For instance, under the assumption that sedentary behavior affects a given health outcome only after a minimum duration (i.e. the minimum bout length), we are interested in the total duration of the sedentary behavior time performed in any bout, while excluding sedentary behavior performed in brief episodes. In the same line of reasoning, when looking into the volume of a behavior, the ratio between total time (performed in any episode) and total time spent in bouts can be an interesting segmentation index: a high ratio points to high level of segmentation, meaning that larger parts of the behavior time is performed in long episodes.

A second approach focuses on the empirical distribution of bout durations and on metrics that can capture the shape of this distribution. In this approach, the distribution of the durations reflects the segmentation profile of an individual and therefore might contain meaningful physiological information in it [51]. Let us consider the example plotted in Figure 5 and suppose that we set our minimum bout length to be of one minute. In scenario (a) six activity bouts of 10 minutes are interrupted five times by inactivity bouts of 10 minutes. In scenario (b) six activity episodes are also interrupted by five inactivity episodes, but the *distributions* of these episodes differ. We have the same number of breaks/transitions, the same total time of bouts, but different bout distributions. Insofar as one long sedentary bout together with five small bouts reflects a different physiological reality that six bouts of the same intermediate length, this difference is captured by properly characterizing the empirical distribution of the bout lengths. To re-emphasize the importance of bout parameters in segmentation profiling, it should be noted that setting our minimum bout length to five minutes would result in yet another picture: scenario (a) would contain six valid bouts adding up to one hour, but (b) would contain only one bout with a total duration of 50 minutes.

Empirical observations show that the distributions of bout durations roughly follow a power law ($p(x) = a \cdot x^{-\alpha}$, where x is the bout duration, $p(x)$ is the density and a a constant). Simply put, there are few long bouts for many short ones. Thus, Chastin and Granat proposed to estimate the exponent of the power law to quantify segmentation: a high α points to a high level of fragmentation, i.e. the volume is accumulated through many short episodes. Likewise, they proposed the Gini index, originally measuring economic inequality, as segmentation index: when the Gini index is high, the volume is made in a few bouts, pointing to low segmentation and vice-versa (see Figure 6).

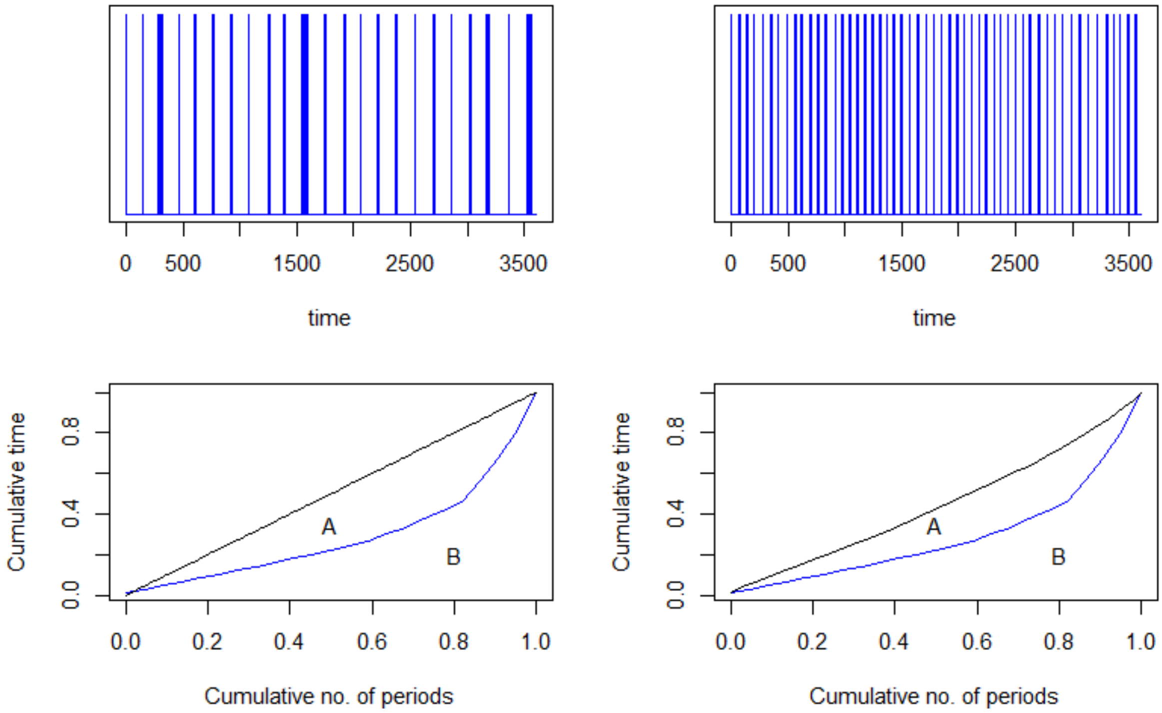

*Figure 6: **Succession of periods of active and inactive time** (blue and white) over 3600 epochs (upper part) and the corresponding cumulative active time by no. of periods (lower part) in two scenarios (left and right) drawn from different distributions of activity durations. In both scenarios, the total active time is equal, but in the right-hand side activity is less segmented, with fewer, longer activity periods than in the left-hand side. The blue curve is called the Lorentz curve; the closer this curve to the diagonal black curve (right-hand side compared to left-hand side), the more even the distribution, the more segmented the activity time. The Gini index is equal to A/(A+B); a low value (right) points to stronger segmentation and a high value (left) to a weaker segmentation.*

There exists no consensus on how to extract information regarding segmentation patterns from bouts and literature is scarce on this topic. Whereas the first approach exposed here uses bouts to quantify *which* of the behavior is performed in prolonged periods, another looks at bouts to quantify *to what extent* the behavior is performed in prolonged periods. The first approach, by discarding short, sporadic periods of behavior can overlook an important constituent of an activity profile, while the second approach, looking at sufficiently small bouts, is able to capture a more general picture. Yet, properly characterizing a distribution with meaningful, interpretable metrics is a challenge that has not been properly addressed in the field yet. Despite the consequences that choices concerning segmentation measurement have on the conclusions drawn from activity data, only a few studies discussed this issue [46, 52].

Unlike Alternburg and Chinepaw, who suggest harmonizing methods of bout analysis for better transparence [50], I think that such a harmonization needs to be preceded by a discussion on the definition of the bout, the health outcome under study and the way that the metric chosen is thought to model it. Different parameters of the bout detection algorithm result in different distributions of bout durations and consequently call for different bout analysis strategies. Likewise, each research questions requires specific methods. Standard guidelines regarding bout analysis are important to improve comparability, but they need to emerge from a systematic discussion of the strategic choices made by the authors. To this day, apart from a few articles pointing to the discrepancies in methods of bout analysis and their potential effect on epidemiological results, such a systematic discussion is still missing.

Bouts are important for segmentation analysis as they retain significant behavior episodes only. In fact, when we ignore micro-sequences, we obtain the greater picture of the activity profile over the day. However, micro-sequences that cannot be captured in bouts can have a relevance for health as well. Thus, characterizing the distribution of sequences *of any length* in the same way as we do for bouts can reveal new aspects that should not be ignored. Nevertheless, approaches to segmentation based on bout analysis only have dominated the literature so far,

In this thesis, I chose a small bout length (1 minute) in order to limit information loss; choosing longer bouts (e.g. 10 or 20 minutes) would have discarded many micro bouts (<2 minutes or <4 minutes) that may be of relevance. In addition to bouts, the distribution of sequences of any length (even micro-sequences, e.g. 1 or 5 seconds) is investigated. With this approach, I highlight both the greater lines of daily activity over the day with bout analysis, while keeping in mind small bursts of behavior.

Summary

Apart from the volume of MVPA, this first part of the introduction discussed two aspects of activity that need to be considered: volumes of physical inactivity/sedentary behavior, which have drawn much attention in addition to the traditional emphasis on volume of MVPA, and the segmentation patterns of these two volumes. First, I emphasized the importance of investigating the effects of different volumes (sedentary behaviors, standing, light-intensity physical activity, MVPA…) as different part of the *same* time budget (i.e. components of a composition). The distinction between sitting and lying should also be developed and their distinct effects investigated. Second, I explained that the segmentation pattern and the volume of a behavior are to be modeled as two distinct effects. In addition, different approaches to segmentation were discussed. Despite abundant literature on the question of physical behaviors and health, it is still unclear whether there are two independent effects of volume and segmentation of physical inactivity/sedentary behavior (Scientific Report - 2018 Physical Activity Guidelines). To clarify this question, studies need to discuss and integrate metrics of segmentation and volume systematically into their models. Moreover, it is not clear whether quiet standing, as opposed to sedentary behavior, is beneficial to health as segmenting behavior or as volume.

Detecting physical behavior with accelerometers

In order to understand the effects of the various aspects of physical behaviors on health, as discussed in the previous section, researchers need to set up experiments and study protocols in which physical behaviors can be measured. Epidemiological studies focusing on free-life conditions have long relied on questionnaires to assess physical activity levels [12]. Although very practical for its low cost and easy implementation, this method was criticized because subjects are known to be inaccurate in their reports regarding volumes of activity [12, 53]. It is needless to emphasize that this recall bias is probably even more severe when investigating the exact patterns of segmentation and time budgets including refined categories of behaviors (e.g. standing or sitting), whose importance was addressed in the previous section. As these dimensions of physical activity gained importance in the scientific scene, the need for a precise, continuous monitoring tool becomes more acute.

In the late two decades, an increasing number of studies relied on accelerometer devices mounted on the subject's body to assess physical activity. One of the early and most notorious study was the 2003-2006 NHANES study, which aimed to estimate physical activity habits of Americans and collected accelerometer data from nearly 15000 individuals aged six years and older [54]. Following the NHANES study, accelerometers gained popularity, their price and size reduced, and they are considered today a standard tool for physical activity assessment in epidemiological studies [55, 56]. In this section, I will introduce the core topics of physical behavior detection with accelerometer. I will discuss technical specifications of accelerometers, the challenges in signal analysis and models for recognition of human activity using these devices. Chapter III of this thesis will report the results of my detection algorithm on empirical data and will further discuss some of the major challenges in this field.

Accelerometer: working principles

Accelerometers are sensors that measure proper acceleration and, consequently, when attached to another body, the acceleration of this body. The accelerometers contain the piezoelectric materials (such as crystals or certain ceramics) or micro-machined microelectromechanical systems (MEMS), which produce electricity in response to mechanical stress along a certain axis. The electricity produced is measured by the device at a certain rate and converted to data bytes (the output) corresponding to the acceleration magnitude. These data are stored in the device's memory and can be accessed at a subsequent stage by the user.

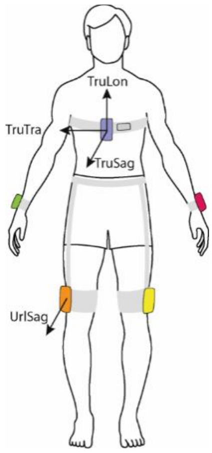

*Figure 7: **Tri-axial accelerometer sensors mounted on different parts of the body**. The three axes (sometimes called longitudinal, sagittal and transversal) are shown for the device on the man's chest.*

The acceleration is typically read in m/s^2 or in acceleration due to gravity, g (≈9.81 m/s^2). Accelerometers measure the acceleration due to Earth's gravity, and hence will read 1 g (or -1 g, depending on the convention used) along an axis when they are at rest on a surface, perpendicular to the ground, and 0 g when parallel to the ground. When the sensor is in movement, the accelerometer will add the corresponding acceleration to the acceleration due to Earth's gravity.

Accelerometers have three important specifications that need to be considered. First, the amplitude of acceleration that can be read differ across devices, and typically goes from ±2 g in the simple devices to ±16g in the performing ones. When this range is too small, the device might not be able to record peak accelerations associated with extreme activities. Second, devices measures acceleration at discrete time points, typically varying between 30 and 100 Hz. In order to capture all information of human motion, the sampling frequency must be at least twice as high as the highest frequency found in the data (Nyquist-Shannon sampling theorem). For instance, if the frequency of human movements does not exceed 15 Hz, we would ideally need to sample acceleration at 30

Hz. Regarding these two specifications, recommendations suggest an ideal range of ±12 g when the sensor is mounted at the ankle, i.e. where acceleration is strongest, and a much smaller range when mounted on the upper body or the head (±4 g or ±5 g) [57]. However, good results can be obtained even when the amplitude is suboptimal [58]. Regarding sampling frequency, Bouten and colleagues recommend at least 20 or 30 Hz [57]; in fact, most of the information concerning human locomotion is contained in frequencies up to 10 Hz [59]. Third, the number of axes used varies across devices; early commercial accelerometers used in health sciences had one axis, but more recent devices read acceleration along two or even three orthogonal axes [60]. A higher number of orthogonal axes allows a better estimation of the device's orientation and the direction of the motion relative to a reference frame (Figure 8). Three-axis accelerometers have progressively become standard in in the beginning of the 2010's, as they allow a significantly better detection of the orientation and motion in space [61].

Processing raw accelerometer data

Windowing

Simply put, the objective of algorithms for activity detection is to apply a function to the accelerometer signal that outputs the corresponding activity behavior. Since human movements are performed in lower frequencies than the sampling frequency, we need to group raw signals in relatively long 'windows' in order to be able to translate them accurately to human activity. A first, common approach bins the signal into sliding windows of the same length. As the window length increases, patterns specific to certain behaviors emerges more clearly and discrimination between behaviors becomes easier. On the one hand, long windows help detect low-frequency behaviors. In addition, when we have high-frequency behaviors that do not manifest themselves in a clear way, long windows, by containing several samples of this repetitive behavior, increase our confidence in identifying them correctly. On the other hand, longer windows result in a lower resolution, as a single behavior is assigned to a long time laps, and therefore may not be adapted to short sporadic behaviors. Thus, window lengths typically vary between 1 and 10 seconds across studies (although much longer windows do exist), and the accuracy of activity recognition was shown to depend on both window length and the specific behavior to be detected [62, 63]. In some studies, the sliding windows overlap to a certain extent (e.g. to 50%) [58], in order to make sure that entire patterns are detected and not split up over two separate windows.

The necessary trade-off between window length and resolution has led researchers to look for other segmenting methods. Some researchers have recommended that the window length be determined dynamically, either by external signals (e.g. when the GPS points to a dislocation) or by features of the accelerometer signal itself [62, 64]. Despite its advantage, dynamic windowing is less frequent in the literature, as it requires complex, uncertain algorithms or external input.

The segmentation discussed here should not be confused with behavior segmentation discussed in the first section. Behavior segmentation relates to the patterns by which humans distribute their behavior over time and bears a physiological significance per se. Here, segmentation merely relates to a technique of signal analysis necessary in order to recognize the human behavior performed.

Posture and motion intensity

As explained above, the accelerometer readings consist of two components: the static acceleration and the dynamic acceleration. The dynamic acceleration corresponds to the intensity of body movements, and the static acceleration to the orientation relative to gravity. Several approaches exist for separating the dynamic acceleration from the total acceleration. Van Hees estimates motion intensity by taking the Euclidian norm of the signals and subtracting 1 g due to gravity($\sqrt{a_1^2 + a_2^2 + a_3^2} - 1$). The same author suggests that a high-pass filter applied to the signal (e.g. Butterworth filter) can improve motion measurement, given that typical human motion is performed at a higher frequency than 0.2 Hz [65]. Others proposed motion metrics based on the standard deviation of the signal over the epoch [66]. Likewise, the mean of the signal or a low-pass filtering applied axis-wise can isolate the gravity component; once this component is given, simple trigonometric formulae can yield the inclination angle of the body relative to gravity.

Most epidemiological studies conducted on physical activity used activPAL or ActiGraph devices [61]. These devices yield 'activity counts', which represent another method of measuring motion intensity over a certain epoch (typically 1 to 60 seconds). This method is proprietary to the manufacturer [67], although Brønd and colleagues succeeded in estimating it [68]. Because of the immense popularity of the devices, activity counts have become a standard measure of activity intensity that has occupied much of the discussion in the field so far. These counts were criticized for being opaque and lacking straightforward meaning; in addition, they are not equivalent across devices [66]. The ubiquitous use of the counts has led to a confusion among researchers and to a dependency to manufacturers using it. Therefore, it seems that research should focus away from counts and use clear, interpretable metrics computable from the raw data.

This basic processing of the data provides the experimenter with measures of motion intensity, which can be used as such or converted to energy expenditure using well-known formulae [61] as well as with measures of inclination, or body postures. Using cut-points, basic behavioral categories can be created out of these measures (e.g. inactivity, very light intensity light-intensity and moderate-to-vigorous behavior, depending on a measure of motion intensity, or standing or sitting, depending on leg inclination). As researchers aim to use refined categories of behavior, they will need to extract complex features from the raw data, which will be fed into advanced algorithms for behavior classification. The next section briefly presents these features.

Time and Frequency domain features

Over a certain window, features of the signal can be computed in the time domain and the frequency domain [58]. Signal data are recorded in the time domain, with each data point corresponding to the acceleration along a certain axis at a certain time. With the Fourier transform, this time series can be described as the sum of a series of sinusoidal waves of different frequencies. Depending on the nature of the signal, each wave will have a certain amplitude. Thus, each time series, belonging to time domain, can be transformed into a series of amplitudes in the frequency domain, representing how much of each frequency is found in the signal.

Features computed in the time domain, i.e. on the original series, aggregate information about acceleration values over time. They are typically the mean, the standard deviation, different quantiles, correlation between signals, etc. For instance, a high standard deviation is due to strong oscillation over the time window, hence pointing to a movement of the sensor. Features computed in the frequency domain aggregate information about the periodicity of the acceleration values. They are typically the maximum amplitude (which is the strongest frequency in the signal), entropy (measuring the disorder, i.e. how clear the periodicity in the signal is), the total energy (i.e. the sum of all squared amplitudes), etc.

The Fourier transform provides information about frequencies for the signal of the whole signal, but not about *where in time* a certain frequency is to be found. Yet, reducing window size for more time precision results in loss of information about lower frequencies. To cope with this limitation, a few studies succeeded in extracting features from a wavelet transform of the data [69, 70]. A wavelet transform yields a representation of the frequency amplitudes *in time*, which enables a better identification of the time at which an activity, corresponding to a certain frequency level, is performed.

Creating an algorithm for behavior detection

In the previous section, we have seen how raw data is processed to create complex features, aiming at capturing the information contained in a certain window/epoch. Once these features are ready, they will be fed into algorithms that will decide, based on the information extracted from the acceleration, what is (the likeliest) physical behavior associated with it (Figure 9).

Relatively simple algorithms for classification can be created by human intelligence. As explained above, we can, for instance, detect movement and inclination of the body using formulae, and impute simple behaviors using this information. However, as the behaviors of interest become more complex, machine learning models represent a very popular and performant alternative [71, 72].

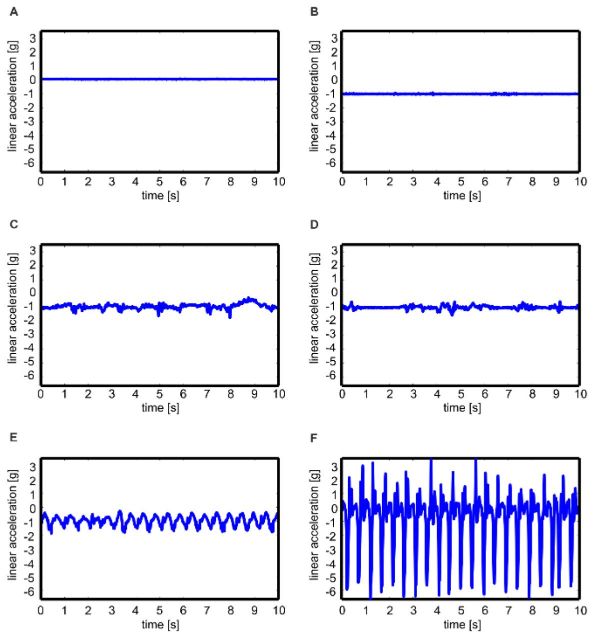

*Figure 8: **Example of raw accelerometer signal** (hip, sagittal axis) for different behaviors. (A) Lying, (B) Standing, (C) Vacuuming, (D) Sweeping, (E) Walking, (F) Rope jumping. Models aim to classify behaviors based on the difference in the patterns of the signals (Image from [73]).*

Machine learning models learn in an unsupervised or supervised way [74]. In our case, unsupervised learning means that the machine will creates its own categories of physical behavior by looking for the most meaningful classification strategy based on the data given. Supervised learning, on the contrary, requires a learning set, in which both the statistical features and the target observed behaviors ('labels') are given. In a training phase, the machine will look for a classification strategy -- based on the features -- that best matches the given observed behaviors. Once the training is done, the optimized strategy will be used to classify features, this time without knowing the real behaviors. A few attempts were made to use unsupervised learning for physical activity detection with accelerometer data [75, 76], but they remain an exception, mainly because investigators aim to obtain pre-defined, well-known categories of physical behavior for their subjects. Thus, supervised learning in human activity recognition has become very popular in the field and has yielded excellent results, as far as the internal validity is concerned [61, 71].

Human activity recognition using supervised machine learning can be divided into two main approaches. The first, more common and traditional, relies on handcrafted features extracted from the raw signal, such as those introduced in the previous section (mean, variance, max. amplitude of the Fourier transform...). Classification models (e.g.

decision tree, support vector machine, naïve Bayes classifier…), are trained to predict the correct behavior for a sequence of the signal using the features extracted from the signal [77] (Figure 10). The second approach relies on artificial neural networks (in a method referred to as deep learning) using the raw signal itself. The signal is fed directly into the neural networks, whose successive layers of neurons are trained to detect meaningful features of the data, eventually allowing for a classification of the signal sequences by behaviors [78](Figure 11).

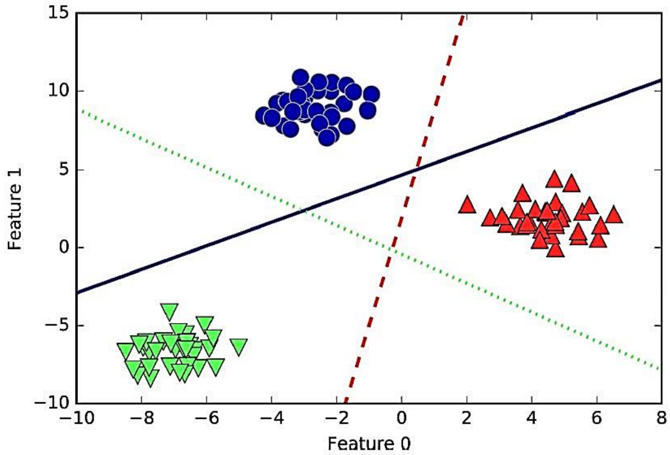

Figure 9: ***A classification model based on handcrafted feature extraction****. Each observation belongs to one of three classes (here physical behaviors), marked with different shapes and colors. Based on two features defined and computed by the researcher (feature 0 and feature 1 - e.g. the mean and the median of the signal over the window), a good classification model draws decision boundaries that separate the plane into decision areas corresponding to each class (Adapted from [79])*

Whereas in the first approach models based on handcrafted features rely on *a priori* domain knowledge of the data to create features, deep learning models have the advantage of being able to detect features that the researcher has not thought of. In addition, features crafted by humans are usually simpler than features created by a deep network of neurons [80]. As researchers aim to discriminate between similar behaviors, they must think of increasingly complex features and in such a large number that the time needed for computation and training grows exponentially, resulting in thousands of features which must be fed into the classification models [81]. *Feature selection* among those initially computed can speed up training process and improve classification accuracy, but designing a good selection algorithm represents a considerable challenge in itself [74]

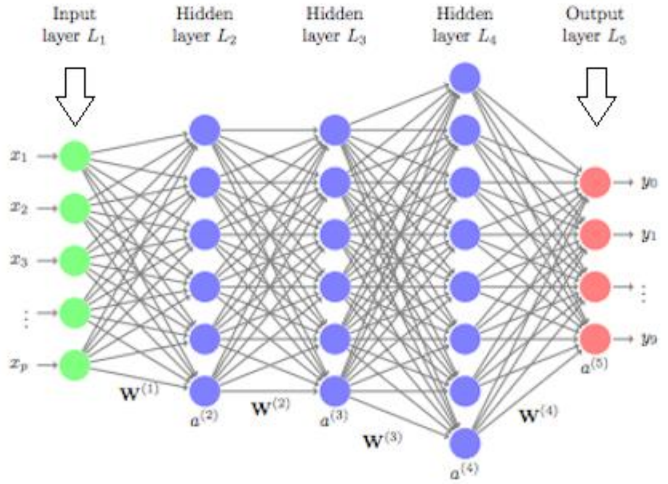

Figure 10: ***Illustration of a neural network****: The signal x is put in the first layer and transformed through the successive layers. By adjusting the weights W iteratively, the model is trained to detect automatically the features of the signal that allow a correct classification into one of the classes y in the output layer.*

Despite these advantages, there are two main downsides to deep learning in this context. First, neural networks are complex and difficult to tune, and the reported results of the best system typically do not include details about the strenuous selection process of the model [82]. Second, the very fact that features detected by neural networks are so complex makes these models very specific and less generalizable to similar but not identical tasks. For instance, Awais and colleagues highlighted the discrepancy between the contexts in which learning data is generated – controlled or semi-controlled laboratory structure – and real-life conditions [71]. Whereas in conventional models based on handcrafted features the human intelligence can ponder whether features with high discriminative power used in the model are generalizable, it is much less the case in the opaque models created in deep learning approaches. In chapter II, we propose a hybrid model combining conventional feature extraction and deep learning methods. The model unites the advantages of the two approaches and represents a performant alternative to pre-exiting models while remaining very simple and versatile.

Number and placement of the devices on the body

To detect human physical activities, accelerometers have been placed at different locations on the body, typically one or several of the following: ankle, thigh, hip, lower back, chest, upper arm and wrist [83]. One can expect that the classification accuracy depend on the number of devices employed and their location, and that a specific activity could be better recognized when the device is placed on specific part of the body [84]. In addition, an important factor to consider is the subjects' comfort, both physical and social, as it can affect their level of compliance with the wear time directives [85].

As the following two examples illustrate it, the optimal choice of placement and number of accelerometers to be used depends on the target behavior and the detection model chosen. Cleland and colleagues compared the accuracy of the best classifier into relatively basic categories (lying, sitting, standing, walking, running, stairs up and stairs down) for all possible combinations of number and locations (left ankle, left thigh, hip, lower back, chest and left wrist) [83]. For a single accelerometer, classification was best when the device was placed at the hip (97.81%) and worst at the ankle and wrist (95.63% and 95.88%). Interestingly, for two and more devices, the difference across combinations of locations was slight and insignificant. Moreover, from two devices up, adding a device did not improve accuracy significantly (and even slightly decreased it from three up [83]. Zdravevski and colleagues looked at a wider array of activities including, on top of the activities used by Clealand and colleagues, rope jumping, vacuuming, sweeping, dish washing, as well as bicycling at 50W and 100W resistance [81]. Accelerometers *and gyroscopes* were placed at the ankle, wrist, chest and the hip. Chest was the best (89.1%) at predicting activities when taking a single location. The best combination of two locations was ankle+wrist and chest+wrist (91.8%). For three locations, ankle+wrist+chest (93.4%) was best. With the four devices, the accuracy dropped to 93.0%. Moreover, the same study showed that different combinations predict best with different machine learning models.

In summary, the choice of the number and placement of accelerometers should account for the specific set of activities under consideration and practical aspects for the wearer [86]. In general, the marginal gain in accuracy decreases beyond two or three locations, and the impact of placement choice becomes less decisive as the number of devices increase.

Conclusion

This section has shown that accelerometers represented a widespread, very good trade-off between accuracy and feasibility of physical behavior monitoring in patients. In the choice of devices, technical specifications such as amplitude and sampling frequency can be relevant, but most modern commercial sensors for health research are good in that respect. More important is the choice concerning placement and number of devices. Detecting behaviors using raw acceleration signals should be encouraged over opaque proprietary metrics. While traditional outcomes such as motion intensity are straightforward, for complex behaviors, judicious choices are to be made regarding raw data processing techniques (windowing, feature extraction) and the statistical classification model. Chapter III takes up this challenge and proposes a simple and performant classification model of human behaviors based on raw inertial data.

The impact of built environment on physical activity

The previous sections have dealt with physical behaviors, how to measure them and what their health effects can be. Yet, what can be said about the determinants of physical behaviors? Many studies pointed to many determinants of physical activity. Multiple biological, environmental and behavioral factors were mentioned in the literature: season and weather [87], psychology [88], various life course events [89], age, sex, education, employment or health condition [52]. My thesis focuses specifically on the effects of characteristics of the built and social environments in cities. This section introduces key findings and some of the main challenges in investigating the effect of urban built and social environment on physical behavior, both at a theoretical and practical level.

Urban environment: an intervention space to fight physical inactivity epidemic

Changing the built environment, especially in cities, is one of the best-studied strategies to fight against the pandemic of physical inactivity [14, 90, 91]. Urban populations are usually much less active than their rural counterparts, and mass urbanization in developing countries plays a key role in the spread of the pandemic [92–94]. As populations move to cities, their physical environment changes radically, potentially affecting their opportunities to engage in physical activity. Aspects of the urban landscapes such as traffic, greenspaces or even aesthetics can play a role in the level of physical activity [93, 95]. Therefore, integrating programs aiming at promoting physically active lifestyles into the priorities of urban planning is a health public concern of the first order.

The literature assessing the effects of variables of the urban environment on physical activity levels is abundant. Below I will briefly give an overview of the environmental factors known to be associated with physical activity. The features mentioned here should not be regarded as exhaustive; whether a variable can be considered as a feature of the urban environment can sometimes be a matter of debate, and many attributes of the environment appear in different categories depending on the author. My aim here is only to illustrate the importance and the nature of the impact of some characteristics of the urban environment on physical activity of those exposed to them.

Studies often distinguish between *utilitarian* (or transportation) *physical activity*, and *recreational* (or leisure) *physical activity* (e.g. [93]). Environmental features can positively affect utilitarian physical activity levels through choice of *active means of transportation* (i.e. transports using only physical activity for locomotion, e.g. walking or bicycling instead of driving), as they affect the choice to engage in physical activity during leisure time. Special attention should be paid to public transportation. Although not active, it differs in several ways from passive means transportation such as car driving, as we will see.

Before introducing the main findings of the literature, it should be emphasized that the literature often distinguishes between children, adolescents, adults and older adults regarding the effects of certain environmental characteristics on activity [93]. In fact, different ages have different ways of life, which interacts with environment in different ways. My thesis focuses on adults and older adults, and evidence presented here refers to these age groups.

The main environmental attributes found to affect, or at least to correlate with, physical activity are the following:

- Walkability: Defined as a composite measure of residential density, street connectivity and a various land-use mix [96], this characteristic was found to be a strong correlate of adults' total levels of physical activity [97, 98] and in particular with older adults' walking [99, 100].
- Active transports infrastructure: More sidewalks, bicycle lanes and traffic free areas were reported to affect utilitarian walking [91, 97].
- Access to destinations: Areas comprising numerous destinations such as shops, services, cultural locations etc. were shown to be associated with walking in older adults [99], and overall walking in adults of all ages [95]. Expectedly, evidence was particularly strong for utilitarian walking [91, 101], but newer evidence points to a similar correlation with recreational walking in older adults [100].

- Aesthetics: Aesthetically pleasing streetscape (beautiful architecture, green- and waterways, absence of litter and vandalism…) was shown to be associated with both recreational and utilitarian walking [93, 97, 99, 100].
- Parks, open spaces and fitness facilities: Proximity to these locations was shown to encourage recreational and planned physical activity [91].
- Safety: Perceived safety from crime was shown to be strongly correlated with walking, especially for older adults; it is supposed that perceived crimes is a barrier for older people to get out of their homes [97].
- Proximity to public transportation: There exists some evidence to the effect of better access to public transportation on utilitarian [97, 102, 103] and leisure walking [100].

Evidence about the effect of environment on physical activity should meet, at least ideally, the criteria of objectivity and causality. By objectivity, I mean an objective assessment of one's built environment and physical activity. By causality, I mean the ability to determine the direction of the relationship between environment and physical activity, as we will see later. Yet, when examining the studies reviewed here (which are themselves reviews of dozens of studies), we see that only a little minority meet the criteria. In fact, designing a study that is based on objective evidence and allowing causal inference represents a major challenge. In the rest of this section, I will discuss the challenges associated with a robust study design and will introduce some of the recent advances in the field.

Study design

All studies investigating the relationship between environment and physical behavior rely on the assumption that individuals are exposed to environmental stimuli that might affect, or not, their behavior. In order to measure this exposure, one must locate individuals and assess the stimulus under consideration in this location. Along with the exposure, i.e. the explanatory variable, one must measure physical behavior, i.e. the outcome variable, and link between them both.

Localizing study subjects for measuring their exposure

Measuring environmental exposure starts by determining the location in which the study subjects are. Most past research used individuals' homes as a proxy to their localization, and most evidence to the effect of environment on physical activity was cumulated based on this proxy. As individuals are mobile over the day and spend most of their time outside of the home area, this method has been criticized for its inability to capture individuals' true exposure to environmental attributes [104, 105]. A practical and cheap solution consisted in asking individuals to report their activity spaces over the study period, using a web mapping tool [106]. For precise and reliable results, a few studies equipped individuals with Global Positioning System (GPS) in order to record their momentary localization and linking it to physical activity, thus investigating people's "spatial energetics" [107].

Environmental measures

Naturally, localizing subjects is not enough to assess exposure to environments that promote physical activity, and the researcher must compute the level of exposure in this localization by using geographic data regarding the environmental attribute of interest. Especially when localization is momentary, such spatial analysis can be complex. Thus, researchers typically use sophisticated Geographic Information System (GIS) tools allowing them to merge layers of information regarding individuals' positioning in space and environmental data [108].

An additional step in building a model consists in choosing a judicious method of aggregation of environmental variables around the location. First, one needs to determine the *area* around the localization that is to be considered in the model. Most studies define the area using administrative boundaries (such as census units), but these areas do not necessarily correspond to the real exposure experienced by the subject [97]. Other studies use buffers drawn around the spot of interest. This area is usually referred to as "buffer" in the literature [109]. This *size* of the area depends, of course, on the nature of the environmental attribute(s) studied. The *shape* of the area needs to be carefully chosen as well. As a matter of practicality, researchers tend to search for an environmental attribute within a certain radius of Euclidian distance around a location point. However, this method relies on a very strong assumption as to individuals' perception of their environment; especially in urban landscapes, a radius of Euclidian distance fails to reflect accurately the field of influence on individuals' behavior. For measuring exposure to shops or parks, for example, a Manhattan distance ("city block distance"), taking into account the layout of the grid street, is arguably more appropriate. Sometimes, a large highway can block the access to attributes that are spatially very close to the individual's location. To address this issue, Chaix and colleagues asked subjects to draw the boundaries on a web-based application [110]. Thus, they had an accurate representation of the subjective exposure area. This approach can be criticized for its lack of objectivity – the subjective activity space is perhaps influenced by the attributes themselves – and is applicable only when looking at a fixed location, typically individuals' homes, but remains appealing for directly addressing the thorny question of exposure spaces.

A second choice that the researcher needs to make, once the exposure area is defined, pertains to the function of aggregation of the attributes of interest in the area. To cite a few examples, researchers can choose an indicator function (we label an area as 1 or 0 depending on whether the number or quantity of attribute exceeds a certain threshold), an average function, a sum function etc. As for the question of the buffer parameters, one should remember that the choices of the aggregation function depends, beyond issues of practicality, on many factors, such as the attribute under consideration, the location's topography, whether a fixed location or a trajectory is studied, etc. (see Figure 7).

Thus, beyond issues concerning data sources of GIS, the parameters chosen for spatial analysis - buffer size and shape, aggregation function etc. - is a delicate issue that researchers need to discuss and justify [109, 111]. As objective assessment of exposure to environmental attributes develops, more awareness of the consequences of the parameter choice in exposure measurement will prevent the great discrepancies in evidence that we have witnessed in the field so far.

Figure 11: ***Map of green areas****: When assessing exposure to green areas around an individual's home, choices concerning the exposure area yield different values. If we measure the fraction of the green area in the exposure zone, different radiuses for the buffer (based here on Euclidian distance) will yield very different values. Likewise, a measurement based on administrative boundaries (here the limits of the neighborhood) will yield yet another value. If we are only interested in the presence/absence of green areas, the smaller buffer zone does not contain any, while the larger ones and the neighborhood do.*

Physical activity measures

The previous section has discussed challenges in measuring exposure to environmental factors. In order to link it to physical activity, the researcher must be able to possess reliable and accurate data on the physical activity performed and to be able to attribute variation in the observed activity to the environment. Direct observation of physical activity in free-life conditions is practically impossible. Consequently, some studies have equipped their subjects with activity monitors (usually accelerometers) to measure their physical activity levels in time in an objective manner [95]. Yet, despite a growing number of studies using objective measures of physical activity, the body of evidence about the effects of environment on physical activity is largely made up of studies relying on individuals' self-reports [97]. As explained above, self-reports regarding physical activity suffer from a considerable bias which makes it difficult to draw meaningful epidemiological conclusions [54].

The problem is even more acute when investigating the effects of environment on postures or accumulation patterns of physical behaviors, as it is reasonable to assume that the recall bias is even more important when it comes to reporting these variables. As explained above, physical behaviors are complex and multidimensional: beyond levels of MVPA, postures such as sedentary behaviors are also relevant for health. Likewise, segmentation was shown to influence health that is independent from the total volume. In fact, to the best of our knowledge, no study investigated the relationship between environmental attributes and body postures or segmentation patterns so far.

Assessing the effect of environment on physical activity requires identifying consistent patterns between measures of individuals' environment and their physical activity. To do so, one needs to define a common statistical time unit of both variables. The immense majority of research focus on exposure to environmental variables around a single fixed location, such as the home or workplace, and aggregate physical activity over long periods, such as days or the entire monitoring periods (e.g. [93, 97, 99–101]). In contrast, a few studies looked at individuals' physical activity over very short epochs (one second to one minute) and linked them to the immediate environment to which the individuals were exposed in order to obtain high-resolution evidence to the relationship between the two [112, 113].

This tempting design, referred to as the "contemporaneous design" in this thesis, was applied a few times, such as in a study on the influence of exposure to greenspace on MVPA in children [113], which used 10-second periods as statistical units. This approach was later criticized for its inability to address causality, as I will explain below. Another approach, developed by Chaix and colleagues, exploited the same high resolution to chop up these observations into meaningful spatio-temporal *life-segments* [114]. These life-segments could be, for instance, activity places (shopping malls, home etc.) or journeys (going from home to work). Looking at life-segments allows the researcher to answer very specific questions about the relationship between environment and physical activity *in a specific context*, and to some extent, to circumvent causality issues, as explained below.

The problem of causality in cross-sectional studies

Epidemiological studies ideally identify causal relationships between risk factors and health. In the case of environmental epidemiology of physical activity, cross-sectional relationships observed between environmental attributes and physical activity, two causal self-selection mechanisms can compromise inference.

A first mechanism, well documented in the literature, is known as the *neighborhood self-selection bias*. When a cross-sectional relationship is found between an environmental attribute around an individual's residence and her/his physical activity level, one can argue that not the attribute caused the high activity level, but the individual chose to establish residency in this area precisely because she/he was prone to perform a high level of activity [95]. The most common solution to this bias is to design a longitudinal study, in which individuals' physical activity is analyzed as a function of the time-changing environments to which they are exposed. Another similar solution is quasi-experiments, in which the subjects' physical activity is measured shortly before and after a sudden significant change in an environmental attribute. For instance, Brown and colleagues measured the effect of a light-rail in Salt Lake City, Utah, by measuring residents' physical activity shortly before and after the construction [102]. However, designing a longitudinal study or finding a quasi-experiment is difficult, and such studies are very rare. As an example, in a recent review about environment and physical activity in older adults, 94 articles were cross-sectional, five longitudinal and one quasi-experimental [97].

In many articles, neighborhood self-selection is addressed by controlling for the reported motivation in choosing the neighborhood in the first place [95]. In addition, contemporaneous designs linking between momentary exposure and physical activity represent a promising solution to this problem: one can limit the frame of analysis to places in which self-selection is not likely to happen, such as to life-segments that are remote from the subjects' residential neighborhood.

A second self-selection mechanism, less discussed in the literature, is the *momentary* (or *daily*) *self-selection bias*. This bias occurs in contemporaneous designs, when a correlation between exposure and activity is observed. One can argue that the change in activity observed is not due to the exposure to the environmental attribute, but rather to the person's self-selection of a specific place based on their personal preference for physical activity [115]. Here again, the concept of life-segments presented above proves useful, as it enables an analysis on a selection of places in which such a selection is not likely to occur (for example in places to which one does not choose to go, such as workplace). This technique is used and developed in chapter IV.

Conclusion

Environmental interventions, especially in urban contexts, can represent a good way to promote an active lifestyle. To do so, epidemiological studies are needed to investigate the effects of variations in environmental attributes on residents' physical behavior using objective measurements. However, there are several designing challenges: extracting GIS features and aggregating them in space, aggregating data in time and circumventing causality issues inherent to cross-sectional designs. Especially informative are studies linking exposure to environmental features by location tracking and their contemporaneous physical activity. In this framework, slicing meaningful life-segments in time and space is a promising idea, since it can reveal new effects of specific environments to physical activity and isolate contexts in which the causal direction is unambiguous. Using objective measurements of physical behaviors (including postures) and location within the paradigm of life-segment analysis, chapter IV estimates the causal effects of various attributes of the urban environment on physical behaviors.

Thesis objectives

In this introduction, we have seen how the theoretical frameworks of physical activity epidemiology have evolved over the past decades. From a framework in which planned continuous bouts of moderate/vigorous physical activity were viewed as the only relevant factors, more recent studies hint to the potential relevance of *any kind of activity* of *any duration*, all along the day. From these new insights, it follows a scientific need for a conceptual and methodological toolbox capable of dealing with nuanced categories of activities and complex patterns of activity accumulation over time. Likewise, for empirical studies, we need new material monitoring devices capable of accurately recording large arrays of activity in real time. As we shift towards this new paradigm, upstream research investigating the environmental causes of physical activity must account for *continuous exposure* to different environments and study their effect on different types of activity, all along the day.

The present thesis positions itself in this new paradigm by undertaking a comprehensive study of physical activity in free-life conditions, its causes and effects, accounting for a large spectrum of activities, in all places and at all times of the day. To do so, I carried out an extensive analysis of data from the Paris RECORD epidemiological cohort, comprising medical information as well as location and physical activity records in more than 150 individuals. To accomplish this task, I contributed to the development of new conceptual and methodological tools fitting into this framework. More concretely, the location tracks of individuals of the RECORD cohort was combined with geographical information in order to assess continuously the exposure levels to characteristics of the urban environment. Exposure to the environment was then linked to a detailed nomenclature of their contemporaneous physical activities. These were in turn linked to medical records, thus aiming at uncovering the entire pathway leading from environment to health through physical activity.

This project was organized around three main research questions: I. How are physical behaviors related to health? II. How can we improve measurement of physical behaviors? III. How does the urban built environment affect residents' physical behaviors? Answering these questions yielded the five chapters of this thesis, as illustrated in Figure 12.

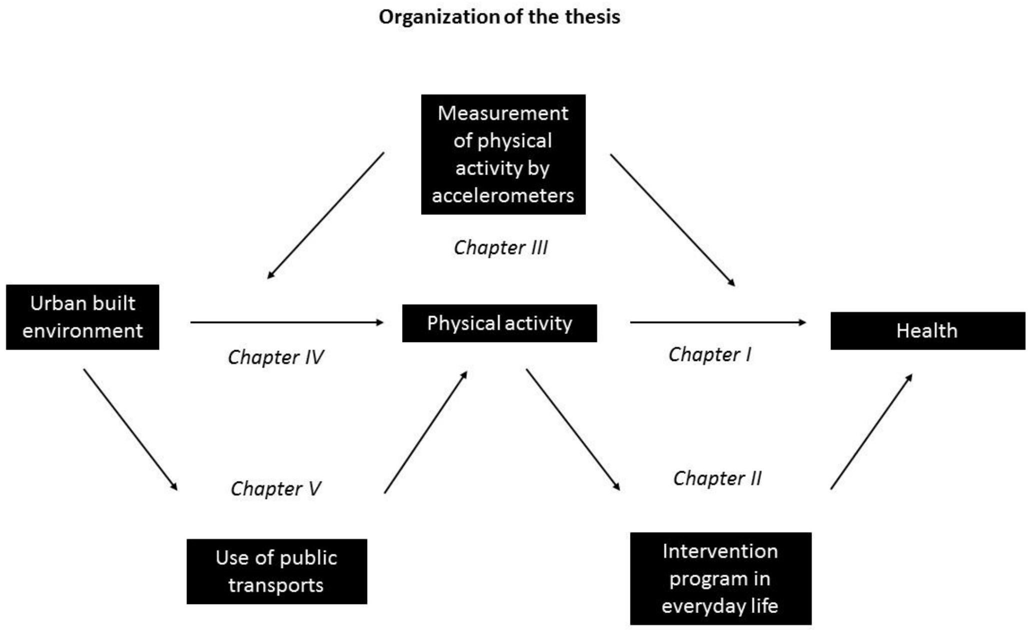

Figure 12: ***Organization of the chapters in the thesis.***

Following a logical flow, this thesis could have been organized by presenting my work in investigating the environmental effects on physical activity first, and then on the effects of physical activity on health. For two reasons, this thesis is organized the opposite way, starting with the effects of physical activity on health and ending with the effects of environment on physical activity. The first reason is didactic: the innovation of the section on environment and physical activity, as well as some concepts used in it, will be better understood by introducing the challenges of research on physical activity and health in the first place. The second reason relates to the history and the natural development of the discipline: epidemiological research started by focusing on the health hazards caused by high levels of physical inactivity in the population; it is only when this problem had been established that researchers, motivated by a need for intervention, began to study its causes in the environment in an extensive manner.

The detailed layout of this thesis is as follows:

Chapter I investigates the relationships between various aspects of physical behaviors on health variables using observational data in free-living conditions from the RECORD epidemiological study. It innovates in the following domains. (i) It distinguishes between lying and sitting, two behaviors that are studied together as “sedentary behaviors” in past literature, and standing, an in-between category whose relationship with health in free-living conditions remains uncertain. (ii) It disentangles the effects of volume and segmentation by integrating advanced metrics of both. (iii) It considers different volumes – components of the behavioral time budget – as compositions, thus accounting for the interdependency between components of the budget of physical behaviors.

Chapter II presents the results of an interventional study conducted in Denver, Colorado, USA, in which participants were asked to integrate episodes of physical activity into their daily routines, following different segmentation patterns. This study brings new insights into the following issues. (i) It assesses the ability to implement an interventional program based on different schemes of sedentary behavior segmentation. (iii) It investigates the effects of different segmentation patterns on subjects’ well-being, fatigue, vigor and index of metabolic health (glucose and insulin levels).

Chapter III is a methodological intermission about derivation of physical behaviors from accelerometer signals. It follows from the need to assess nuanced categories of behavior in observational studies, as stated in the first part. It proposes a new algorithm aiming to detect a very detailed nomenclature of physical behaviors from accelerometer data. The algorithm combines several state-of-the-art techniques and outperforms algorithms proposed by previous studies in accuracy and computational efficiency.

Chapter IV examines effects of various attributes of the urban environment on mobility and physical behaviors. It proposes a pioneering framework with the following innovations. (i) It combines a high-resolution, contemporaneous design linking objective measurements of environmental attributes and physical activity. (ii) It investigates environmental effects on unseen categories of behaviors: sedentary behavior, standing and MVPA. (iii) It circumvents issues of causality bias by adopting a life-segment approach.

Chapter V builds on results from chapter IV suggesting that an efficient infrastructure of public transports is a key variable in the link between urban environment and physical activity. With the RECORD study, it investigates the passengers' time budget of physical behaviors by transportation mode (public transports, car…) using the same design as in the previous chapter. It is one of the very few studies to address this question, and it represents a precious source of information for estimating the effects that urban planning policies concerning transportation can have on residents' physical behaviors.

CHAPTER I

Associations between physical behavior and various health outcomes

Reference:
Debache I, Bergouignan A, Chaix B, Sneekes EM, Thomas F , Sueur C. Associations of Sensor-Derived Physical Behavior with Metabolic Health: A Compositional Analysis in the Record Multisensor Study. International journal of environmental research and public health 16 (5), 741. 2019.

Vocabulaire technique utilisé

Activité physique d'intensité modéré à vigoureuse : Activité physique nécessitant une dépense énergétique de 3 MET (3 fois le métabolisme basal) ou plus, par exemple : marcher à 5 km/h ou plus, passer l'aspirateur.

Comportement sédentaire : Activité physique quasi nulle, nécessitant une dépense énergétique de 1,5 MET ou moins tout en ayant une posture assise ou couchée. Dans la littérature traditionnelle, il est cependant usuel de considérer les activités effectuées debout au repos comme des activités sédentaires.

Comportement physique : Manière du corps de se mouvoir (activité) et de se positionner dans l'espace (posture), par exemple : marcher, sauter, être assis, être allongé.

Volume d'activité ou *volume de comportement* : Le temps total passé à effectuer un comportement physique sur la période étudiée (typiquement plusieurs jours).

Budget-temps des comportements physiques : L'ensemble des volumes des comportements, dont la somme est égale au temps total de la période étudiée.

Chapitre I : Associations entre l'activité physique et des indices de santé (résumé français)

De nombreuses études ont démontré les relations existant entre l'activité physique et diverses variables de santé. Il a été démontré qu'une activité physique insuffisante entraînait un risque accru de maladies cardiovasculaires, diabète de type II, ainsi que les cancers du côlon et du sein. Des études plus récentes ont avancé que, pour un volume total d'activité physique d'intensité modérée-à-vigoureuse constant, un large volume de temps sédentaire représenterait un risque supplémentaire pour la santé cardio-vasculaire. Enfin, ces dernières années ont vu émerger une littérature importante sur l'effet de la segmentation temporelle des comportements physiques sur la santé. D'une part, le postulat selon lequel le volume d'activité physique modérée-à-vigoureuse recommandée devait être accumulée en périodes plus ou moins longues (par exemple 10 ou 20 minutes) a été récemment remis en question, certains auteurs arguant que ce volume recommandé pouvait être effectué aussi bien en séquences longues que courtes. D'autre part, ce sont précisément les périodes *prolongées* de comportements sédentaires qui ont été mises en cause dans les études, indépendamment du budget-temps des comportements physiques. En supposant que le volume d'activité modérée-à-vigoureuse soit suffisant, et que le temps sédentaire soit suffisamment segmenté, il n'est en effet pas certain que le volume de temps sédentaire ait des effets nocifs sur la santé.

La plupart des études réalisées en milieu de vie libre sur les activités sédentaires et leur lien avec la santé souffrent de plusieurs limites. Premièrement, la distinction entre divers comportements à faible dépense énergétique, à savoir les activités réalisées debout, assis ou couché, n'a pas été suffisamment étudiée jusqu'à présent. Deuxièmement, les effets des comportements sont le plus souvent étudiés séparément les uns des autres, alors que c'est l'ensemble des comportements qui doit être considéré comme un budget-temps (c'est-à-dire une composition), où le volume d'un comportement est nécessairement alloué sur le budget aux dépens d'un autre. Troisièmement, la segmentation temporelle des comportements a été étudiée de façon peu systématique. En mesurant une relation entre un comportement et la santé, la plupart des études ne séparent pas clairement la part due à son volume temporel total et à la façon dont celui-ci est segmenté. Par ailleurs, en agrégeant les comportements physiques sur des pas de temps relativement long (par exemple une minute), les grandes études ont ignoré les micro-séquences de comportement, bien que celles-ci puissent avoir un effet important sur la santé.

Pour remédier à ces lacunes, nous avons entrepris d'analyser, de façon transversale, des données de l'étude de cohorte RECORD effectuée en région parisienne. Dans le cadre de cette étude, 154 adultes en bonne santé (64% d'hommes, âgés de 34 à 83 ans) ont porté des accéléromètres sur le torse et la cuisse pendant une semaine, dans des conditions de vie libre. Grâce à l'accélération enregistrée tout le long de la journée, nous avons pu dériver les comportements physiques des sujets à une résolution d'une seconde et établir un profil d'activité détaillé et représentatif de leur vie quotidienne. Cette résolution nous a permis de dresser un portrait non seulement du budget-temps des comportements physiques (temps passé couché, assis, debout, effectuant des activités d'intensité légère ou modérée-à-vigoureuse) mais aussi de la façon dont ces volumes de temps ont été cumulés sur la journée. Dans la mesure où ces profils d'activité reflétaient les habitudes comportementales des sujets, il nous a semblé intéressant de les mettre en relation avec diverses mesures physiologiques collectées au début de l'étude. Les deux dimensions du comportement, le volume et la segmentation, ont été donc intégrées dans nos modèles d'analyse, qui ont également respecté le caractère compositionnel des budget-temps (c'est-à-dire de la codépendance des composantes du budget).

Au-delà des relations bien connues entre l'activité modéré-à-vigoureuse et la santé, notre analyse a pu mettre en évidence des relations moins connues entre les budgets-temps à haute composante « debout au repos » et un meilleur profil lipidique : concentration sanguine en triglycérides plus faible, concentration plus haute de lipoprotéines de haute densité (HDL). Ainsi, nos modèles d'analyse suggèrent que les profils où une grande partie du temps « assis » est remplacé par du temps passé « debout au repos » bénéficiaient d'un bilan lipidique tout aussi bon que les profils majoritairement sédentaires à niveau d'activité physique modérée-à-vigoureuse satisfaisant. De plus, nos modèles ont montré une corrélation importante entre les mesures d'adiposité (indice de masse corporelle et tour de taille) et les individus qui remplaçaient le temps passé généralement assis par du temps passé couché.

Les relations entre la segmentation des différents comportements et les mesures de santé sont ressorties moins clairement. Nous avons constaté qu'à budget-temps constant, un volume sédentaire fragmenté en nombreuses micro-séquences était associé à une glycémie plus faible. D'autre part, une accumulation du temps non-sédentaire (debout ou actif) par de nombreuses micro-séquences était associée à un indice de masse corporelle plus élevé. Enfin, aucune corrélation entre les modes de segmentation de l'activité modérée-à-vigoureuse et nos indices de santé n'a pu être observée.

Notre étude souffre de limites évidentes. D'abord, sa conception transversale nous interdit toute inférence causale. De plus, la puissance statistique relativement faible de l'étude ne nous permet pas de déterminer de façon catégorique que l'absence de corrélation dans notre échantillon reflète une réalité physiologique et n'est pas due à la petite taille de nos effectifs. Enfin, certaines relations observées entre la segmentation des comportements sédentaires et des indices de santé semblent complexes et nécessiteraient une investigation plus approfondie pour être tout à fait élucidées.

Malgré ces limites, cette étude reste innovatrice sur plusieurs plans. A partir d'observations en conditions de vie réelle, elle démontre que certaines allocations du budget-temps postural pouvaient partiellement compenser un manque d'activité physique d'intensité modérée-à-vigoureuse. De plus, elle met en avant l'importance cardinale d'une prise en compte des micro-séquences dans l'étude des relations entre segmentation du comportement et la santé. Nous suggérons que de nouvelles études continuent à développer les approches analytiques proposées en les appliquant à des jeux de données plus grands, de préférence dans le cadre d'un suivi longitudinal.

Article

Associations of Sensor-Derived Physical Behavior with Metabolic Health: A Compositional Analysis in the Record Multisensor Study

Isaac Debache [1,*], Audrey Bergouignan [1,2], Basile Chaix [3], Emiel M Sneekes [4], Frédérique Thomas [5] and Cédric Sueur [1,2]

1 Institut Pluridisciplinaire Hubert Curien (IPHC), UMR 7178 Centre National de la Recherche Scientifique (CNRS), Université de Strasbourg, 67000 Strasbourg, France; audrey.bergouignan@iphc.cnrs.fr; (A.B.); cedric.sueur@iphc.cnrs.fr (C.S.)
2 Division of Endocrinology, Metabolism, and Diabetes and Anschutz Health and Wellness Center, School of Medicine, University of Colorado, Aurora, CO 80045, USA
3 INSERM, Sorbonne Université, Institut Pierre Louis d'Epidémiologie et de Santé Publique, IPLESP, Nemesis team, F75012 Paris, France; basile.chaix@iplesp.upmc.fr
4 Department of Rehabilitation Medicine, Erasmus MC, 3000 CA Rotterdam, The Netherlands; e.sneekes@erasmusmc.nl
5 Preventive and Clinical Investigation Center, 75116 Paris, France; Thomas@ipc.asso.fr
* Correspondence: isaac.debache@iphc.cnrs.fr; Tel.: +33 3 88 10 6931

Abstract

Previous studies about the effects of physical activity and sedentary behaviors on health rarely recorded the exact body postures and movements, although they might be of metabolic relevance. Moreover, few studies treated the time budget of behaviors as compositions and little was done to characterize the distribution of durations of behavior sequences in relation with health. Data from the RECORD (Residential Environment and CORonary heart Disease) study of two combined VitaMove accelerometers worn at the trunk and upper leg for a week by 154 male and female adults (age = 50.6 ± 9.6 years, BMI = 25.8 ± 3.9 kg/m²) were analyzed. Using both iso-temporal substitution and compositional analysis, we examined associations between five physical behaviors (lying, sitting, standing, low physical activity, moderate-to-vigorous activity) and seven health outcomes (fasting serum glucose, low- and high-density lipoprotein, and triglycerides levels, body mass index, and waist circumference). After adjustment for confounding variables, total standing time was positively associated with better lipid profile, and lying during the day with adiposity. No significant association was observed between breaking up moderate-to-vigorous physical activity and health. This study highlights the importance of refined categories of postures in research on physical activity and health, as well as the necessity for new tools to characterize the distribution of behavior sequence durations, considering both bouts and micro-sequences.

Keywords: sitting; standing; low physical activity; moderate-to-vigorous physical activity; blood lipids; glucose; HDL; compositional analysis; iso-temporal substitution

Introduction

Physical inactivity has been recognized as a major health hazard for several decades [1–3]. More recently, research highlighted prolonged sedentary behavior (SB) as a risk factor for developing coronary heart diseases, obesity, diabetes [4–7], and some cancers [8,9]. This risk factor is thought to operate independently from the level of physical activity (PA) and through different metabolic mechanisms [10,11]. Strictly defined, SB refer to sitting or reclining postures with low energy expenditure (<1.5 metabolic equivalent) [5,12,13]. However, most objective evidence to their adverse effects on health were obtained using a looser definition of SB, based on the sole movement intensity and without distinguishing between quiet standing, sitting and lying. As a consequence, the extent to which the risks associated with SB are distinct from physical inactivity in a narrow sense is still being debated [13,14]. Even among the studies that explicitly distinguished between standing and sitting time, for example with regard to glucose or lipid profile [15,16], only a few investigated the associations between postural behavior and metabolic outcomes in natural, free-living conditions [17,18]. Moreover, the distinction between lying and sitting cannot be properly addressed, even with newer thigh-worn devices such as ActivPal®. The study presented here uses a double accelerometer, worn on the subjects' trunk and thigh, which allows for a precise derivation of body postures and movements.

Independent of the total time spent in MVPA (Moderate to Vigorous Physical Activity) and SB, shorter SB bouts are thought to have a positive effect on cardio-metabolic biomarkers [17,19–21]. Yet, although the patterns by which a given SB time is partitioned into sequences of varying durations is relevant for health, most past studies used very simple indices to characterize partitioning patterns, such as the median or the mean bout duration. In the present study, we tested more sophisticated partitioning indices: beside the median bout duration, we used the Gini index of the sequence length distribution [19], and the ratio of behavior time spent in bouts to the total behavior time (spent in bouts or not). By a behavior bout, we mean a sequence of significant duration, e.g., 1 minute, during which the dominant observed activity is the activity of interest. Unlike most past studies, we examined the relationship between partitioning patterns and health not only for SB, but also for other behaviors, such as standing or MVPA. In addition, we did not focus exclusively on bouts, but also on sporadic behavior sequences and their distribution.

The associations between the total time volumes spent in each behavior (sitting, standing, MVPA etc.) and a certain health outcome are often estimated in separate models. Whether expressed as time or as proportions, the different behaviors add up to a constant (the total time studied or to 1), and are therefore to be regarded as a time budget with relative and codependent parts: a growing volume of one part always comes at the expense of another. Ignoring this sum constraint often leads to erroneous estimates and interpretations [22–24]. Some studies have acknowledged this issue and used iso-temporal substitution techniques [25, 26]. These models estimate the effect of time reallocation from one part to another on the outcome variable, while all other parts remain constant. However, others argue that these models still fail at treating these data as proportions, which constitute a sample space, known as the Aitchison simplex, with its own mathematical properties and methods of analysis [24,27,28]. As Biddle and colleagues did in a recent article [18], we address this issue by using both iso-temporal models and methods of compositional analysis.

By using precise categories of physical activities and postures, a thorough approach to duration distribution of activity sequences and appropriate techniques for analyzing time compositions, this study proposes a novel, comprehensive framework for examining the associations between time spent in different physical behaviors, their daily patterns and key health outcomes.

Material and Methods

Study Subjects

The present study uses data from the MultiSensor sub-study [29] of the RECORD (Residential Environment and CORonary heart Disease) Cohort study. From February 2007 to March 2008, individuals that came to four of the IPC (Investigation Préventive et Clinique) Medical Centers for a free medical examination offered by the French National Insurance System for Employees were invited to enter the RECORD study. Eligibility criteria were age 30–79 years, residence at baseline in 10 districts of Paris (out of 20) and 111 other municipalities of the Ile-de-France region and sufficient cognitive and linguistic abilities. During the second wave of the RECORD Study (September 2013 to June 2015), after completing their medical checkups, participants were systematically invited to enter the RECORD MultiSensor Study whenever monitoring devices were available. In this study, 154 participants (97 men and 57 women), aged 34–83 years, accepted to carry a GPS receiver and the two combined accelerometers placed at the trunk and on the lower limb. The study protocol was approved by the French Data Protection Authority (Decision No. DR-2013-568 on 2/12/2013). All participants signed a written informed consent form.

Anthropological and Biological Data

During the screening visit, participants underwent a medical examination including anthropological measurements and a blood draw in a fasting state. Details about collection of anthropological and biological data can be found elsewhere [30]. Anthropological measurements were made by trained nurses at the medical centers. WC (Waist Circumference) was measured using an inelastic tape placed midway between the lower ribs and the iliac crest, on the mid-axillary line. In this study, we used the following outcomes: serum glucose concentration, plasma triglyceride, high- and low-density lipoprotein (HDL and LDL), as well as body mass index (BMI) and waist circumference (WC).

Physical Activities and Postures

Physical activities and postures were derived from two tri-axial Vitamove Research-V1000® (Vitabase v2.0 B5, Temec Instruments, Herleen, The Netherlands) accelerometers. Participants were requested to wear one at the trunk and the other on the right upper leg during wake time for seven days, as they carried out usual daily activities in free-living conditions (except for water-based activities). In addition, they wore a small GPS device and kept a log of the places they visited and the wearing times of the devices. We invalidated days with less than ten hours of wear time, and subjects with less than four valid days out of seven [31]. Twenty-two subjects out of 154 did not meet these criteria.

The base software of the sensors, VitaScore®, has a large nomenclature of physical activities and postures. To simplify the analysis, we combined them into five categories: lying (trunk in horizontal or nearly horizontal position), sitting (trunk upright or nearly upright), standing, light physical activity (LPA, including slow movements) and moderate-to-vigorous physical activity (MVPA, including walking, running, biking etc.).

To identify bouts, we used a modified version of the function *guideline.bouts* from the R-package *activpalProcessing* [32]. The minimum length of the bouts was set to 60 seconds and the threshold for the proportion of the behavior of interest was kept at 0.8 (see Figure 1). For the sake of simplicity, we did not analyze partitioning patterns for each behavior, but on three broader categories: SB (including lying and sitting), non-sedentary behaviors (NSB) including standing, LPA and MVPA, and MVPA only.

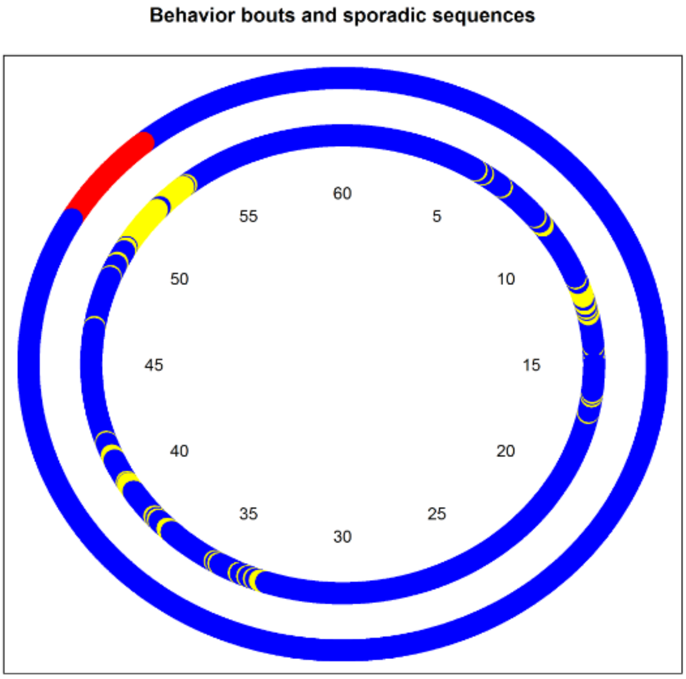

Figure 1. ***A random sedentary-behavior chart over an hour*** *(based on the empirical distribution of sequence duration). In the inner circle, non-sedentary time is colored in yellow and sedentary time in blue. The outer circle represents, for the same data, the time that is regarded as non-sedentary bout is in red, and as sedentary bouts in blue. This study takes into account both bouts and sporadic sequences, although the latter is disregarded by traditional methodology. Here, 7.6 minutes were spent in NSB (yellow), and 3.4 in a NSB bout (red). The ratio time in bouts to total time is of 0.42.*

Other Data

Individuals that joined the study answered a questionnaire regarding socio-demographics, dietary and health habits, from which we used data about age, educational level and annual total income. As nutrition is thought to be correlated with both sedentary time and health [33], we also added to the models information regarding nutritional and health habits. These variables are described in details in Table C1 in Appendix C. A discussion about the validity of the questionnaire is to be found in another paper devoted to the RECORD study [34].

Data Processing

BMI and triglycerides data were log transformed in the models. The information about nutrition and health habits was reduced and expressed as the first two dimensions of a principal component analysis performed over the array of all relevant variables in the questionnaire mentioned above. They appear in the tables as "nutritional index".

Statistical Analysis

The following three models, run with individuals as statistical units, addressed the two questions at hand: the relationship between the behavior time budget and health (the first two models below), and the relationship between the behavior partitioning patterns and health for a given behavior time budget (third model below). In all models, sex, age, annual income, education, and the two nutritional indices were added as control variables. BMI was added as control variable in all the models, except in those whose response variable was BMI or waist circumference.

- Iso-temporal substitution models: they estimate the change in the health outcome variable associated with time reallocation (in proportion) from a behavior to another, while all other behavior time volumes remain constant. Thus, the models preserve the compositional structure of the data.
- Compositional models: they are identical to the familiar linear regression models, but before including them as regressors in the models, the compositions are

transformed from coordinates in the Aitchison simplex for composition S^D to the coordinates in the real space R^{D-1} (here, we chose the isometric log ratio (ilr) transformation [35]). Once the coefficients for the compositions are estimated by the models, they are back-transformed to the Aitchison simplex. The independent variable (here, the health variable) is fitted in the same way as in a traditional linear model, but using the Aitchison geometry for compositions [22] (i.e., by taking the Aitchison inner product of the compositional vector and the corresponding coefficient vector, see Appendix A). Thus, we can estimate our health response variable for any composition, or the change in the response variable following any change in a composition, while operating in the appropriate mathematical framework for these data. To illustrate the change in a health outcome associated with a change in a time budget, we created four hypothetical profiles, which represent archetypes of physical activity patterns, and compared the predicted health outcomes for these profiles against the average profile. The four profiles are 'couch potato'—a time budget with a strong component of lying/reclining postures (lie = 30%, sit = 50%, stand = 10%, LPA = 5%, MVPA = 5%); 'office worker'—strong component of sitting (lie = 5%, sit = 70%, stand = 10%, LPA = 5%, MVPA = 10%); 'doorman'—strong component of standing (lie = 5%, sit = 15%, stand = 70%, LPA = 5%, MVPA = 5%); active—strong component of MVPA (lie = 5%, sit = 40%, stand = 30%, LPA = 5%, MVPA = 20%). We implemented the models using the R-package *compositions* [36] and the handbook by van den Boogaart and Tolosana-Delgado [23].

- Linear models for behavior partitioning: these are traditional linear models, which estimate the change in the health outcome associated with the change in a partitioning index. We did not calculate the indices for each behavior, but rather for three broader categories of behaviors (SB, non-SB, and MVPA). To make sure that the association of behavior partitioning with health is independent of the behavior time volumes, we added the behavior time budget (expressed as ilr) to the model as control variable. As partitioning indices, we use the median length of the behaviors bouts, the ratio of the behavior time in bouts to the total behavior time (spent in bouts or not, see Figure 1), and the Gini index of the total time distribution of sequences of different durations (see Figure 2).

All analyses were performed using the R statistical system (version 3.3.2) [37]. Statistical significance was set at 0.05.

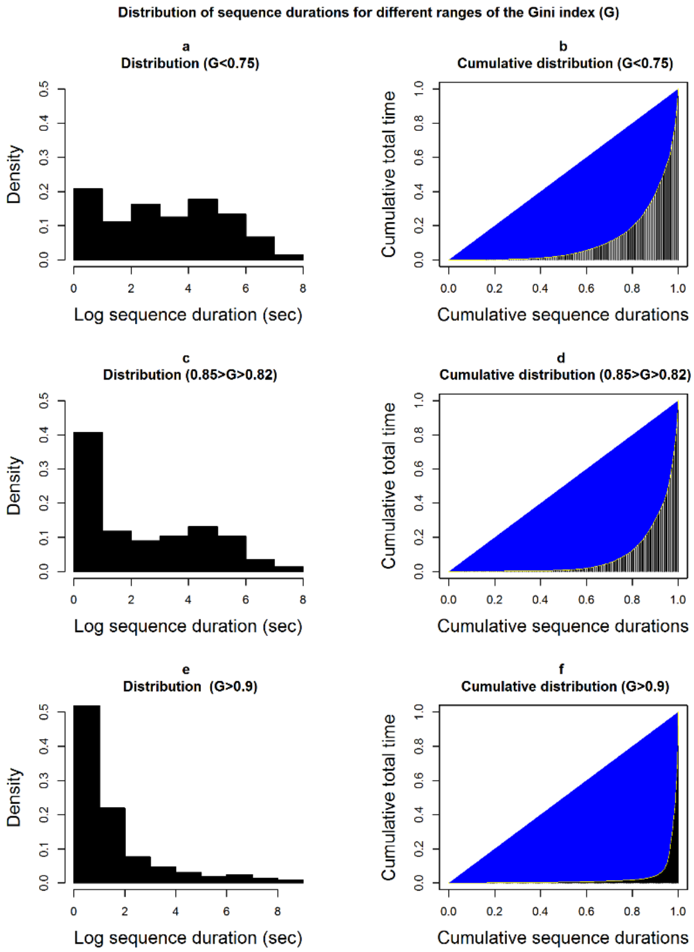

Figure 2. ***Random non-sedentary sequence durations*** *(sub-figures* ***(a)****,* ***(c)****,* ***(e)****) and their corresponding Lorenz curves, i.e. the total time accumulation by sequence duration (sub-figures* ***(b)****,* ***(d)****,* ***(f)****), for three different ranges of the Gini index: low (G<0.75; sub-figures* ***(a)*** *and* ***(b)****), medium (0.85>G>0.82; sub-figures* ***(c)*** *and* ***(d)****) and high (G>0.9; sub figures* ***(e)*** *and* ***(f)****). The durations are randomly drawn from the empirical distributions observed in our population and they add up to the same total time. The Gini index increases as contributions of sequences to the total time are less even (top to bottom). It represents the area between the diagonal and the Lorenz curve (right-hand column) divided by the whole area under the diagonal.*

Results

Anthropological, Demographics and Biological Characteristics of the Participants

The final population was made up of 131 subjects, 64% of them men, aged 50.5 ± 9.6 (arithmetic mean ± standard deviation) years. We removed twenty-two participants for insufficient wear time and one for incomplete biological data. Participants were, in average, slightly overweight with a BMI of 25.8 ± 3.9 kg/m². Seventy-six of the participants were overweight (BMI > 25 kg/m²) but only 16 were obese (BMI > 30 kg/m²). The others were in normal ranges (20 kg/m² < BMI < 25kg/m²). Three participants had metabolic syndrome, as defined by the International Diabetes Foundation [38]. In average, the other health variables examined were in normal ranges. The socio-economic status of the participants was, however, somewhat higher than the French average [39].

Daily Pattern of Physical Activity and Sedentary Behaviors

The mean daily wear time was 14.34 ± 2.08 hours. On average, our population spent 8.04% ± 3.30% of their wake time in MVPA, 3.58% ± 1.40% in LPA, 27.13% ± 9.61% in quiet standing, 51.57% ± 12.08% sitting and 9.68% ± 9.60% lying. Lying time was subject to high inter-individual variability, with values ranging from 0% to 53.63%. The closed geometric mean (which is usually preferred over the arithmetic mean for compositions [23]) was (lie = 5.64%, sit = 54.72%, stand = 27.94%, LPA = 3.64%, MVPA = 8.01%). The covariance matrix, accounting for co-dependencies between the parts of the composition, is shown in Table B1 in Appendix B.

With regard to partitioning patterns, the median bout duration was of 1.8 ± 0.66, 4.37 ± 1.59 and 6.58 ± 2.72 minutes, for MVPA, NSB and SB, respectively. Although not necessarily related to the median length, the Gini index also points to different partitioning patterns for SB and NSB time than for MVPA, the former being accumulated through fewer, longer sequences (0.8 ± 0.06, 0.83 ± 0.05, and 0.6 ± 0.09, respectively). While the largest share of SB and NSB time was spent in bouts longer than 1 minute (0.98 ± 0.02 and 0.96 ± 0.03), the share was much smaller and more variable for MVPA (0.53 ± 0.15).

Detailed descriptive statistics of the physical behaviors and the related indices used in this study are shown in Table 1 and for health and potentially confounding variables in Table 2.

Table 1. *The top section of the table shows the arithmetic mean, standard deviation (SD), minimum and maximum for the time proportion devoted to each physical behavior (n = 131). The bottom sections show the same statistics for various partitioning indices of sedentary time (SB, i.e., lying or sitting), non-sedentary time (NSB, i.e., standing, Light Physical Activity (LPA), Moderate to Vigorous Physical Activity (MVPA)), and MVPA time.*

Descriptive statistics of physical activity and postures	Mean	SD	Min	Max
PHYSCIAL ACTIVTIES & POSTURES (time proportions)				
Lying	0.0968	0.0960	0.0002	0.5363
Sitting	0.5157	0.1208	0.2275	0.7544
Standing	0.2713	0.0961	0.0925	0.6570
LPA	0.0358	0.0140	0.0109	0.0950
MVPA	0.0804	0.0330	0.0168	0.1798
PARTITIONING INDICES (sedentary)				
Median length (minutes)	6.58	2.72	1.73	15.43
Gini	0.7990	0.0608	0.6703	0.9368
Ratio (bouts/total)	0.9770	0.0196	0.8574	0.9958
PARTITIONING INDICES (non-sedentary)				
Median length (minutes)	4.37	1.59	1.52	10.90
Gini	0.8367	0.0502	0.7224	0.9365
Ratio (bouts/total)	0.9613	0.0336	0.7788	0.9964
PARTITIONING INDICES (MVPA)				
Median length (minutes)	1.80	0.66	1.00	5.18
Gini index	0.5954	0.0948	0.3301	0.8159
Ratio (bouts/total)	0.5280	0.1541	0.0797	0.8828

Table 2. *Arithmetical mean, standard deviation, minimum and maximum values of the health and control co-variables (n = 131).*

Descriptive Statistics of health and control variables	Mean	SD	Min	Max
Glucose (mg/dL)	96	9	75	123
LDL (mg/dL)	160	38	82	252
HDL (mg/dL)	53	13	25	98
Triglycerides (mg/dL)	109	54	40	306
BMI (kg/m^2)	25.77	3.89	16.03	37.56
Waist Circumference (cm)	87.34	12.17	57.00	116.00
Sex (0 = female)	0.64			
Age (years)	50.55	9.57	34.00	83.00
Education (categorical)	5.66	2.16	0.00	9.00
Income (categorical)	6.66	2.71	0.00	9.00
Nutritional index 1	0.00	1.55	-3.94	3.08
Nutritional index 2	0.00	1.37	-3.19	4.62

Associations between Behaviors and Health Outcomes

The following section presents the results of our models by health variable. Tables 3–5 include full results of the iso-temporal, compositional and partitioning models, respectively. Table 4 also includes the differences (or ratios) in the health values between an average time budget and the four hypothetical time budgets mentioned above.

Blood Glucose Concentration

No significant association was observed between the time volume of any behavior and blood glucose concentration. However, for a given behavioral time budget, partitioning patterns of both SB and NSB time were correlated with glucose level (Table 5). The Gini index was inversely correlated with glucose concentration: glucose level tended to be higher when short and long sedentary sequences contributed to the total time in an equal manner. For example, an increase of 0.1 in the Gini index (see Figure 2) was associated with a decrease of 3.0 mg/dL in glucose concentration. This counter-intuitive result was confirmed by the negative correlation with the ratio (sedentary time in bouts/total sedentary time): as the share of sedentary time spent in bouts decreased, the glucose level increased. A shift from a ratio of 0.9707 (1st quartile) to 0.9892 (3rd quartile) was associated with a decrease of 1.3 mg/dL in glucose level. Although the quadratic term for the ratio was significant, the relation was always negative in the observed range of the ratio values.

NSB (mostly standing) partitioning patterns also correlated with glucose concentration. The relation between the median bout durations of NSB and glucose was U-shaped, with a minimum median length reached at around 6 minutes. A median bout duration of 3 or 9 minutes was associated with –3.7 mg/dL and –3.3 mg/dL blood glucose, respectively.

***Table 3.** Estimated coefficient of linear iso-temporal substitution models. The coefficients are the estimated change in y due to reallocation of a time unit from one state (column) to another (row). Here, a unit represents the whole time budget. Hence, a reallocation of 1% (0.01) of the total time from sitting to standing, is associated with a change in fasting high density lipoprotein concentration (HDL) of 23.93 × (0.01) ≈ 0.24 mg/dL. Levels of p-values: † <0.1; * <0.05; ** <0.01; *** <0.001.*

Results of iso-temporal substitution models (coefficients and 95% confidence intervals)					
Health outcome	**Behavior**	**lie**	**sit**	**stand**	**LPA**
GLUCOSE (mg/dL)	sit	−1.67 [−19.71; 16.37]			
	stand	−9.76 [−33.16; 13.63]	−8.09 [−27.34; 11.15]		
	LPA	61.27 [−81.08; 203.62]	62.94 [−78.69; 204.57]	71.03 [−80.05; 222.12]	
	MVPA	−18.55 [−74.84; 37.74]	−16.88 [−73.26; 39.5]	−8.79 [−66.32; 48.74]	−79.82 [−254.1; 94.45]
LDL (mg/dL)	sit	−60.66 [−134.8; 13.48]			
	stand	−41.59 [−138.16; 54.98]	19.07 [−60.75; 98.89]		
	LPA	136.35 [−453.69; 726.4]	197.02 [−390.29; 784.32]	177.95 [−448.61; 804.5]	
	MVPA	−251.97 * [−484.86; −19.08]	−191.31 [−425.06; 42.45]	−210.38 † [−448.89; 28.14]	−388.32 [−1111.13; 334.48]
HDL (mg/dL)	sit	−5.07 [−27.36; 17.22]			
	stand	18.87 [−10.17; 47.9]	23.93 * [−0.07; 47.94]		
	LPA	−83.91 [−261.32; 93.5]	−78.84 [−255.43; 97.75]	−102.77 [−291.17; 85.62]	
	MVPA	32.1 [−37.92; 102.13]	37.17 [−33.11; 107.46]	13.24 [−58.48; 84.95]	116.01 [−101.32; 333.34]
log TRIGLYCERIDES (mg/dL)	sit	−0.58 [−1.35; 0.19]			
	stand	−1.33 ** [−2.33; −0.32]	−0.74 † [−1.58; 0.09]		
	LPA	3 [−3.14; 9.15]	3.59 [−2.53; 9.7]	4.33 [−2.2; 10.86]	
	MVPA	−4 *** [−6.43; −1.57]	−3.42 ** [−5.85; −0.98]	−2.67 * [−5.16; −0.19]	−7 † [−14.53; 0.53]
log BMI (kg/m²)	sit	−0.45 *** [−0.7; −0.2]			
	stand	−0.33 † [−0.67; 0.01]	0.13 [−0.16; 0.41]		
	LPA	−1.04 [−3.13; 1.05]	−0.59 [−2.68; 1.5]	−0.71 [−2.94; 1.51]	
	MVPA	−0.32 [−1.15; 0.51]	0.13 [−0.7; 0.97]	0.01 [−0.84; 0.86]	0.72 [−1.85; 3.3]
WAIST CIRCUMFERENCE (cm)	sit	−33.53 *** [−49.18; −17.87]			
	stand	−34.28 *** [−55.5; −13.06]	−0.75 [−18.39; 16.88]		
	LPA	−6 [−136.1; 124.11]	27.53 [−102.5; 157.56]	28.28 [−110.33; 166.89]	
	MVPA	−55.69 * [−107.27; −4.1]	−22.16 [−74.12; 29.79]	−21.41 [−74.38; 31.56]	−49.69 [−209.93; 110.54]

Table 4. *The top section of the table shows the estimated coefficient vectors $\hat{\beta}$ of compositional linear models. If **x** is the composition of behavior times, and **z** a vector of co-variables, the predicted outcome Y for individual i will be:* $\hat{Y}_i = \alpha + \langle\hat{\beta}, x_i\rangle_A + \langle\hat{\gamma}, z_i\rangle + \epsilon_i$*. The middle section of the table shows the normalized coefficient vectors, representing the direction along which a composition must be perturbed in order to achieve the largest effect* $|\hat{\beta}|$ *on* $\hat{Y}$*. The bottom section of the table show the change in* $\hat{Y}$ *associated with four scenarios of departure from the mean composition. The change is expressed as a difference* $\hat{Y}_i - \hat{Y}_M$ *or as ratio* $\frac{\hat{Y}_i}{\hat{Y}_M}$.

Results of compositional models						
COEFFICIENT VECTORS $\hat{\beta}$						
	Glucose	LDL	HDL	Log Trigl.	Log BMI	Waist Circum.
lie	0.3547	0.0198	0.0094	0.209	0.2053	0.748
sit	0.2318	0.0001	0.0008	0.23	0.1948	0.0563
stand	0.0361	0.0051	0.7822	0.1814	0.2002	0.0305
LPA	0.3041	0.975	0.0075	0.2188	0.1968	0.14
MVPA	0.0732	<0.0001	0.2001	0.1608	0.203	0.0253
p-value of the model	0.5326	0.2033	0.1858	0.0097	0.0208	0.0006
NORMALIZED COEFFICIENT VECTORS						
lie	0.2854	0.2347	0.1521	0.2183	0.3327	0.3924
sit	0.2305	0.1721	0.0975	0.3025	0.0987	0.1542
stand	0.0906	0.2167	0.3394	0.1347	0.1855	0.1236
LPA	0.2642	0.296	0.1461	0.2551	0.1253	0.2143
MVPA	0.1293	0.0805	0.265	0.0895	0.2578	0.1155
Vector norm $|\hat{\beta}|$	1.99	16.82	5.51	0.29	0.04	2.77
PREDICTED VALUES compared to mean composition						
Composition (%) [lie, sit, stand, LPA, MVPA]	Diff.	Diff.	Diff.	Ratio	Ratio	Diff.
[30,50,10,5,5]: 'couch potato'	3.21	13.81	−6.25	1.35	1.03	5.46
[5,70,10,5,10]: 'office worker'	1.36	−3.67	−4.22	1.12	0.99	0.58
[5,15,70,5,5]: 'doorman'	−1.35	12.38	6.35	0.86	1.02	0.07
[5,40,30,5,20]: 'active'	−0.82	−9.79	2.96	0.8	1.01	−1.16

*Table 5. Coefficients and 95% confidence intervals of linear regression models of various partitioning indices against health variables. The top section refers to sedentary bouts (lying or sitting). The middle section refers to non-sedentary behaviors (standing, Light Physical Activity (LPA) or Moderate to Vigorous Physical Activity (MVPA)). The bottom section refers to MVPA. Quadratic terms are reported when they significantly improve the model. Levels of p-values: † <0.1; * <0.05; ** <0.01; *** <0.001.*

Results of partitioning models (coefficients and 95% confidence intervals)						
Index	**Glucose**	**LDL**	**HDL**	**log Triglycerides**	**log BMI**	**Waist circumf.**
SEDENTARY BEHAVIORS						
Median (min.)	−0.41 [−1.1; 0.28]	−0.98 [−3.88; 1.92]	−0.51 [−1.37; 0.36]	0 [−0.03; 0.03]	0 [−0.01; 0.01]	−0.12 [−0.77; 0.53]
Gini	−29.8 * [−56.73; −2.87]	52.19 [−61.82; 166.21]	−6.82 [−40.93; 27.3]	0.13 [−1.07; 1.34]	−0.16 [−0.57; 0.25]	−11.2 [−36.87; 14.46]
Ratio	−3944.26 * [−7423.65; −464.88]	24.26 [−424.35; 472.86]	28.55 [−105.22; 162.33]	0.79 [−3.94; 5.53]	1.45 † [−0.12; 3.03]	26.26 [−72.96; 125.49]
Ratio²	2047.4 * [195.41; 3899.39]					
NON-SEDENTARY BEHAVIORS						
Median (min.)	−4.75 * [−9.31; −0.19]	2.31 [−3.07; 7.68]	−1.61 * [−3.19; −0.02]	0.02 [−0.04; 0.08]	0.02 ** [0.01; 0.04]	0.83 [−0.34; 1.99]
Median² (min.)	0.39 * [0.01; 0.77]					
Gini	13.31 [−20.3; 46.93]	123.44 † [−15.92; 262.81]	8.68 [−33.42; 50.77]	1.13 [−0.34; 2.61]	0.67 *** [0.21; 1.14]	25.05 † [−4.61; 54.7]
Ratio	−46.65 [−107.6; 14.3]	−33.56 [−291.17; 224.06]	−63.7 † [−139.72; 12.31]	0.61 [−2.11; 3.32]	0.28 [−0.64; 1.21]	−2.76 [−60.41; 54.9]
MVPA						
Median (min.)	−0.03 [−2.52; 2.46]	2.29 [−8.12; 12.69]	−0.14 [−3.25; 2.97]	−0.03 [−0.14; 0.08]	−0.01 [−0.05; 0.02]	−1.79 [−4.09; 0.52]
Gini	−1 [−21.87; 19.86]	18.45 [−68.83; 105.73]	9.6 [−16.41; 35.6]	−0.38 [−1.3; 0.54]	−0.12 [−0.44; 0.19]	−8.47 [−27.91; 10.96]
Ratio	−1.09 [−13.77; 11.6]	13.5 [−39.56; 66.57]	3.82 [−12.02; 19.65]	−0.2 [−0.76; 0.35]	−0.05 [−0.24; 0.14]	−2.35 [−14.22; 9.52]

Low-Density Lipoprotein (LDL), High-Density Lipoprotein (HDL), and Triglycerides

Time volumes of MVPA, but also to quiet standing, were clearly associated with triglycerides level, both in the iso-temporal and compositional models (Tables 3 and 4), and HDL in the iso-temporal model. In this model, reallocation of 1% of the time budget from sitting to standing was associated with an estimated increase in HDL level of 0.2 mg/dL (Table 3). The compositional model for triglycerides concentration confirmed the importance of standing and MVPA (Table 4). The predicted triglycerides concentration for a hypothetical profile ‘doorman’ dominated by standing was 14% lower than that predicted for the average time budget, while the predicted concentration for the profile “active” was 20% lower than for the average profile.

The models including partitioning indices suggest (Table 5) that, independently from the time budget, longer bouts of NSB were associated with lower HDL: an increase of three minutes in the median NSB bout durations was associated with a decrease of 4.8 mg/dL of HDL.

Body Mass Index and Waist Circumference

Only lying time was significantly associated with WC. In the iso-temporal model, reallocating 1% of the total time from lying to sitting, standing, or MVPA was associated with a decrease in WC of 0.34, 0.34 and 0.57 cm, respectively (Table 3). These results are supported by the corresponding compositional model (Table 4). Lying time was clearly associated with higher BMI (reallocating 1% of the total time from lying to standing associated with a decrease of 0.5% in BMI). Interestingly, neither BMI nor WC were associated with MVPA.

For a fixed time budget, the models suggest that longer median NSB bouts are associated with a strong increase in BMI, but not WC: an increase of 3 minute in the median length is associated with 6.2% increase in BMI. The model including the Gini index also points to a relationship between NSB partitioning patterns and BMI: as NSB time is accumulated through a smaller number of longer episodes, BMI increases.

Discussion

Distinguishing between a number of body postures such as standing, sitting and lying, alters our understanding of the relationship between physical activity and health. The beneficial effect of MVPA on the lipid profile has already been well established [40], and we also found positive associations of MVPA time with LDL and triglycerides levels. However, quiet standing, which was often classified as a sedentary behavior [20], was shown here to have similar positive associations with lipid profile (HDL and triglycerides levels). Our models suggest that persons standing during the day (such as the 'doorman' in Table 4) have a lower triglycerides level than the average, sedentary individuals.

Hence, increasing standing time proportion (e.g., at work) should be studied as a practical alternative to long periods of MVPA. In fact, our finding regarding the importance of standing is supportive of a few other studies, some of which were designed as interventions in the workplace [15,17]. The muscle activation required for posture maintenance, which is more important in standing than sitting[41, 42], could explain, at least partly, the beneficial association between standing time and lipid profile. Although both standing time was associated with a better lipid profile, it was not associated with glucose level. This can be explained by the lack of concentric muscle activation in standing, and a relatively low glucose uptake involved with Glucose Transporter type 4 (GLUT-4) translocation [43]. Likewise, the distinction between lying and sitting, which is made possible by the trunk sensor, reveals that lying diverges from sitting in the nature of its associations with BMI and WC in a slightly overweight population. Reallocation of time from lying to all other behaviors, including sitting, was significantly associated with a decrease in WC. This may also be related to a difference in energy expenditure between sitting and lying positions. More generally, the results suggest that, for some health-related aspects, physical activity should be regarded as a gradient in the following order: lying-sitting-standing-MVPA.

Results of the models accounting for the partitioning patterns of behavior time volumes shed light on several associations between physical postures/activities and health. Surprisingly, a certain pattern of partitioning of SB time, namely the existence of a large number of very short sequences of SB beside long sequences (expressed as a high Gini index and a relatively low ratio of bouts to total time) positively associated with fasting plasma glucose concentration. Moreover, glucose concentration was lowest in the population exhibiting a balance between short and long episodes in the accumulation of NSB time. To our knowledge, two studies found a negative correlation between blood glucose concentration and breaking up of SB: Carlson et al. [44] and Bellettiere et al. [17]. The former used a hip-worn accelerometer with the count per minute method, which cannot accurately distinguish between sitting and standing and does not account for sporadic behavior sequences, and the latter did not adjust for total sitting time. In fact, by looking only at bouts of SB, past studies overlooked sporadic sequences of SB, which should be regarded as interruptions in NSB behaviors (i.e., higher levels of physical activity). Many short sequences of SB very likely indicate that the subject does not perform long sequences of NSB, which could easily explain the higher glucose level.

In the same way, the NSB median bout duration was negatively associated with glucose level, up to a certain point. However, the U-shape relationship between NSB median bout duration and glucose suggests that several factors are at play. In fact, we found a positive association between the median duration of NSB bouts and BMI, as well as a negative one with HDL. These results supports the hypotheses proposed by Miles-Chan and Dulloo, according to which it is the efforts associated with frequent transitions between standing and sedentary postures that have a beneficial effect on these variables [41]. These hypotheses might also explain the U-shape observed for glucose level: prolonged standing and PA bouts might be beneficial for health, but shorter bouts points to a higher number of transitions from SB to NSB.

No significant association between health and partitioning patterns of MVPA time. This supports recent similar evidence observed in older British men [45] and is in accordance with the second 2018 US guidelines on physical activity, which removed the recommendation to perform MVPA in bouts of minimum 10 minutes duration [46]. In other words, any sequence of MVPA matters, regardless of its duration. However, it should be noted that the lack of correlation observed here might be due to the fact that our population did not suffer from severe weight issues.

Overall, compositional models agreed with iso-temporal substitution models. Besides being mathematically appropriate, the former allows a response fitting for any time budget while the latter estimate the response associated with time reallocation from a single part to another. Yet, the interpretation of the compositional models remains less straightforward. In addition, the model provides a significance level for the whole model, but not for each component [24]. We believe that combining both types of models can improve our understanding of the complex relationship between the behavioral time budget and health, and that further elaboration of these procedures of analysis should be a focus of future research.

A main limitation of this study is the impossibility to infer cause-and-effect relationships. In fact, causal links are often counterintuitive. For example, Ekelund et al. showed that it is adiposity that affects the volume of sedentary behaviors, and not inversely [47]. In addition, the relatively small sample size did not allow for investigation of possible effects of interactions with sex or age. Finally, we report results for a healthy population, which might not be extrapolated to other populations, such as those suffering from severe obesity, diabetes or other conditions.

Conclusion

By distinguishing physical activity and postures, the present study unmasked associations between standing time and lying time with key clinical outcomes, indicating that components other than MVPA play a key role in health. It also showed that the duration of MVPA bouts had no influence on health outcomes. These observations support the newly released U.S. physical activity guidelines that recommend to “Sit Less and Move More” and emphasized the importance of moving all along the day without necessarily trying to attain a specific duration bout of MVPA. Results also suggest that the relationships between fragmentation of SB/NSB and health are more complex than previously assumed and needs to be further investigated. In particular, very short behavior sequences, which have been overlooked in past studies, should be taken into consideration. Future research should focus on innovative ways to link patterns of behavior partitioning to health, use a refined categorization of behaviors and look for ways to implement the new resulting guidelines in the population.

Author Contributions: Conceptualization, I.D., A.B., B.C. and C.S.; methodology, I.D. and E.M.S.; software, I.D.; formal analysis, I.D.; writing—original draft preparation, I.D.; writing—review and editing, A.B., B.C., and C.S.; funding acquisition, B.C. and C.S; data collection—B.C. and F.T.

Funding: The RECORD study was funded by the National Research Agency (ANR), Institute for Public Health Research (IReSP), National Institute for Prevention and Health Education (INPES), National Institute of Public Health Surveillance (InVS), French Ministry of Research, French Ministry of Health, National Health Insurance Office for Salaried Workers (CNAM-TS), Regional Direction of Health and Social Affairs in Ile-de-France (DRASSIF), Regional Group of Public Health in Ile-de-France (GRSP), and Regional Direction for Youth and Sports in Ile-de-France (DRDJS).

Acknowledgments: The authors thank the Center for Preventive and Clinical Intervention (IPC) in Paris, the participants in the study and the collaborators who made this study possible.

Conflicts of Interest: The authors declare no conflicts of interest.

Appendix A:

A few explanations about operations in the Aitchison simplex

- Closure C:

$$C(\boldsymbol{a}) = \frac{\boldsymbol{a}}{\mathbf{1^t} \cdot \boldsymbol{a}}$$

- Perturbation $\oplus$**:**

$$\boldsymbol{a} \oplus \boldsymbol{b} = C[a_1 \cdot b_1, \dots, a_D \cdot b_D] \,,$$

where D is the length of the vector.

- Aitchison's inner product $\langle \mathbf{a}, \mathbf{b} \rangle_A$:

$$\langle \mathbf{a}, \mathbf{b} \rangle_A = \frac{1}{2D} \sum_{i=1}^{D} \sum_{j=1}^{D} \frac{\ln a_i}{\ln a_j} \frac{\ln b_i}{\ln b_j}$$

Appendix B:

Descriptive statistics using a compositional approach

Compositional Mean:

[Lie = 0.05645, Sit = 0.5472, Stand = 0.2794, LPA = 0.0364, MVPA = 0.0805]

Covariance Matrix:

Table B1. *Covariance matrix of the budget time of lying, sitting, standing, light physical activity and moderate-to-vigorous activity in the population.*

Behavior	Lie	Sit	Stand	LPA	MVPA
Lie	1.3334	−0.3297	−0.3355	−0.3016	−0.3666
Sit	−0.3297	0.2170	0.0380	0.0208	0.0539
Stand	−0.3355	0.0380	0.1636	0.0865	0.0475
LPA	−0.3016	0.0208	0.0865	0.1265	0.0678
MVPA	−0.3666	0.0539	0.0475	0.0678	0.1974

Appendix C:

Table C1. *Potentially confounding variables controlled for in the models.*

Name of variable	Type of variable	Value
Education	Continuous	0 = No diploma 1 = Certificat d'études primaire (completion primary school) 2 = Brevet élémentaire ou équivalent (completion of four first year of secondary education) 3 = Certificat d'aptitude professionnelle/Brevet d'études professionnelles (completion of 6 years of secondary vocational education) 4 = Baccalauréat professionnel (completion of vocational secondary cycle) 5 = Baccalauréat general (completion of general secondary cycle) 6 = Bac + 2 (completion of two years of higher education) 7 = Bac + 3 or Bac + 4 (Bachelor's degree) 8 = Bac + 5 or doctorat (Master's degree or higher)
Income (Total net revenues of household)	Continuous	0 ≤ 500 € 1 = 500–1000 € 2 = 1000–1500 € 3 = 1500–2000 € 4 = 2000–3000 € 5 = 3000–4000 € 6 = 4000–5000 € 7 = 5000–6000 € 8 = 6000–7000 € 9 ≥ 7000 €
Nutritional habits	The first two dimensions of a principal component analysis including all variables were used as continuous variables.	Intake of olive oil, vegetables, fruits, juice, meat, dairy products, desserts, sodas, wine, legume, fish, pizza, lean meat, nuts, commercial desserts; preference of olive oil over other oils; whether usually eats between meals

CHAPTER II

Implementation of physical activity episodes in people's daily life, following different segmentation schemes: consequences and feasibility

Reference:

Jong N, Debache I, Pan Z, Garnotel M, Lyden K, Sueur C, Simon C, Bessesen D, Bergouignan A. Breaking up sedentary time in overweight/obese adults on work days and non-work days: results from a feasibility study. International journal of environmental research and public health 15 (11), 2566 (2019).

Chapitre II: Implémentation de périodes d'activité physique dans la vie quotidienne suivant différents régimes de segmentation : études des conséquences et de faisabilité (résumé français)

Dans le chapitre précédent, nous avons étudié l'importance de divers aspects du comportement physique (la posture et l'activité, leur volume et leur segmentation) sur la santé. Malgré les conséquences sur la santé, le niveau d'activité moyen dans de nombreux pays, surtout dans le monde économiquement développé, reste bien en deçà des recommandations formulées par les spécialistes de la santé. Ce chapitre porte sur une expérience d'implémentation de deux programmes d'activité physique, suivant chacun un régime de segmentation différent, dans la vie quotidienne de personnes essentiellement sédentaires et en surpoids. Le but de l'étude était d'estimer le niveau d'adhérence des sujets en fonction du programme, déterminer leurs impressions des programmes d'activité physique et enfin d'analyser les conséquences sur des mesures clef de santé métabolique (index de la sensibilité à l'insuline).

L'étude a été menée sur 22 adultes (19-45 ans, 10 hommes et 12 femmes) de la région de Denver au Colorado (USA) souffrant de surpoids ou d'obésité modérée (indice de masse corporelle 27-33 kg/m²), ayant une profession sédentaire et ne satisfaisant pas les niveaux d'activité physique recommandés. Les individus participant à l'étude ont dû suivre trois programmes d'activité physique, d'une durée de trois jours chacun: 1) SED : les individus mènent leur vie sédentaire habituelle, en évitant des séances d'exercice structurées ; 2) MICRO : les individus doivent effectuer, pendant 9 heures consécutives de la journée, des séances d'activité physique modérée (marche rapide) de 5 minutes toutes les heures et maintenir leur mode de vie sédentaire pendant le reste de la journée; 3) ONE : les individus doivent effectuer une séquence d'activité physique modérée (marche rapide) de 45 minutes en une seule fois et maintenir leur mode de vie sédentaire pendant le reste de la journée. Au matin du 4ème jour, la concentration sanguine (à jeun) en glucose et en insuline a été mesurée. Sur cette 4ème journée, une auto-évaluation subjective du niveau de fatigue et de vigueur a également été collectée auprès des sujets. Les individus ont porté un accéléromètre à la taille et un inclinomètre sur la cuisse durant les trois programmes, afin d'observer le niveau d'adhérence aux instructions et les changements de l'activité physique journalière.

Les résultats ont montré que, par rapport à la condition de contrôle SED, le volume d'activité physique modérée-à-vigoureuse total mesuré par l'accéléromètre a augmenté aussi bien dans le programme ONE (+40.2 et +36.0 minutes en moyenne selon que le jour était ouvré ou pas) que dans le programme MICRO (+23.4 et +21.6 minutes, respectivement). Bien que les programmes d'activité n'aient pas induit une baisse de la glycémie à jeun, la concentration d'insuline a baissé dans ONE et MICRO. La différence entre ces deux conditions, quant à elle, était quasiment nulle. Les jours ouvrés, le niveau de vigueur perçu était significativement plus élevé dans ONE et MICRO que dans SED. Quant à la fatigue perçue, les jours ouvrés, elle était plus élevée dans ONE que dans MICRO.

Au niveau de l'adhérence, nos résultats montrent clairement que l'implémentation d'une seule longue séquence d'activité physique dans la vie quotidienne résultait en un niveau est plus facile que l'implémentation du même volume réparti en courtes séquences. Alors que le volume d'activité *prévu* était le même, le volume *effectué* était presque deux fois plus important dans ONE que dans MICRO. Cependant, le niveau de fatigue légèrement plus élevé dans ONE que dans MICRO (les jours ouvrés) pourrait signifier que la condition ONE exige une endurance plus importance et qu'elle serait donc plus difficile à implémenter au long terme.

Nos résultats suggèrent que les instructions sont mieux suivies sur le lieu de travail les jours ouvrés que lorsque les individus restent chez eux le weekend. Ce type d'intervention devrait donc être privilégié dans le cadre du travail.

Concernant les conséquences de santé, le fait que, dans les deux programmes actifs, la concentration en insuline a baissé alors que la glycémie est restée constante indique que l'activité physique augmenterait la sensibilité à l'insuline. L'effet de segmentation n'a pas pu être établi dans cette étude, puisque les mesures de santé ne diffèrent pas entre ONE et MICRO. Cependant, il faut tenir compte du fait que la baisse de la concentration d'insuline a été pratiquement la même dans les deux conditions alors que le volume d'activité était nettement plus élevé dans ONE que MICRO.

Il convient de souligner que la période d'expérimentation de chaque programme a été relativement courte (3 jours). En effet, une période plus longue aurait permis de mesurer l'endurance dans l'effort et donner une meilleure idée de la faisabilité d'une implémentation de programme d'activité comme solution permanente au problème de santé publique que représente l'inactivité physique. De plus, le fait que les sujets savaient que leur activité était enregistrée sur l'accéléromètre pourrait donner une estimation biaisée de l'effet qu'aurait un tel programme sans monitoring par accéléromètre.

Malgré son faible effectif, cette étude ouvre des perspectives intéressantes sur les possibilités de combattre la pandémie de l'inactivité physique à travers des séquences d'activité structurées, en particulier dans le milieu du travail. De plus, les résultats encourageants sur l'effet de tels programmes sur le moral des participants indiquent que la perte de temps pourrait être au moins partiellement compensée par un bien-être et une productivité accrue au travail.

Article

Breaking up Sedentary Time in Overweight/Obese Adults on Work Days and Non-Work Days: Results from a Feasibility Study

Nathan P. De Jong [1], Isaac Debache [2,3], Zhaoxing Pan [4], Mael Garnotel [5,6], Kate Lyden [7], Cédric Sueur [2,3], Chantal Simon [6,7], Daniel H. Bessesen [1,8,†] and Audrey Bergouignan [1,2,3,*,†]

[1] Division of Endocrinology, Metabolism, and Diabetes and Anschutz Health and Wellness Center, University of Colorado, School of Medicine, Aurora, CO 80045, USA; nathan.dejong@ucdenver.edu (N.P.D.J.); daniel.bessesen@ucdenver.edu (D.H.B.)
[2] Institut Pluridisciplinaire Hubert Curien, Université de Strasbourg, CNRS 67000 Strasbourg, France; isaac.debache@iphc.cnrs.fr (I.D.); Cedric.sueur@iphc.cnrs.fr (C.S.)
[3] UMR 7178 Centre National de la Recherche scientifique (CNRS), 67000 Strasbourg, France.
[4] Department of Biostatistics and Informatics, Anschutz Medical Campus, University of Colorado, Aurora, CO 80045, USA; zhaoxing.pan@ucdenver.edu
[5] CARMEN, CRNH, INSERM U1060/University of Lyon 1/INRA U1235 Lyon, France; ext-mael.garnotel@chu-lyon.fr
[6] Laboratoire de Biochimie CHLS 69310 Pierre Bénite, France; chantal@simon-bertrand.com
[7] KAL Research and Consulting LLC, Denver, CO 80002, USA; katelyden6@gmail.com
[8] Denver Health Medical Center, Denver, CO 80204, USA.
† These authors contributed equally to this work.
* Correspondence: Audrey.bergouignan@iphc.cnrs.fr; Tel.: +33-388-1069-14

Abstract

Office workers are vulnerable to the adverse health effects of sedentary behavior (i.e. sitting time). Increasing physical activity and preventing time spent sitting is an occupational health priority. This randomized crossover design study compared the short-term (3-days) effects of hourly interruptions of sedentary time with 5-min micrrobouts of activity for 9 hours (MICRO) to a sedentary control condition (SED) and a duration-matched continuous single bout of physical activity (45-min/d, ONE) condition on inclinometer-derived sitting-time on work and non-work days in sedentary overweight/obese adults. Differences in sitting/lying, standing, stepping, number of sit/stand transitions, time spent in moderate and vigorous activity (MVPA), energy expenditure, self-perceived vigor and fatigue, and insulin sensitivity were also examined. Twenty-two participants (10M/12F; 31.7 ± 1.3 year old BMI 30.4 ± 0.5 kg/m^2) completed all conditions. No between-condition effects were observed in sitting-time and sit/stand transitions. Both interventions increased daily steps, MVPA and energy expenditure with increases being greater in ONE than MICRO. Feelings of vigor and fasting insulin sensitivity were also improved. Participants reported less fatigue with MICRO than SED and ONE. Both interventions increase physical activity and energy expenditure in occupational and leisure-time contexts. The sustainability of these effects over the long term and on health outcomes will need to be tested in future studies.

Keywords: sedentary behaviors; sitting; microbouts; physical activity; MVPA; activity energy expenditure; vigor; fatigue; insulin sensitivity

Introduction

Sedentary behavior, i.e. sitting time, has been associated with adverse health outcomes including body mass index, cardio-metabolic outcomes, mental health and premature mortality [1–9] and has emerged as an important public health concern [10]. In addition to total daily sitting time, prolonged unbroken sitting time has been negatively associated with cardiometabolic health biomarkers [11,12].

Over the past few decades, advances in technology and computer-based tasks have increased time spent sitting at the workplace [13]. It has been found that office-based employees spend 66% of their total work time sitting with 25% of total sitting time in bouts longer than 55 minutes [14]. These changes in the workplace have been associated with reduced daily occupational energy expenditure. Since the 1960s, in the USA and the UK, population levels of occupational physical activity have declined by more than 30% [15]. Facing this developing public health challenge, the World Health Organization has recently published new guidelines for employers to promote healthier occupational environments [16]. Among the four major components of the guidelines, limiting prolonged sitting and increasing physical activity is one of them. While guidelines exist, they still need to be translated into practical strategies that can be implemented on a large scale. In this context, there has been increasing interest in understanding the efficacy of a broad range of interventions targeting sedentary behavior in the workplace.

A growing number of studies have examined environmental changes in the occupational setting to reduce sitting time such as active workstations and include sit-to-stand desks, treadmill desks and seated active workstations utilizing portable pedal machines [17–19]. These interventions have shown mixed results. While individual sit-to-stand desk interventions have not been shown to decrease sedentary time [20], interventions with multi-level components targeting the individual but also the social and built environment showed that stand-up desk options reduce sitting and increase standing time [21]. However, no effect on stepping time was observed. A personalized consultation with weekly emails that aimed to reduce prolonged sitting time did not decrease total daily sedentary time but reduced the occurrence of sedentary bouts of more than 30 min [22,23]. Another study using hourly computer screen prompts and text messages to break up sitting decreased total time spent sitting and increased the number of daily steps, but failed at increasing the number of sit-to-stand transitions [10]. Another goal of these interventions is to increase energy expenditure. The implementation of treadmill desks and seated active workstations can reduce daily sitting time, increase time spent in physical activity [24,25] and almost triple the energy expenditure of that measured while sitting. For example, walking at 1.8 km/h can induce an expenditure above 0.41 MJ/h, which could beneficially impact energy balance if sustained for several hours per day [26,27]. However, long-term adherence to these interventions (12 months) are poor [24,25], treadmill desks are costly and present a safety hazard. Therefore, a cost effective, easy to implement intervention that can reduce total time spent sitting, prevent prolonged sitting bouts as well as increase time spent active and energy expenditure is still needed. Implementing frequent short bursts of walking could fulfill these requirements.

Such interventions have already been tested in the laboratory setting. Past studies showed beneficial effects of frequent interruptions of sitting time with short bouts of activity varying in mode, frequency, duration and intensity on metabolic, cognitive and hemodynamic outcomes [28–39]. Regardless of adiposity, sex and age frequent interruptions of sedentary actives with walking breaks have been associated with attenuated postprandial plasma glucose and insulin concentrations in obese and type 2 diabetic adults [31–40]. We have shown that interrupting sedentary behavior with short bursts of treadmill walking increases self-perceived feelings of energy, vigor and mood and decreases feelings of fatigue throughout the day in normal weight adults [0]. The effect of such an intervention on the profile of physical activity and energy expenditure in free-living conditions is unknown.

While the workplace has been identified as a priority setting for addressing sedentary behaviors, it may be important to target sedentary behaviors in other contexts such as on non-working days. Non-working days also comprise a large portion of a working adult's week and have also been associated with a large amount of time attributed to sedentary activities [40]. Because workers who spend more time in sedentary pursuits during work hours do not compensate by being more active in non-working periods [20], there is a need to test interventions that aim at reducing time spent sedentary both during work days and non-work days outside of the controlled laboratory environment.

Based on the data generated by the past intervention studies conducted in the laboratory setting and the real-world, we hypothesized that an intervention aimed at breaking up sedentary time with short bouts of activity could attenuate time spent sitting, increase daily physical activity and energy expenditure, and positively impact metabolic health and well-being in office workers. The purpose of this study was to test the feasibility to implement such an intervention over a short period of time (3-days) in the daily life of overweight sedentary male and female adults during work days and non-work days. To test whether the effects on time spent sitting, time spent physically active and energy expenditure were due to the frequent interruptions of sedentary time with short bouts of activity or to the total time spent active, we used a three arm cross-over randomized design. Frequent interruptions of sedentary time with short bouts of physical activity were compared to a duration-matched single continuous bout of physical activity, and a sedentary control condition. Further, we compared the effect of the interventions on self-perceived vigor and fatigue and an index of insulin sensitivity. Finally, we assessed how difficult it was for participants to implement these interventions in their daily life on work days and non-work days.

Methods

Participants

This study was approved by the Colorado Multiple Institutional Review Board (COMIRB) and was in accordance with the Declaration of Helsinki (COMIRB# 14-0429). Eligible participants were between 19-45 years old with an occupation that requires sitting time, had a body mass index (BMI) between 27–33 kg/m^2, were weight stable for at least 3 months, insulin sensitive (fasting plasma insulin concentration below 25 μIU/mL), and self-reporting > 6hrs/day of occupational sitting. All women enrolled in the study were pre-menopausal and could use birth control medications. Exclusion criteria included clinically diagnosed diabetes, taking glucose- and/or lipid-lowering medications, dyslipidemia, smoking, or meeting the American College of Sports Medicine (ACSM) physical activity recommendations (>150 min/week MVPA). Participants were recruited between October 2014 and October 2016 from newspaper advertisements, public announcements, and flyers in the Denver and Aurora areas in Colorado, USA. Participants were randomized to one of three possible trial-condition orders using balanced blocks separately prepared for male and female participants. The study statistician (Z.P.) prepared the computer-generated randomization lists and sealed envelopes for randomization [41].

Study Design

Eligible volunteers completed three separate 3-day trial phases under free-living conditions. The study phases were separated by a 28-day wash out period and women were all studied in the follicular phase of their menstrual cycle. All the study related visits were conducted at the Clinical and Translational Research Center of University of Colorado (CTRC). The three trial conditions were administered in random order:

Sedentary (SED)*:* Free-living subjects maintained their usual levels of daily activity during the three days of measurement and were asked to refrain from structured exercise.

Sedentary + 1 continuous bout of activity (ONE)*:* During the 3-days of measurement, subjects were asked to perform 45-min of moderate-intensity walking once per day and maintain their usual sedentary lifestyle the rest of the day.

Sedentary + microbouts of activity (MICRO)*:* During the 3-days of measurement, participants were asked to perform a 5-min bout of moderate-intensity walking bout each hour for 9 consecutive hours throughout the day and maintain their usual sedentary lifestyle the rest of the time.

For both interventions, the intensity of the activity was defined during the screening visit. On each day of measurement, participants were asked to complete a diary log and record the time the participant went to sleep and woke up from sleep, the time the bouts of physical activity were performed and if it was a work day or not.

Screening Visit

Subjects were screened, consented and underwent a review of medical history and physical examination and a blood draw to verify fasting plasma insulin concentrations for eligibility. Resting Metabolic Rate (RMR) was measured by indirect calorimetry for 30 minutes in the fasted state, under resting conditions and at thermoneutrality. Body composition including fat-free mass (FFM) and fat mass (FM) was measured by dual energy X-ray absorptiometry (DXA, Hologic Delphi-W, Bedford, MA, USA). The short version of the International Physical Activity Questionnaire (IPAQ) was completed to assess habitual physical activity and time spent sitting [42]. Subjects then performed an incremental exercise test on a treadmill (increments of 0.3 miles/hr every 2-min) to determine a walking pace that was then prescribed for ONE and MICRO conditions. For each exercise level, subjects rated their perceived effort on a Borg scale from 0 (very light) to 20 (maximal exertion). The aim was to identify the walking speed that subjects associated with a perceived exertion level of 13 (somewhat hard). Subjects were instructed to walk at this pace for each bout of activity during the intervention.

Measurement of Time Spent Sitting/Lying, Standing, Stepping and Daily Steps

Time spent sitting/lying, standing, stepping and daily steps were quantified using an ActivPAL™ triaxial accelerometer/inclinometer (PAL Technologies Ltd, Glasgow, Scotland) during the three days of measurement in each condition. Participants were instructed to wear the monitor at all times. The device was worn midline on the anterior aspect of the thigh and wrapped with a nitrile sleeve, allowing for 24 hr measurement. The monitor produces a signal related to thigh inclination and is a valid and reliable measurement tool for determining posture and motion during activities of daily living [43]. When the monitor is oriented horizontally, it classifies the activity as sitting/lying. Vertical positioning of the monitor is classified as standing. Step cadence and number of steps were recorded by the monitor when a participant was walking.

The ActivPAL™ has been validated for use in adults to distinguish between sitting/lying, standing, and stepping activities [44–47]. Data event files from the

ActivPAL™ were used to quantify sitting/lying, standing, and stepping time. In these files, the ActivPAL™ records each time an activity changes and the time that the activity changed. Sitting/lying, standing, and stepping time were calculated by summing the duration of each event and the number of breaks from sitting time were quantified as a transition from sitting/lying to either standing or stepping. Sitting bouts lasting longer than 30-min and 60-min were also used to test the effect of the conditions on the sitting bout length. A customized R program (www.r-project.org) was used to convert the event data file to a second-by-second data file to estimate additional metrics of sedentary behaviors and time spent sitting/lying, standing, stepping. The following metrics of sedentary behaviors were computed over 24 hr: total sedentary time (total time spent in sitting/lying events), total breaks in sedentary time (number of times a sitting/lying event was followed by a standing or stepping event), and time (minutes/day) in sedentary bouts ≥30 and ≥60-minutes. The same outcomes were also reported as percentage of waking time. Because sleep time was removed, we assumed that sitting/lying time mainly corresponded to sitting time during waking hours. The R package (PAactivPAL) is available for researchers to generate these metrics [47].

Measurement of Physical Activity Intensity, Activity Energy Expenditure and Physical Activity Level

Activity energy expenditure (AEE) and time spent in different activity intensities were determined using the ActiGraph GT3X tri-axial accelerometer (ActiGraph, Pensacola, FL, USA). Participants were instructed to wear the accelerometer during wake time by attaching it to their right hip directly above their right knee using an elastic belt that was provided. A sampling rate of 30-Hz was used. After each of the 3-day study conditions, data were downloaded using the Actilife 6.13 software provided by the manufacturer and AEE per minute (J/kg/min) was estimated using the 'Freedson vector magnitude combination model' [49,50]. Total energy expenditure (MJ/d) was calculated as (AEE + RMR) / 0.9, where RMR was resting metabolic rate (MJ/d). Physical activity level (PAL) was calculated as the ratio between TEE over measured RMR. Cut-points of <1.5 and <3 METs and >3METs (metabolic equivalents) were used for very light intensity activity, light intensity activity and moderate-to-very vigorous activity, respectively.

Minute-data during waking hours were summed to obtain data per day. Although sedentary behavior has been defined as activities with an energy expenditure below 1.5 METs while in a sitting, reclining or lying posture [50], activities with METs <1.5 were referred to as very light intensity activity in our study. By only measuring energy expenditure without recognition of the concomitant posture, we are including activities such as standing that are not sedentary activities. By choosing the term "very light intensity activity" we are more conservative and avoiding any misinterpretation.

Perception of the Challenges Associated with the Conditions, Self-Perceived Vigor and Fatigue

At the end of each intervention or control day participants filled out online 100 mm visual analog scales (VAS) designed to capture their perception of the study condition [51]. The VAS addressed the following question *"Please indicate on the scale how challenging you found the day."* The anchors for this question were "*Extremely Easy"* and *"Extremely Challenging."* Immediately after the first survey, participants then completed an online modified version of the Perception of Mood survey (POMs) to assess changes in feelings of vigor and fatigue [52]. Only the POMs-Fatigue (POMs-F; $n = 7$ items) and the POMs-Vigor (POMs-V; $n = 8$ items) subscales were used for analysis.

Plasma Metabolic Outcomes

The morning after each 3-day trial, the participants reported to the CTRC for a fasting blood collection which was analyzed for glucose and insulin. Whole blood was added to a preservative (3.6 mg EDTA plus 2.4 mg glutathione in distilled water). Insulin concentrations were measured using a standard double antibody radioimmunoassay (EMD Millipore, St. Charles, MO, USA). Serum glucose concentrations were determined using the hexokinase method (Wako Diagnostics, Mountain View, CA, USA). These analyses were performed on the Beckman Coulter AU480 Chemistry Analyzer (Brea, CA, USA).

Statistical Analysis

Based on the diary log information, data were recorded on 43, 47 and 43 work days while in SED, ONE and MICRO conditions, respectively. Consequently, 23, 19 and 23 study days were non-work days when participants were in SED, ONE and MICRO conditions, respectively. The analysis of the working status effect (work day versus non-work day) was *a posteriori* analysis. This is why the number of work days and non-work days are unbalanced across the three conditions and the work status.

If there was more than one measure assessed at different days per condition and work status, the mean value of the repeated measures served as outcome in the model. Linear mixed models were used to test differences in the two activity monitor outcomes, self-perceived challenge, vigor and fatigue, with sequence, period, condition (SED, MICRO and ONE), work status (work day vs non-work day) and condition-by-work status interaction as fixed effects and subjects as random effect with a compound symmetry covariance. Contrasts were used, under this model, to test for the between work status difference under each condition, the between condition differences separately on workdays and non-work days and the between work status difference with respect to the between condition difference. No correction for multiple comparisons was applied. Fasting plasma insulin and glucose concentrations measured on the morning of day 4 were also analyzed using linear mixed model but work status was not considered. Indeed, within the three days prior to the blood draw, days could have been randomly spent at work or not, it was therefore impossible to know if the interaction between the condition and the work status had any influence on index of insulin sensitivity. Data are expressed as mean ± SD, unless otherwise stated. All statistical analyses were performed with SAS 9.4 (SAS Institute, Cary, NC, USA).

Results

Subjects' Characteristics and Compliance with the Interventions

Subjects' characteristics are displayed on Table 1. On average over the 3-days of intervention, participants performed 97.6 ± 0.0% and 98.4 ± 0.1% of the prescribed physical activity bouts in MICRO and ONE, respectively. High levels of compliance with both interventions was attained despite reporting that performing the physical activity interventions was more challenging than spending a day being sedentary (Intervention effect: $p = 0.007$; Figure 1). While participants reported that MICRO was challenging to perform on work days ($p = 0.004$ vs. SED), ONE was perceived to be more challenging to comply with on non-work days compared to both SED ($p = 0.05$) and MICRO ($p = 0.04$).

Table 1. *Study participant's anthropological characteristics and habitual sitting time.*

Parameters	Males	Females	All
n	10	12	22
Age (year)	31.5 ± 7.4	32.0 ± 6.1	31.8 ± 6.6
BMI (kg/m²)	28.8 ± 2.9	31.7 ± 1.8	30.5 ± 2.7
FM (kg)	24.6 ± 4.3 ***	36.0 ± 4.7	30.9 ± 7.3
FFM (kg)	63.1 ± 9.9 ***	49.9 ± 5.0	56.0 ± 10.1
FM (%)	28.1 ± 2.4 ***	41.8 ± 2.4	35.6 ± 7.4
Self-reported sitting time (h/d)	9.0 ± 3.2	10.6 ± 1.1	9.5 ± 4.1

Data are presented as mean ± SD. $p < 0.0001$ vs. Female. *n*, number of subjects; BMI, body mass index; FFM, fat-free mass; FM, fat mass; Self-reported sitting time was estimated from the IPAQ, international physical activity questionnaire.

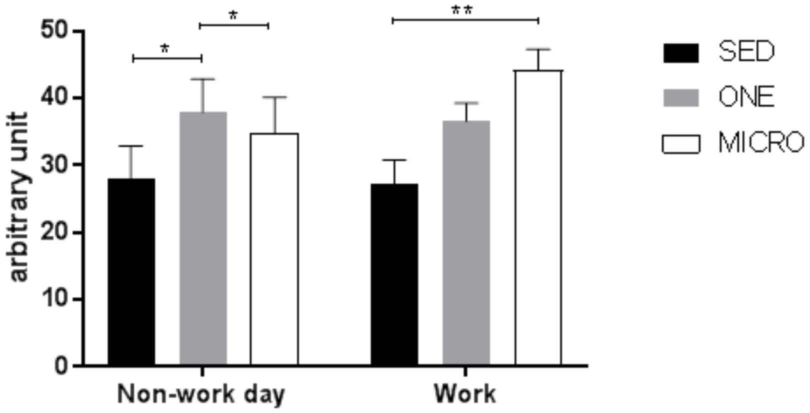

***Figure 1. Visual analog scale representing the perception of the challenges associated with the conditions**. At the end of each intervention or control day participants filled out online 100 mm visual analog scales (VAS) designed to capture their perception of the study condition. The VAS addressed the following question "Please indicate on the scale how challenging you found the day." The anchors for this question were "Extremely Easy" and "Extremely Challenging." SED, indicates the sedentary condition; ONE, indicates the one-bout intervention; MICRO, indicates the microbouts intervention. * $p < 0.05$, ** $p < 0.01$ vs. sedentary control.*

Effect of the Physical Activity Interventions on Time Spent Sitting/Lying, Standing and Stepping

Time spent sitting/lying, standing and stepping over 24 hr is reported in Table 2. One ActivPAL™ was lost and two were defective, we are therefore reporting data obtained in 19 subjects. Both MICRO (11. 4 ± 4.7 vs. 9.2 ± 3.4%, p = 0.009) and ONE (13.9 ± 3.5% vs. 9.2 ± 3.4%, p < 0.0001) increased the percentage of waking time spent stepping compared to SED on work days but not on non-work days. This resulted in 0.4 ± 0.1 hour more spent stepping in ONE than in MICRO (p = 0.01). As a result, the number of daily steps increased from 7125 ± 2554 to 12,257 ± 3145 in ONE (p < 0.0001) and 10,036 ± 4262 in MICRO (p = 0.0002) on work days; participants took more steps when performing ONE than MICRO (p = 0.005). Both ONE (+2967 ± 456, p = 0.005) and MICRO (+2841 ± 552, p = 0.02) led to a greater number of daily steps compared to SED on non-working days. However, time spent sitting and standing, the average duration of the sedentary bouts and the number of transitions from the sitting to standing position (index of breaking up prolonged sitting) were not significantly different across conditions and days (p > 0.05 for all). Surprisingly, the sitting bouts of more than 30 minutes tended to occur more often in MICRO than in both SED (p = 0.057) and ONE (p = 0.051) when in leisure contexts

Table 2. *Time spent sitting/lying, standing and stepping over 24hr and as percent of wake time.*

Physical Activity Outcomes	SED		ONE		MICRO	
	Non-work day	Work	Non-work day	Work	Non-work day	Work
Sitting/lying (hr/d)	9.8 ± 2.0	10.6 ± 2.3	9.6 ± 1.9	10.2 ± 2.4	9.6 ± 2.5	10.5 ± 2.2
Standing (hr/d)	3.5 ± 1.8	3.4 ± 1.8	3.0 ± 1.8	3.4 ± 1.5	3.6 ± 2.1	3.2 ± 1.9
Stepping (hr/d)	1.4 ± 0.5	1.4 ± 0.5	1.7 ± 0.4	2.1 ± 0.5 ***	1.7 ± 0.4	1.7 ± 0.7 **δ
Sitting (% waking time)	66.6 ± 14.2	68.4 ± 13.5	67.2 ± 12.7	64.5 ± 10.4	64.0 ± 15.3	67.8 ± 14.1
Standing (% waking time)	23.9 ± 12.3	22.3 ± 11.6	20.7 ± 11.9	21.4 ±8.8	24.0 ± 14.5	20.7 ± 12.6
Stepping (% waking time)	9.4 ± 3.6	9.2 ±3.4	11.9 ± 2.4	13.9 ± 3.5 ***	11.9 ± 2.9	11.4 ± 4.7 **δ
Sit-to-stand transitions (#)	48.8 ± 15.1	47.2 ± 17.7	42.5 ± 13.6	50.1 ± 22.3	46.1 ± 12.4	50.7 ± 21.3
Sitting bouts > 30-min (#)	5.6 ± 1.7	6.2 ± 2.2	5.5 ± 1.7	6.1 ± 1.7	6.7 ± 2.7 *δ	7.4 ± 2.7
Sitting bouts > 60-min (#)	3.1 ± 1.4	3.1 ± 1.5	2.6 ± 1.1	3.1 ± 1.6	2.3 ± 1.6	2.8 ± 2.0
Step count (#)	6409 ± 2843	7125 ± 2554	9376 ± 2387 **	12,257 ± 3149 ***	9250 ± 2291 *	10,036 ± 4262 **δδ

Data are presented as the mean ± SD. * p < 0.05, ** p < 0.01, *** p < 0.0001 compared to SED control within the same location. δ p < 0.05, δδ p < 0.01 different from ONE within same location. Sitting/lying (hr/d), number of hours per day spent siting; Standing (hr/d), number of hours per day spent standing; Stepping (hr/d), number of hours per day spent standing; Sitting (% waking time), percent of waking hours spent sitting; Standing (% waking time), percent of waking hours spent sitting; Stepping (% waking time), percent of waking hours spent stepping; Sitting bouts >30-min, number of sitting bouts lasting at least 30 minutes; Sitting bouts >60-min, number of sitting bouts lasting at least 60 minutes; Sit-to-stand transitions, number of times a participant rose from a seated position; Step Count (#), is the number of steps taken per day.

Effect of the Physical Activity Interventions on Time Spent in Very-Light, Light, Moderate and Vigorous Intensity Physical Activity

Time spent in very-light, light, MVPA during waking hours is shown in Figure 2. One ActiGraph GT3X was lost; data are reported for 21 subjects.

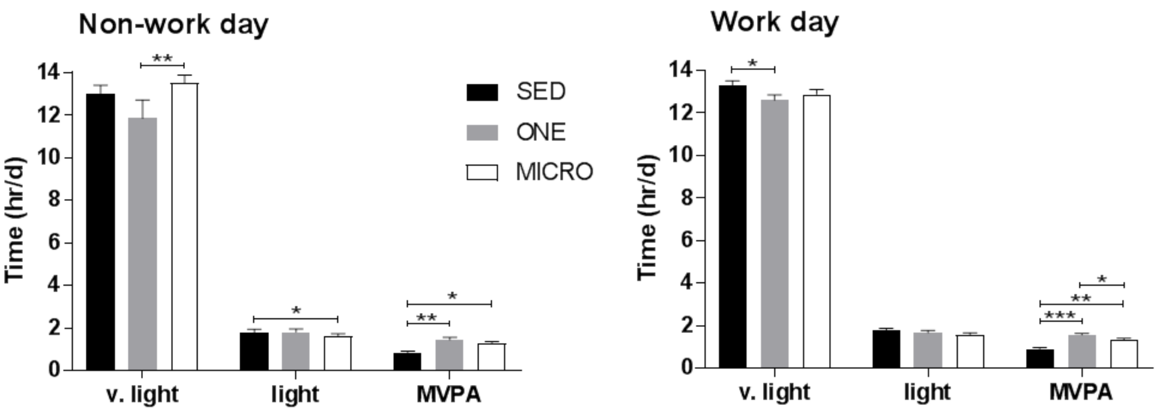

*Figure 2: **Waking time per day performing very light, light and moderate-to-vigorous intensity physical activity**. Accelerometry data collected from ActiGraph GT3X tri-axial accelerometer are displayed by location (work or non-work day) and by physical activity intensity. V. light, very light intensity physical activity; MVPA, moderate-to-very vigorous intensity physical activity; SED, sedentary condition; ONE, one-bout intervention; MICRO, microbouts intervention. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.0001$ vs. sedentary control condition.*

On work days, waking time spent in very light intensity activities tended to be lower in ONE compared to SED (12.5 ± 1.3 vs 13.5 ± 1.1 h/d, $p = 0.055$), but not different between MICRO and SED or ONE. Light intensity activities were not different across conditions ($p > 0.05$ for all). On non-work days, MICRO significantly reduced time spent in light intensity activities compared to SED (1.5 ± 0.5 vs. 1.9 ± 0.8 h/d, $p = 0.040$), but was associated with more time spent in very light intensity activities than ONE (13.7 ± 0.5 vs. 11.5 ± 0.5 h/d, $p = 0.002$). Both MICRO (work day: +23.4 ± 6.6 min, non-work day: +21.6 ± 8.4 min) and ONE (work day: +40.2 ± 6.6, non-work day: +36.0 ± 9.0 min) significantly increased time spent in MVPA compared to SED on both non-work and work days ($p < 0.01$ for all). On work days, MVPA was even greater in ONE than in MICRO ($p = 0.02$).

Effect of the Physical Activity Interventions on 24hr Activity Energy Expenditure and Physical Activity Level

Changes in MVPA induced by the physical activity interventions translated into parallel changes in AEE (Figure 3). Both MICRO and ONE significantly increased AEE compared to SED on both work and non-work days ($p < 0.05$ for all). Physical activity level (PAL) was significantly lower in SED compared to ONE on non-work days (SED: 1.46 ± 0.04, ONE: 1.62 ± 0.04, $p = 0.004$) and compared to both ONE and MICRO on work days (SED: 1.43 ± 0.03, ONE: 1.65 ± 0.03, $p < 0.001$, MICRO: 1.55 ± 0.03, $p = 0.003$). PAL was further higher in ONE than in MICRO on work days ($p = 0.008$).

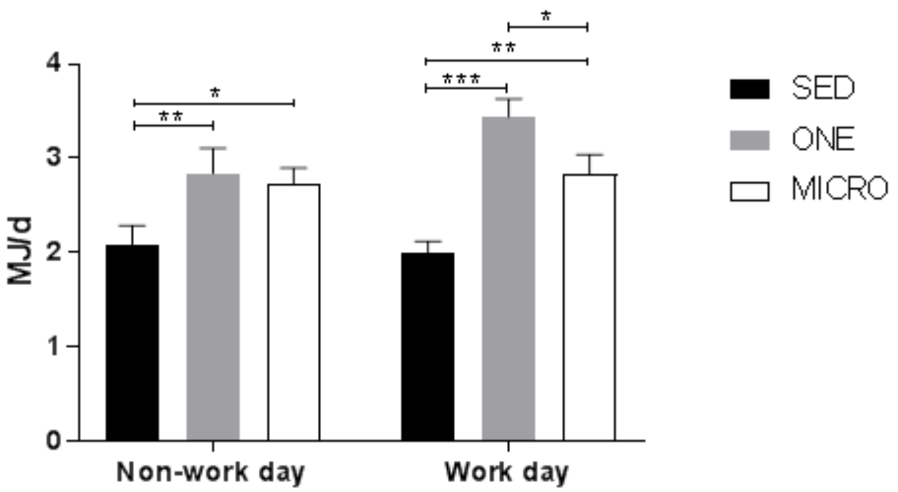

Figure 3. ***Activity energy expenditure.*** *The activity energy expenditure (MJ/d) estimated from ActiGraph GT3X tri-axial accelerometer is displayed by location (work or non-work day). SED, sedentary condition; ONE, one-bout intervention; MICRO, microbouts intervention. * p < 0.05, ** p < 0.01, *** p < 0.0001 vs. sedentary control condition.*

Effect of the Physical Activity Interventions on Self-Perceived Vigor and Fatigue

No significant differences in self-perceived vigor were noted across conditions on non-work days ($p > 0.05$ for all, Figure 4). On working days, participants reported a greater level of self-perceived vigor at the end of the day in both MICRO (386.7 ± 27.9, $p = 0.01$) and ONE (403.4 ± 28.1, $p = 0.002$) compared to SED (314.1 ± 28.0). They further reported feeling less fatigue on work days after a day performing MICRO than after a day performing ONE (–119.7 ± 52.5, $p = 0.03$). On non-work days, they tended to feel less fatigue on MICRO compared to both SED (-128.9 ± 65.6, $p = 0.054$) and ONE (–124.5 ± 67.3, $p = 0.069$).

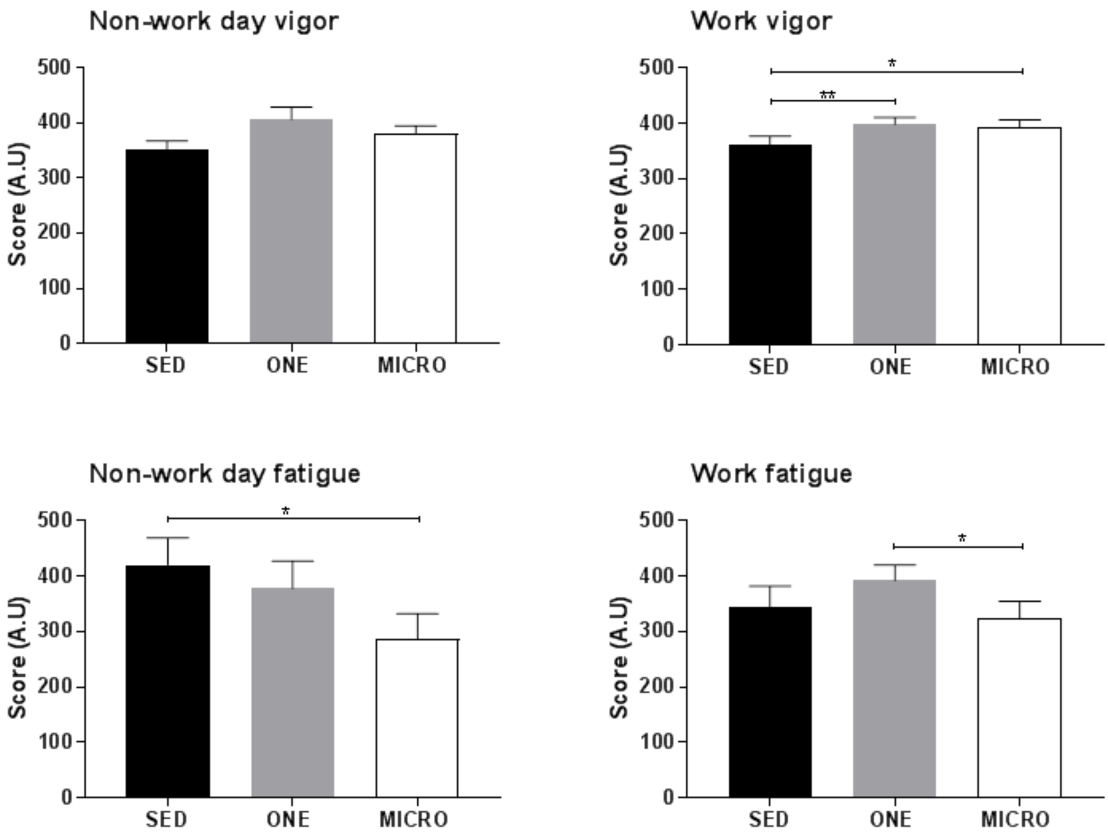

Figure 4. ***Self-perceived fatigue and vigor****. At the end of each study day participants rated their self-perceived feeling of fatigue and vigor (arbitrary unit). SED, sedentary condition; ONE, one-bout intervention; MICRO, microbouts intervention. * p < 0.05, ** p < 0.01 vs. sedentary control condition.*

Effect of the Physical Activity Interventions on Index of Insulin Sensitivity

On the morning of day 4, fasting insulin and glucose concentrations were measured (Table 3). MICRO and ONE significantly decreased fasting insulin concentration by 37.3% ($p = 0.03$) and 43.6% ($p = 0.02$) respectively compared to SED. Fasting glucose concentrations remained unchanged. As a result, insulin:glucose ratio, an index of insulin sensitivity, was reduced by both MICRO ($p = 0.03$) and ONE ($p = 0.02$) compared to SED, suggesting an improvement in insulin sensitivity. No differences were observed between the two active conditions.

Table 3. Fasting plasma glucose and insulin concentrations.

Parameters	SED	ONE	MICRO
Fasting glucose (mg/dL)	**90.1 ± 7.3**	**88.4 ± 7.7**	**88.7 ± 10.6**
Fasting insulin (uI/mL)	**10.8 ± 8.9**	**6.1 ± 3.0 ***	**6.7 ± 6.1 ***
I/G	**0.121 ± 0.101**	**0.069 ± 0.341 ***	**0.075 ± 0.063 ***

Data are presented as the mean ± SD. * $p < 0.05$ compared to SED control. I/G, insulin/glucose ratio.

Discussion

In this randomized feasibility study, we showed that sedentary, physically inactive, overweight/obese individuals were able to implement physical activity interventions consisting either of frequent bouts of activity or one continuous bout, the latter being more commonly promulgated by public health promotion initiatives and healthcare providers. Overall these two physical activity interventions had similar effects. Both interventions increased daily steps, MVPA, AEE and PAL on both working and non-working days compared to the sedentary control. These increases were more pronounced with a daily single bout of physical activity as compared to microbouts. The greater physical activity and energy expenditure were further associated with higher self-perceived feelings of vigor at the end of the day and improved fasting insulin sensitivity. Microbouts of activity were also associated with lower feelings of fatigue at the end of the day both on work days and non-work days. Neither of the interventions decreased time spent sitting or standing, the number of breaks from the sitting position and the average duration of a sitting bout.

Because office employees are vulnerable to the adverse health effects of prolonged sitting, an increasing number of interventions have targeted the work environment [53]. Strategies that promote body movements, such as passive pedaling or treadmill desks have been shown to increase physical activity and energy expenditure and to some extent reduce time spent sitting [18–20,25–27]. However, they are relatively expensive, can be a safety hazard and may be impractical to implement on a large scale. Therefore, we proposed that an intervention involving frequent short bouts of brisk walking could be an inexpensive, safe, easy to implement physical activity promotion intervention.

Contrary to our hypothesis, microbouts of activity spread out across the day did not reduce the number or duration of sitting bouts and did not increase the number of transitions from the sitting position to standing or stepping. This may be because asking individuals to break-up prolonged sitting nine times a day, every hour for nine consecutive hours to perform 5-min of walking is not a sufficient stimulus. In support of this interpretation, a recent study used hourly computer screen prompts or text messages to break up sitting. Sitting time was broken up with 7 minutes of walking to accumulate 30–60 minutes of walking per day. Additionally, there was an additional 6000 step count goal. This intervention was 7 days measured in overweight/obese and resulted in a decrease in total sitting time by 1.85 h/d on average [110]. Despite the frequency of activity being more frequent (every 30–60 min), this study also failed to show an increase in the number of sedentary breaks (sit-to-stand transitions) [10]. In our study, the number of sitting bouts longer than 30-min was even greater when participants were asked to perform microbouts of activity compared to single bouts on non-work days. This suggests that people tend to stay seated until they have to stand up and be active. Therefore, future studies may need to test specific interventions that primarily target breaks from sitting in addition to sitting time, daily steps or bouts of physical activity.

The American College of Sports Medicine, the American Heart Association and the American Diabetes Association recommend that adults perform at least 150-300 min/wk (21.4–42.8 min/day) of MVPA to maintain and promote cardiovascular health and insulin sensitivity [54]. Implementing frequent short bouts of 5-min brisk walking across the day in our study led to a significant 22.5–min/day increase in MVPA on average. In addition, the microbouts intervention produced an increase in AEE of 0.54 MJ/d (129 kcal/d) on non-working days and 0.78 MJ/d (187 kcal/d) on work days. It has been proposed that a very small energy gap – the difference between energy intake and energy expenditure –plays a role in weight gain [55]. A difference of 100 kcal/day at the population level could theoretically prevent weight gain in 90% of the U.S. adult population. Consequently, the increase in AEE along with the suppressive effect on appetite previously reported with microbouts of activity (at least in normal-weight individuals) [27,30] may help mitigate weight gain. Implementing microbouts of activity at work could be a viable strategy, among other strategies, to slow down weight gain. In addition, large prospective cohort studies of diverse populations have shown that an AEE of approximately 4.18 MJ/wk (1000 kcal/wk) is associated with lower rates of cardiovascular disease and premature mortality [54]. It would therefore be important to study the effect of this intervention over the long-term and verify whether a 1000 kcal/wk energy expenditure could be reached. Finally, our feasibility study showed that three days of microbouts of activity performed in daily life improves insulin sensitivity, which adds to the increasing body of data collected in the laboratory settings on the beneficial effect of frequent interruptions of prolonged sitting on insulin action [30–40,56]. This is the first study to show that an intervention using small bouts of activity promotes overweight-obese sedentary adults to comply with the current physical activity guidelines, at least in the short term. As a result, this strategy may have positive effects on body weight control and cardiometabolic health. However, we need to acknowledge that the single bout intervention we tested in the same subjects induced greater increases in MVPA (40 min/work day) and AEE 1.41 MJ/work day (+337 kcal/work day). The subjects thus attained a PAL of 1.65 that is characteristic of people who are moderately active. Future studies are needed to test the long-term effects of the microbouts of activity versus single bout of activity on the daily pattern of physical activity and energy balance regulation (appetite, energy intake, energy expenditure).

The long-term goal will be to test this type of intervention in the public on a large scale. The modern occupational environment promotes increased sedentary time [57], and has therefore been identified as an ideal environment to target sedentary behaviors. This is even more important because adults who spend more time sedentary at work do not compensate by being more active during non-working periods [58]. Interestingly, we showed that the beneficial effects of the microbouts intervention on physical activity and self-perceived fatigue were observed on both work and non-work days. This means that if implemented in occupational contexts this intervention, if sustained on weekends, could also increase physical activity on non-work days. A limitation is that instead of shifting time from very light to MVPA intensity activities as observed with the single bout of activity, the microbouts of activity increased MVPA in detriment of light intensity activity on non-work days. Another potential issue for future implementation of such intervention is the fact that participants reported the microbouts of activity to be more challenging to perform at work. But in our study, participants were the only employees performing these activities at their workplace. If the environment was designed to support breaking up sitting, participants may find this approach less challenging. It is well known that socio-ecological approaches acting on both the micro environment (individual) and macro environment (socio-professional environment, office layout, alternative work stations, active vs sitting meetings, etc.) are key when aiming to implement new interventions that change behavior for a sustained period of time. Developing strategies to self-motivate individuals in adopting this new behavior is also crucial [59]. The fact that our overweight/obese participants perceived less fatigue at the end of a workday performing the microbouts than a single continuous bout of activity, as we previously reported in normal weight individuals, could be used to encourage employers to incorporate microbouts of activity into the daily routines of their office employees [0]. Additionally, strategies aiming to reduce time spent sitting have not been shown to affect productivity or cognitive functions [28–30]. Most likely, a combination of the two interventions to target both occupational and non-work time may be the best approach. It could also provide individuals with different tools to choose from according to their mood that day at the office or outside the office.

Several limitations need to be acknowledged. The main limitation is that the study was conducted over 3-days and so conclusions about whether the weekly level of recommended MVPA could be reached and sustained for longer time periods cannot be made. The comparison between work days and non-work days was not *a priori* powered and led to an unbalanced number of days spent in the two different settings. Because participant's knew their physical activity was being tracked by two physical activity monitors there could have been an effect of increased activity [61]. Indeed Clemes et al. showed that wearing activity monitors for three days induces a spike in physical activity levels that regresses back to the mean after 7-days [60]. However, other studies have shown no evidence of reactivity to physical activity monitors [61,62]. In addition, the cross-over design may have limited the reactivity effect to the monitors. Another strength was that the pattern of physical activities was assessed using two complementary activity monitors, one specifically designed to detect changes in sitting and the other one designed to determine time spent in activities of different intensities and the associated energy expenditure. Finally, this feasibility study testing a novel lifestyle intervention to prevent sedentary behavior was conducted in overweight/obese, sedentary, physically inactive adults, which represent a high-risk group for metabolic diseases.

Conclusions

This feasibility study showed in overweight/obese physically inactive sedentary adults that regardless of the terms of the intervention, promoting physical activity led to an increase in physical activity and energy expenditure, and improved insulin sensitivity and vigor. However, none reduced total daily sitting time or the length of sitting bouts. This suggests that more efforts are needed in the workplace to increase physical activity along with a concomitant reduction in the number and duration of sitting bouts. It may be that frequent prompts to rise from sitting in combination with encouragements for either microbouts or single bouts of activity may represent the best overall strategy. This will need to be tested as part of a multicomponent intervention at the organizational, environmental and individual levels. Therefore, the overall public health message should communicate that any increase in physical activity can be beneficial when performed consistently over time.

Author Contributions: AB and DHB designed the study, collected, analyzed and interpreted data, and drafted the manuscript. NPDJ collected, analyzed and interpreted data, and drafted the manuscript. ID, MG, KL, CSu, CSi analyzed and interpreted data and drafted the manuscript. ZP performed statistical analysis and drafted the manuscript. All authors gave final approval of the manuscript to be published and agreed to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Funding: This work was supported by the National Institutes of Health grant numbers K99DK100465 (AB) and UL1 TR001082 (AB), NORC 5P30DK048520 (AB, DHB), CTRC UL1TR002535 (AB, DHB), 5P30DK048520 (ZP). This publication was also supported by Grant Number T42OH009229-10 (AB) from CDC NIOSH Mountain and Plains Education and Research Center. CSueur and ID are funded by a grant from the French national Research Agency (ANR-15-CE36-0005) and by the Grand Est Region (France). The contents of this publication are solely the responsibility of the authors and do not necessarily represent the official views of the funding sources.

Acknowledgements: The authors thank study volunteers for their time and participation. They are also very thankful for the University of Colorado Hospital CTRC staff.

Conflict of interest: The authors declare no conflict of interest. The funding sponsors had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, and in the decision to publish the results.

CHAPTER III

Deriving physical activity with wearable accelerometers

Chapitre III: Dériver l’activité physique à partir d’accéléromètres (résumé français)

Les études épidémiologiques visant à estimer les effets du comportement physique, mesuré objectivement, sur la santé s’appuient principalement sur des accéléromètres portés au corps. Cet outil représente une solution fiable et peu coûteuse et, grâce à sa petite taille, occasionne un inconfort relativement léger. L’utilisation de l’accélérométrie en épidémiologie s’est beaucoup répandue grâce à des études à grande échelle aux Etats-Unis (les différentes vagues de la National Health And Nutrition Survey, notamment).

Malgré sa popularité, il n’existe pas de consensus quant aux méthodes permettant de transformer le signal d’accélération en mesure de l’activité physique. Certes, un savoir-faire propre au domaine de l’épidémiologie de l’activité physique s’est développé avec la démocratisation de cet outil dans les années 2000 et au début des années 2010. Mais ce savoir-faire répondait aux besoins définis par les paradigmes de l’époque, axé sur la l’activité physique modérée-à-vigoureuse, et sur des spécifications techniques des appareils dépassés aujourd’hui : accéléromètre uniaxial, une capacité de stockage de données limitée et une voluminosité de l’appareil exigeant un port à la taille. De plus, l’utilisation massive d’appareils de type ActiGraph estimant l’intensité du mouvement à travers un système de *counts* (méthode d’agrégation opaque du signal d’accélération sur des périodes plus ou moins longues) a entravé le développement d’une accélérométrie propre à l’épidémiologie de l’activité, alors que l’importance de l’étude d’une nomenclature d’activités détaillée à haute résolution temporelle était progressivement mise en avant dans la littérature.

Pourtant, une littérature abondante sur l’accélérométrie, majoritairement indépendante du monde de l’épidémiologie, a pris son essor avec la popularisation d’applications, basées sur le signal accélérométrique, développées pour smartphones et smartwatches. Des méthodes d’apprentissage artificiel (*machine learning*) sont largement utilisées et se sont avérées très précises dans leur capacité à prédire le comportement physique à partir du signal. Cependant, ces méthodes demandent des compétences en science des données qui dépassent celles que possèdent la plupart des épidémiologistes. Surtout, l'hyper-paramétrage est tellement chronophage, et le coût en calcul informatique est tel, que ces solutions, malgré leur excellente performance, restent peu applicables et généralisables aux problèmes posés par l'épidémiologie de l'activité physique. Le but de notre étude a été donc de créer un algorithme à la fois performant et simple, rapide et robuste afin de dériver un large éventail de comportements à partir d'accéléromètres posés au corps.

L'apprentissage artificiel, permettant d'attribuer à des séquences de signaux d'accélération des classes de comportements prédéfinis (apprentissage supervisé), nécessite un jeu de donnés annotées comportant à la fois les signaux et leur comportements correspondants. En apprenant sur des exemples annotés, le modèle est entrainé à associer un signal d'entrée à un comportement, pour être finalement capable de prédire pour chaque signal un comportement inconnu. Pour entraîner notre algorithme, nous avons choisi le jeu de données DaLiAc (*daily life activities*), comportant des signaux d'accéléromètres et de gyromètres portés au corps (poignet, poitrine, taille et cheville) de 19 individus alors qu'ils réalisaient une série de 13 activités : 1) être couché au repos, 2) être assis au repos, 3) être debout au repos, 4) faire la vaisselle, 5) passer le balai, 6) passer l'aspirateur, 7) marcher, 8) monter les escaliers, 9) descendre les escaliers, 10) courir, 11) faire du vélo à une résistance de 50 Watt, 12) faire du vélo à une résistance de 100 Watt, 13) sauter à la corde. Ce jeu de données public ayant été souvent testé dans d'autres études, il nous a également permis de comparer notre algorithme à d'autres algorithmes récents.

Notre algorithme propose un traitement relativement simple des signaux. Un filtre passe-bas et passe-haut est appliqué au signal d'accélération, séparant ainsi l'accélération dynamique et celle due à la gravité. Chaque signal d'accélération est donc divisé en deux signaux indépendants, en plus du signal original, qui est conservé lui aussi. On applique une transformée de Fourier aux trois signaux, passant du domaine temporel au domaine fréquentiel. Une série de variables est calculée dans ces deux domaines (moyenne, écart-type, maximum, minimum, entropie des amplitudes fréquentielles etc.). Celles-ci sont fournies aux modèles qui apprennent à les associer aux classes de comportements. La classification s'opère dans un système hiérarchique : à un premier niveau, un modèle de classement général affecte chaque échantillon du signal à une méta-classe regroupant plusieurs comportements détaillés semblables. A un second niveau, un autre modèle « spécialisé « dans une méta-classe attribue un comportement détaillé à cet échantillon parmi ceux présents dans la méta-classe.

En utilisant les signaux filtrés, un système de classification hiérarchique et un petit nombre de variables à haut degré d'information, notre algorithme, basé sur une classification par régression logistique, surclasse les algorithmes présentés précédemment en termes de précision et de coût en calcul informatique. L'algorithme est simple et robuste, et a nécessité aucun hyper-paramétrage ou sélection de variable.

Notre étude teste l'algorithme proposé sur l'ensemble des 15 combinaisons d'appareils à partir des quatre avec lesquels les données ont été collectées (poignet+poitrine+taille+cheville, poignet+poitrine+taille, poignet+poitrine+cheville,..., poignet, poitrine, taille, cheville). Nous montrons qu'en conservant deux appareils au poignet et à la cheville, la précision du classement ne diminue quasiment pas par rapport à un classement basé sur l'ensemble des appareils (96,8% au lieu de 97,3 %). De même, en retirant les signaux de gyromètres, la précision baisse également de 0,5% seulement. Enfin, nous montrons qu'un modèle de régression logistique, malgré sa simplicité, reste plus performant que d'autres modèles : réseau de neurones à convolution, machine à vecteurs de support, *gradient boosting* et la méthode des k voisins les plus proches.

Malgré ces excellents résultats, il faut souligner que les données ont été collectées dans des conditions semi-contrôlées de laboratoire et se prêtent donc bien à l'apprentissage automatique. Des données moins « propres » pourraient accuser un taux de classements corrects plus bas. De même, notre méthode se base sur un fenêtrage des données à 5 secondes, alors que des données prises sur le vif pourraient demander un fenêtrage plus court compte tenu des changements rapides de comportement. En effet, appliquant des fenêtres d'une seconde (ce qui semble plus adapté aux conditions de vie réelles), la précision moyenne du modèle baisse de 2,9%. De plus, la dépense énergétique ne peut être estimée avec cet algorithme qu'en faisant correspondre les comportements prédits à des valeurs de dépense énergétique moyennes correspondantes, connues par ailleurs. Les variations en dépense énergétique au sein d'une activité dues à des intensités variables (par exemple différentes vitesses au sein de la catégorie « marche »), ne sont donc pas directement prises en charge par l'algorithme.

Malgré ces limites, notre algorithme reste, relativement aux autres, plus performant, polyvalent et étonnement simple. Il démontre que des solutions simples basées sur une compréhension profonde du problème peuvent mener à de meilleurs résultats que des méthodes ultra-performantes mais peu adaptées. Cet algorithme, nous le croyons, sera un nouveau point de départ pour de nouveaux modèles développés sur des données plus réalistes, comportant éventuellement dans le futur des informations détaillées sur la dépense énergétique.

A lean and performant hierarchical model for human activity recognition using body-mounted sensors

Isaac Debache[1], Lorène Jeantet[1], Damien Chevallier[1] , Audrey Bergouignan[1,2] and Cédric Sueur[1,3]

1.Institut Pluridisciplinaire Hubert Curien (IPHC) UMR 7178 Centre National de la Recherche Scientifique (CNRS), Université de Strasbourg 67000 Strasbourg, France.
2. Division of Endocrinology, Metabolism, and Diabetes and Anschutz Health and Wellness Center, University of Colorado, School of Medicine, Aurora, CO 80045, USA
3. Institut Universitaire de France, Paris, France

***Corresponding author:** Isaac Debache

Keywords: Accelerometers, sensors, human activity recognition, machine learning

Abstract

This article proposes a new machine learning algorithm for classification of human activities by means of accelerometer and gyroscope signals. Based on a novel hierarchical system of logistic regression classifiers and a relatively small set of features extracted from the filtered signals, the proposed algorithm outperformed previous work with a mean accuracy of 97.3% on the Daily Life Activity dataset. The algorithm represents also a significant improvement in terms of computational costs and required no feature selection and hyper-parameter tuning. The algorithm stilled showed a robust performance (96.8% mean accuracy) with only two devices (ankle and wrist) out of the four (chest, wrist, hip and ankle). The present work shows that low-complexity models can compete with heavy, inefficient models in classification of advanced activities when designed with a careful inspection of the data.

Introduction

Activity monitoring with wearable sensors has various scientific, medical and industrial applications, such as physical activity epidemiology [1], fall detection in the elderly population [2] or for smartwatch applications [3]. Among the existing sensors, accelerometers are regularly used for activity monitoring mainly because of their relatively high accuracy, low price and small size [4], [5]. To improve measurement reliability, accelerometers in activity monitoring are sometimes coupled with gyroscopes (measuring angular velocity), for instance in the smartphone [6]. Methods for human activity recognition (HAR) using wearable motion sensors were thoroughly investigated and reported in the scientific literature, and many studies demonstrated their ability to predict activity with a high level of accuracy [7], [8].

Existing HAR methods usually rely on supervised machine learning models to map between motion signals and activities. All methods rely on the assumption that different physical activities are reflected by different, characteristic signals and that it should be possible to discriminate between activities with appropriate, meaningful features extracted from the signal [8], [9]. HAR models can be divided into two main families: classical machine learning models and neural networks (often referred to as deep learning models) [10]. In the classical approach, activities are discriminated by means of handcrafted features extracted from segments of the signal in the time and frequency domains (e.g. mean, standard deviation or maximum frequency) [8], [9]. Such features have proven useful in discriminating activities in various models, such as tree-based models, support vector machines (SVM), logistic regression, k-nearest-neighbours (KNN), Naïve Bayes Classifier or hidden Markov models (HMM) [7], [9]. In contrast, neural networks can be fed directly with the raw signal and are automatically tuned in order to detect discriminative features [10], [11]. Neural networks have been proposed in different variants, such as convolutional neural networks (CNN) and recurrent networks [11].

The automatic feature detection of deep learning models makes them capable of detecting very complex, highly discriminative features and patterns in the data [10]. CNN drawing upon advances in computer vision have recently proved powerful in HAR and outperformed classical machine learning models (e.g. [12], [13]). Although very performant, deep learning models are very long to train and finding an optimal architecture for the task at hand is most often a tedious process [11]. The effectiveness of automatic feature learning comes thus at a high computational price, which makes it often more efficient to rely on human domain knowledge for feature extraction [10]. Furthermore, the long process of model selection makes the final model hardly generalizable to similar but different tasks [11], [14].

Classical supervised machine learning methods, in contrast, are easier to train but their shallow learning can makes them less performant in difficult classification tasks [10]. To make up for this deficiencies, researchers using classical model must handcraft a very large number of increasingly complex features, sometimes amounting to several thousand [8], [15]. As too many features can impair the performance of the models and makes training computationally impractical, researchers must engage in a process of feature selection in order to form a small subset of highly informative features, which are subsequently fed into the classification models [16]. This process of feature selection can be in itself complex [15] resulting in computationally expensive, inefficient and sometimes unclear classification algorithms.

Several studies demonstrated the usefulness of a hierarchical classification system for HAR in increasing accuracy while maintaining the algorithm reasonably simple [17]–[19]. This system consists in assigning precise target classes to samples in two steps. In a first step, a base classifier discriminates between meta-classes regrouping several similar target classes. In a second step, classifiers specific to each meta-class discriminate between the final target classes. With a strong base-level classifier, such systems can manually prevent potential misclassification [18] and combine different classifiers for different tasks, each 'specializing' in a different problem solving task [17]. Finally, a hierarchical system provides an interesting insight into the performance of the algorithm solving a basic classification, which can represent an objective *per se*.

The goal of this article is to propose a high-performance, fast and yet simple algorithm for HAR based on a careful inspection of the signal and a smart use of classification methods. We rely on a novel hierarchical system and a relatively small set of highly-informative features extracted from the filtered signals. We test our approach on the public *Daily Living Activity (DaLiAc)* dataset presented below [17].

Materials and Methods

The DaLiAc dataset

The DaLiAc (Daily Living Activity) dataset consists of the signals of accelerometers and gyroscopes placed on the chest, wrist, hip and ankle of nineteen adults performing thirteen daily activities in semi-controlled conditions. The activities includes a wide range of simple and complex activities: lying, sitting, standing, dish washing, vacuum cleaning, sweeping, walking, running, ascending stairs, descending stairs, bicycling with a resistance of 50 Watts, bicycling with a resistance of 100 Watts and rope jumping. Details about the subjects and the experimental designs can be found in [17].

Processing

Acceleration signals are known to be composed of a dynamic component (acceleration of the body) and a gravitational one. As a consequence, some authors suggested to apply a low-pass filter to the acceleration signal in order to isolate the gravitational component and infer the inclination of the device in space [8], [20]. Using a Butterworth filter (1^{st} order, with a threshold of 2 Hz), we separated the accelerometer signals into dynamic and gravitational components (AC and DC components, respectively). Unlike the widespread approach, we treated raw acceleration, AC and DC components as three separate signals all along the feature extraction process. AC and DC components reflect two different aspects of physical activity, orientation and motion, and as such should be treated as two independent signals. For instance, periodicity metrics extracted for the signals can be different, but equally interesting, when looking at orientation and motion over time. Thus, we ended up, for each sensor, with the following time-series: three total acceleration signals (along each axis), three AC, three DC and three gyroscope signals. All signals were down-sampled to 51.2 Hz (we sampled every 4^{th} datapoint from the original data) and normalized.

All signals were segmented along the time axis into windows of 5 seconds with a 50% overlap, as done by other authors [21], in order to make evaluation comparable with other algorithms tested on the same data [12].

Feature extraction

For each window, the following statistics were computed:

For the signal x in the time-domain, by window of length N :

- Mean, standard deviation, skewness and kurtosis;
- The following percentiles: [0, 5, 10, 20, 30, 40, 50, 60, 70, 80, 90, 95, 100];
- Range : max(x) – min(x) ;
- RMS: $\sqrt{\frac{1}{N}\sum_{i=1}^{N} x^2}$
- Zero-crossing: The number of times the signal crossed the mean.

In the frequency domain:

For all windows, we applied a Fourier transformation to the mean-subtracted signal $x' = x - \underline{x}$ to obtain the amplitudes {**X'ₖ**} for the frequencies $k \in \left[0\,, +\frac{N}{2} - 1\right]$. The following features were computed for all Fourier transformed series:

- Energy: $E = \sum_k |X'_k|^2$,
- Entropy: for the spectral density $P_k = \frac{|X'_k|}{\sum_k |X'_k|}$, $H = -\frac{1}{\log_2 \frac{N}{2}} \sum_k P_k \cdot \log_2 P_k$,
- Centroid: The sum of the frequencies k associated with the transform $\{\mathbf{X'_k}\}$, weighted by the spectral densities: $C = \sum_k k \cdot P_k$
- Bandwidth: The weighted mean absolute distance from the centroid $\sum_k |f_k - C| \cdot P_k$.
- Maximum frequency: $f_{argmax\,(\{X'_k\})}$

Classification

Classification was done using a two-level hierarchical system illustrated in Figure 1. For all classification tasks in the system, the following classifiers were tested: logistic regression (with a L2 regularization and a penalty coefficient equal to one); KNN with k=5; gradient boosting (500 estimators, selecting 10 features at a time) and SVM. For additional comparability, a convolutional network was also tested (architecture in Figure 2) taking as input the four signals (AC, DC, accelerometer and gyroscope) and their Fourier transform. Classification was done using all 15 possible combinations of device locations on the subjects' body (e.g. ankle, ankle+chest, ankle+chest+wrist,...).

We used Python's Scikit-learn[22] and Tensorflow [23] libraries for the analysis and, unless otherwise specified, their default parameters.

Evaluation method

In order to evaluate the performance of the proposed models, a leave-one-subject-out procedure is followed: models are tested against data from one subject after being trained on all the rest, for each subject of the 19 subjects in the dataset. This procedure was adopted by the first study on the dataset and followed by six subsequent studies (see Table 1), as it reduces bias in the accuracy estimator [17]. Models are tested against data from subjects they have never seen, hence hinting to their generalizability.

For all models, we report the mean and standard deviation of the accuracy (rate of correctly classified samples) for the 19 leave-one subject-out rounds. To present a complete picture, for models based on the four devices, we also present the confusion matrix, and the f-score, which is the harmonic mean of precision (true positives/ (true positives + false positives)) and recall (true positives/ (true positives + false negatives)).

Results

For the five classification models (logistic regression, gradient boosting, KNN, SVM and CNN), accuracy is reported for each combination of devices and for each task in the hierarchical system in Table 1. Overall classification accuracy was highest for logistic regression (based on data of all four devices) with 97.30% accuracy, followed by gradient boosting (all devices) with 96.94%, SVM (all devices) with 96.84%, CNN (three devices at ankle, chest and wrist) with 95.42%, and KNN (three devices at ankle, chest and wrist) with 91.82%. However, when looking at sub-tasks in the hierarchical classification system, gradient boosting is very slightly better than logistic regression in the base-level classification (99.23% vs. 99.21%). GB outperformed logistic regression also in distinguishing between standing and washing dishes (97.40% vs. 97.06%) and between walking, ascending and descending stairs (99.08% vs. 98.72%). When we combined the best classifiers for all sub-tasks, overall mean accuracy rose by 0.04%. As this improvement remains very marginal, we refer to the system based exclusively on logistic regression as the best algorithm. The confusion matrix for the final classification with logistic regression is shown in Table 2.

Discussion

The proposed hierarchical system based on logistic regression classifiers, with 97.30% mean accuracy, represents the best option among those examined here. Compared with previous algorithms tested on the DaLiAc data set (summarized in Table 3), the proposed algorithm represents a threefold improvement. First, our algorithm relies on logistic regression, one of the simplest and most robust tools in machine learning. Unlike other algorithms, the simplicity of our model permitted to reach a very high accuracy, without preliminary hyper-parameter optimization and feature selection. In fact, hyper-parameter optimization of classifiers and feature selection can be a daunting, time-consuming task, and was shown to lead to over-fitting and poor generalization [24]. Second, despite its simplicity, the proposed algorithm performs better than major works tested against the DaLiAc dataset (97.30% versus 96.40% using CNN) (see Table 3). Third, the training of the models themselves is significantly shorter with logistic regression compared to other popular learning algorithms. Using Google Colab (with GPU accelerator) and the parameters mentioned above, training and predicting data following the leave-one-out procedure (i.e. 19 times) lasted 4.5 minutes with logistic regression and KNN, 7.2 minutes for SVM, 10.7 minutes for gradient boosting, and over an hour for CNN. The entire preprocessing phase for the 19 subjects (over 6 hours of observations in total) took only 1.2 minutes.

HAR classification algorithms involve many steps and authors do not always specify all the decisions that they made during data processing before reaching the results. Consequently, it is difficult to fully explain how our algorithm outperformed previous algorithms using classical machine learning classifiers (Table 3) by nearly 3.9%.

We undertook a few steps to identify the innovations that made our algorithm more accurate. First, running our algorithm with a flat classification system instead of the hierarchical system proposed here resulted in 1.81% decrease in mean accuracy. Second, with extracting feature performed on the acceleration signal only, without including the AC and DC components as we did, the difference in accuracy rose to 2.63%. The additional 1.27% difference with the two best-performing algorithms using classical methods by Chen [25] and by Zdravevsky [15] can be attributed to a good trade-off between the number of features and their informativeness. In fact, the former study omitted very important features (no frequency domain features were extracted), while the latter may have had too many of them (4871 before selection). For a better transparency and comparability with other works, the Python script of our model is attached in the Annex.

Despite this promising improvement, a few caveats need to be highlighted. First, very large epidemiological studies interested in physical activity (e.g. the NHANES in the USA) equip their study subjects with accelerometers that are not coupled with gyroscopes [26]. Algorithms evaluated against the DaLiAc data, however, draw upon both. In our algorithm, leaving out gyroscope data resulted in a decrease of 0.5% in mean accuracy. A second issue is related to the nature of our data. HAR algorithms are tested against clean data of activities performed in a characteristic manner as part of a relatively structured protocol. Realistic data, however, can contain less characteristic activities (e.g. slouching, walking a single step...), which represent a greater challenge to classify. In addition, people in real conditions tend to switch rapidly between activities. Consequently, windows of five seconds are probably too long to capture a single activity. A possible solution would be to view sets of activities that are performed together (e.g.

standing and walking around) as activities *per se*. Another solution is to consider smaller windows, for instance of one second. Smaller windows are known to be less good when aiming to capture cyclical activities [21] and can result in a decrease in total accuracy and longer training. In fact, running our algorithm on 1-second windows resulted in a drop of 2.9% and lasted almost 5 times as long as with the 5-second windows commonly used. Limiting this loss in accuracy by applying dynamic windowing methods [21], [27] is an interesting direction for future development. To that extent, very recent attempts to create benchmark activity datasets simulating real conditions [28] are an important development in the field and new algorithms should preferably be assessed using these data.

The high accuracy reached with our algorithm does not mean that logistic regression is best classifier for the task at hand. A better choice of the hyper-parameters of the powerful SVM, GB or CNN models could have resulted in better results. Our point is to emphasize that a simple approach based on domain knowledge can results in a fast, robust and performant model, and that issues of generalisability and tedious processes of model selection must be acknowledged in the evaluation of a new algorithm.

Last, equipping study subjects in free-leaving conditions with four accelerometers as in the DaLiAc dataset can be costly and cause physical and social discomfort to the subjects. In this study, comparing accuracy with different combinations of devices yielded important insights into this matter. Our analysis shows that classification accuracy using only two devices at the ankle and the wrist was practically as good as with the four devices (99.08% accuracy in the basic classification and 96.81% overall accuracy).

Conclusion

In this paper, we propose a novel algorithm for HAR from motion signals (accelerometers and gyroscopes), which significantly improves upon previous work in terms of computational expenses, inferential robustness and classification accuracy. Using a hierarchical classification system with logistic regression, and a relatively small set of features extracted not only from the acceleration signal, but also from low-pass filtered and high-pass filtered signals, proved highly useful in solving the classification task at hand. From a practical perspective, we showed that two devices placed at the wrist and the ankles result in an accuracy that is practically as good as with two additional accelerometers on the chest and the hip, and that using the method proposed here, the additional information brought by the gyroscope was marginal.

Future research should focus on data that better simulate real life conditions, with their swift transitions between activities and less characteristic behaviours. New, simple models should be developed to better adapt to these conditions, while relying, as much as possible, on domain knowledge.

Tables

Table 1: Overview of previous algorithms tested against the DaLiAc dataset using a leave-one-subject-out validation procedure

Authors	Year	Classifiers	Mean accuracy score (%)	Remark
Leutheuser et al. [17]	2013	SVM, AdaBoost, KNN, SVM	89.6	Reference paper
Chen et al. [25]	2016	SVM	93.43	
Nazabal et al. [29]	2016	HMM	95.8	Merged the two bicycle activities
Zdravevski et al. [15]	2017	SVM	93.4	
Hur et al. [12]	2018	CNN	96.4	
Jurca et al. [30]	2018	LSTM	87.16	
Huynh-The et al. [13]	2019	CNN	95.7	
Proposed algorithm	**2020**	**Logistic Regression**	**97.30%**	

*Table 2: **Mean and standard deviation of** accuracy (for 19 leave-one-subject-out rounds) of the different classifiers involved in the hierarchical system by model tested. Results in bold indicate best combination of devices (highest mean accuracy and lowest standard deviation).*

MODEL I : SUPPORT VECTOR MACHINE												
	base		**stand**		**clean**		**walk**		**bike**		**overall**	
	mean	**sd**	**mean**	**sd**	**mean**	**sd**	**mean**	**sd**	**mean**	**sd**	**mean**	**sd**
Device location												
ankle	0.9847	0.0127	0.8449	0.0843	0.7692	0.1165	**0.9872**	**0.0076**	0.9245	0.0600	0.9240	0.0274
chest	0.9711	0.0213	0.9275	0.0769	0.8331	0.1145	0.9640	0.0252	0.9092	**0.0577**	0.9339	0.0330
hip	0.9743	0.0339	0.9343	0.0735	0.7890	0.1045	0.9469	0.0571	0.8422	0.1334	0.9098	0.0525
wrist	0.9441	0.0341	0.8715	0.1045	0.8570	0.0678	0.9015	0.0659	0.8271	0.0955	0.8895	0.0467
ankle|chest	0.9883	0.0125	0.9575	0.0550	0.8928	0.0995	0.9867	0.0139	0.9423	0.0579	0.9617	0.0225
ankle|hip	0.9898	**0.0076**	0.9389	0.0792	0.8337	0.0941	0.9841	0.0165	0.9331	0.0675	0.9488	0.0268
ankle|wrist	0.9893	0.0087	0.9578	0.0499	0.9155	0.0650	0.9797	0.0117	0.9191	0.0665	0.9582	0.0184
chest|hip	0.9843	0.0127	0.9576	0.0535	0.8668	0.0882	0.9700	0.0247	0.9224	0.0817	0.9486	0.0267
chest|wrist	0.9799	0.0180	0.9607	0.0550	0.9213	0.0577	0.9532	0.0372	0.9111	0.0667	0.9476	0.0296
hip|wrist	0.9846	0.0116	0.9595	0.0413	0.9059	**0.0521**	0.9502	0.0367	0.8876	0.0835	0.9426	0.0227
ankle|chest|hip	0.9908	0.0103	0.9671	0.0484	0.8981	0.0925	0.9830	0.0193	**0.9495**	0.0625	0.9648	0.0227
ankle|chest|wrist	0.9889	0.0118	0.9644	0.0579	0.9391	0.0668	0.9792	0.0160	0.9376	0.0761	0.9650	0.0227
ankle|hip|wrist	0.9911	0.0091	0.9575	**0.0365**	0.9268	0.0616	0.9778	0.0165	0.9338	0.0623	0.9623	**0.0166**
chest|hip|wrist	0.9861	0.0124	0.9694	0.0499	0.9248	0.0576	0.9617	0.0296	0.9273	0.0771	0.9567	0.0249
ankle|chest|hip|wrist	**0.9911**	0.0106	**0.9716**	0.0496	**0.9397**	0.0671	0.9790	0.0180	0.9456	0.0713	**0.9684**	0.0219
Best combination	*0.9911*	*0.0076*	*0.9716*	*0.0365*	*0.9397*	*0.0521*	*0.9872*	*0.0076*	*0.9495*	*0.0577*	*0.9684*	*0.0166*
MODEL II : CONVOLUTION NEURAL NETWORK												
	base		**stand**		**clean**		**walk**		**bike**		**overall**	
	mean	**sd**	**mean**	**sd**	**mean**	**sd**	**mean**	**sd**	**mean**	**sd**	**mean**	**sd**
ankle	0.9876	0.0101	0.8949	**0.0498**	0.8504	0.1174	0.9764	0.0187	0.8919	0.0703	0.9334	0.0291
chest	0.9714	0.0308	0.8752	0.1679	0.8484	0.1206	0.9611	0.0390	0.9085	0.0554	0.9258	0.0400
hip	0.9201	0.1040	0.7603	0.1307	0.7199	0.1658	0.8944	0.0900	0.8325	0.1090	0.8409	0.0923

wrist	0.9395	0.0504	0.8452	0.1503	0.8583	0.1071	0.8518	0.1114	0.8538	0.1061	0.8718	0.0632
ankle|chest	0.9872	0.0099	0.9259	0.1390	0.9162	0.0777	0.9773	0.0298	0.9141	0.0616	0.9532	0.0220
ankle|hip	0.9846	0.0150	0.8870	0.1064	0.8497	0.1104	0.9695	0.0374	0.9182	0.0680	0.9365	0.0290
ankle|wrist	0.9880	0.0211	**0.9650**	0.0640	0.9214	0.0757	0.9684	0.0632	0.9074	0.0660	0.9541	0.0283
chest|hip	0.9722	0.0212	0.8657	0.1782	0.8484	0.0994	0.9576	0.0477	0.9088	0.0657	0.9227	0.0313
chest|wrist	0.9817	0.0152	0.9079	0.1331	0.9210	0.0775	0.9526	0.0548	0.9002	0.0711	0.9401	0.0358
hip|wrist	0.9531	0.0595	0.8566	0.1650	0.8887	0.0858	0.9471	0.0507	0.8679	0.0793	0.9082	0.0518
ankle|chest|hip	0.9793	0.0174	0.8635	0.1837	0.8642	0.0977	**0.9799**	0.0220	0.9170	**0.0577**	0.9375	0.0370
ankle|chest|wrist	**0.9896**	**0.0093**	0.9425	0.0811	**0.9364**	**0.0607**	0.9668	0.0703	0.9070	0.0819	**0.9542**	**0.0220**
ankle|hip|wrist	0.9797	0.0258	0.9369	0.0796	0.8838	0.1199	0.9733	0.0479	**0.9259**	0.0811	0.9479	0.0312
chest|hip|wrist	0.9831	0.0156	0.9251	0.0739	0.9213	0.0632	0.9435	0.0832	0.8879	0.0704	0.9372	0.0310
ankle|chest|hip|wrist	0.9787	0.0282	0.9148	0.1024	0.9080	0.1023	0.9771	**0.0168**	0.9137	0.0885	0.9465	0.0313
Best Combination	*0.9896*	*0.0093*	*0.9650*	*0.0498*	*0.9364*	*0.0607*	*0.9799*	*0.0168*	*0.9259*	*0.0577*	*0.9542*	*0.0220*

MODEL III : K NEAREST NEIGHBORS												
	base		**stand**		**clean**		**walk**		**bike**		**overall**	
	mean	**sd**	**mean**	**sd**	**mean**	**sd**	**mean**	**sd**	**mean**	**sd**	**mean**	**sd**
Device location												
ankle	0.9726	0.0158	0.6211	0.0920	0.6782	0.1218	0.9761	0.0157	0.7480	0.1036	0.8343	0.0344
chest	0.9344	0.0357	0.8697	0.0932	0.6988	0.1163	0.9561	0.0279	0.6792	0.1264	0.8534	0.0493
hip	0.9455	0.0402	0.8257	0.1262	0.6878	0.1169	0.9221	0.0481	0.6728	0.1243	0.8344	0.0553
wrist	0.8940	0.0536	0.8510	0.1139	0.6572	0.1581	0.8123	0.0909	0.6180	0.1269	0.7842	0.0712
ankle|chest	0.9828	0.0160	0.9196	0.0766	0.7981	0.0876	**0.9873**	**0.0085**	**0.8042**	0.0851	0.9150	0.0255
ankle|hip	0.9802	0.0138	0.8077	0.1098	0.7298	0.1155	0.9763	0.0163	0.7174	0.0953	0.8683	0.0320
ankle|wrist	0.9808	**0.0128**	0.9164	0.0748	0.7861	0.0882	0.9737	0.0213	0.7478	0.1303	0.8981	0.0327
chest|hip	0.9588	0.0333	0.9127	0.0768	0.7542	0.1117	0.9614	0.0301	0.7268	0.1248	0.8813	0.0516
chest|wrist	0.9592	0.0244	0.9205	0.0759	0.8281	0.0940	0.9280	0.0435	0.7574	0.0979	0.8920	0.0326
hip|wrist	0.9511	0.0358	0.9014	0.0797	0.7805	0.1403	0.9236	0.0561	0.6343	0.1482	0.8562	0.0548

ankle|chest|hip	0.9803	0.0172	0.8990	0.0875	0.8048	0.1024	0.9779	0.0169	0.7801	**0.0754**	0.9047	0.0296
ankle|chest|wrist	**0.9840**	0.0137	0.9280	0.0876	**0.8642**	**0.0633**	0.9799	0.0173	0.7787	0.1136	**0.9182**	**0.0233**
ankle|hip|wrist	0.9828	0.0132	0.9183	0.0792	0.8392	0.0850	0.9726	0.0196	0.7454	0.1140	0.9056	0.0301
chest|hip|wrist	0.9659	0.0287	0.9359	0.0777	0.8246	0.0983	0.9443	0.0402	0.7396	0.1212	0.8952	0.0428
ankle|chest|hip|wrist	0.9839	0.0148	**0.9336**	**0.0742**	0.8566	0.0792	0.9775	0.0140	0.7807	0.0973	0.9179	0.0249
Best combination	*0.9840*	*0.0128*	*0.9336*	*0.0742*	*0.8642*	*0.0633*	*0.9873*	*0.0085*	*0.8042*	*0.0754*	*0.9182*	*0.0233*
MODEL IV : GRADIENT BOOSTING												
	base		**stand**		**clean**		**walk**		**bike**		**overall**	
	mean	**sd**	**mean**	**sd**	**mean**	**sd**	**mean**	**sd**	**mean**	**sd**	**mean**	**sd**
ankle	0.9650	0.0373	0.7377	0.1573	0.7629	0.1254	0.9778	0.0179	0.9101	0.0851	0.8927	0.0431
chest	0.9729	0.0246	0.9249	0.0974	0.8497	0.1243	0.9779	0.0198	0.8990	0.0773	0.9364	0.0344
hip	0.9666	0.0505	0.9212	0.0936	0.7744	0.1431	0.9523	0.0706	0.8431	0.1827	0.9076	0.0646
wrist	0.9331	0.0519	0.8641	0.0905	0.8375	0.0822	0.9016	0.0788	0.7928	0.1126	0.8718	0.0624
ankle|chest	0.9832	0.0243	0.9483	0.0836	0.8812	0.0938	0.9897	0.0078	0.9278	0.0805	0.9537	0.0334
ankle|hip	0.9901	0.0071	0.9185	0.1102	0.8432	0.0886	0.9847	0.0146	0.9143	0.0765	0.9442	0.0305
ankle|wrist	0.9779	0.0274	0.9330	0.0956	0.8612	0.0976	0.9820	0.0183	0.9263	0.0701	0.9451	0.0323
chest|hip	0.9774	0.0369	0.9608	0.0542	0.8546	0.1131	0.9673	0.0753	0.9229	0.0622	0.9444	0.0448
chest|wrist	0.9816	0.0219	0.9494	0.0749	0.9084	0.0780	0.9773	0.0174	0.9056	0.0679	0.9518	0.0272
hip|wrist	0.9774	0.0235	0.9485	0.0757	0.8934	0.0653	0.9745	0.0177	0.8608	0.1066	0.9359	0.0309
ankle|chest|hip	**0.9923**	0.0073	0.9712	0.0381	0.8846	0.0912	**0.9908**	**0.0063**	0.9329	0.0753	0.9639	0.0231
ankle|chest|wrist	0.9846	0.0191	0.9657	0.0626	0.9210	0.0698	0.9893	0.0090	**0.9408**	0.0584	0.9647	0.0242
ankle|hip|wrist	0.9913	**0.0057**	0.9691	0.0373	0.9213	**0.0487**	0.9867	0.0125	0.9134	0.0807	0.9626	0.0184
chest|hip|wrist	0.9846	0.0161	0.9708	0.0525	0.9096	0.0745	0.9818	0.0117	0.9258	**0.0546**	0.9595	0.0196
ankle|chest|hip|wrist	0.9922	0.0064	**0.9740**	**0.0313**	**0.9292**	0.0660	0.9891	0.0093	0.9336	0.0662	**0.9694**	**0.0188**
Best combination	*0.9923*	*0.0057*	*0.9740*	*0.0313*	*0.9292*	*0.0487*	*0.9908*	*0.0063*	*0.9408*	*0.0546*	*0.9694*	*0.0188*
MODEL V : LOGISTIC REGRESSION												
	base		**stand**		**clean**		**walk**		**bike**		**overall**	
	mean	**sd**	**mean**	**sd**	**mean**	**sd**	**mean**	**sd**	**mean**	**sd**	**mean**	**sd**

ankle	0.9828	0.0130	0.7977	0.0951	0.7939	0.0992	0.9814	**0.0099**	0.9333	0.0532	0.9199	0.0261
chest	0.9708	0.0232	0.9415	0.0693	0.8224	0.0933	0.9633	0.0397	0.8704	0.0670	0.9258	0.0288
hip	0.9714	0.0269	0.9257	0.0566	0.7393	0.0931	0.9548	0.0424	0.8048	0.1286	0.8939	0.0438
wrist	0.9243	0.0421	0.8547	0.1250	0.8289	0.1035	0.8865	0.0626	0.8155	0.0738	0.8673	0.0489
ankle|chest	0.9875	0.0130	0.9575	0.0566	0.8825	0.0740	**0.9872**	0.0209	0.9348	0.0716	0.9586	0.0228
ankle|hip	0.9886	0.0104	0.9298	0.0608	0.8048	0.1022	0.9848	0.0182	0.9305	0.0707	0.9435	0.0265
ankle|wrist	0.9908	**0.0069**	0.9631	0.0415	0.9212	0.0522	0.9803	0.0203	0.9528	**0.0493**	0.9681	**0.0135**
chest|hip	0.9805	0.0172	0.9428	0.0660	0.8403	0.0861	0.9802	0.0192	0.9090	0.1002	0.9431	0.0310
chest|wrist	0.9831	0.0141	0.9643	0.0516	0.9336	0.0487	0.9630	0.0339	0.9135	0.0507	0.9546	0.0185
hip|wrist	0.9841	0.0179	0.9653	0.0305	0.9038	0.0592	0.9625	0.0274	0.8885	0.0665	0.9449	0.0295
ankle|chest|hip	0.9888	0.0127	0.9532	0.0625	0.8775	0.0811	0.9848	0.0158	0.9472	0.0734	0.9606	0.0239
ankle|chest|wrist	0.9913	0.0096	**0.9706**	0.0515	**0.9444**	0.0555	0.9854	0.0143	0.9429	0.0723	0.9706	0.0199
ankle|hip|wrist	**0.9921**	0.0073	0.9685	**0.0354**	0.9267	**0.0453**	0.9828	0.0138	0.9426	0.0505	0.9683	0.0153
chest|hip|wrist	0.9881	0.0136	0.9648	0.0593	0.9330	0.0458	0.9753	0.0240	0.9268	0.0875	0.9626	0.0241
ankle|chest|hip|wrist	0.9916	0.0103	0.9669	0.0582	0.9421	0.0475	0.9845	0.0159	**0.9547**	0.0719	**0.9730**	0.0198
Best combination	*0.9921*	*0.0069*	*0.9706*	*0.0354*	*0.9444*	*0.0453*	*0.9872*	*0.0099*	*0.9547*	*0.0493*	*0.9730*	*0.0135*

*Table 3: **Confusion matrix** for the aggregated confusion matrices calculated for all leave-one-subject-out rounds. Class-specific precision, recall and f-score (β=1) are reported for each class.*

		Predicted												
		sit	**lie**	**stand**	**wash**	**vacuum**	**sweep**	**walk**	**Stairs-up**	**Stairs-down**	**run**	**bike 50W**	**bike 100W**	**jump**
Observed	**sit**	**430**	0	17	3	0	0	0	0	0	0	0	0	0
	lie	1	**455**	0	0	0	0	0	0	0	0	0	0	0
	stand	2	0	**442**	8	0	0	1	0	0	0	0	0	0
	wash	0	0	2	**924**	7	4	0	0	0	0	0	0	0
	vacuum	0	0	0	7	**422**	25	0	0	0	0	0	0	0
	sweep	0	0	6	4	23	**704**	4	2	0	0	0	0	0
	walk	0	0	3	1	4	5	**2010**	11	6	1	0	0	0
	stairsup	0	0	0	0	0	1	6	**312**	1	0	0	0	0
	stairsdown	0	0	0	0	0	0	5	2	**266**	0	0	0	0
	run	0	0	0	0	0	0	0	0	0	**910**	1	0	0
	bike 50W	0	0	0	0	0	0	0	0	0	0	**877**	46	0
	bike 100W	0	0	0	0	0	0	0	0	0	0	37	**883**	2
	jump	0	0	0	0	0	0	0	0	0	0	0	0	**243**
	precision	0.99307	1.00000	0.94043	0.97571	0.92544	0.95264	0.99210	0.95413	0.97436	0.99890	0.95847	0.95048	0.99184
	recall	0.95556	0.99781	0.97572	0.98613	0.92952	0.94751	0.98481	0.97500	0.97436	0.99890	0.95016	0.95770	1.00000
	f_score	0.97395	0.99890	0.95775	0.98089	0.92747	0.95007	0.98844	0.96445	0.97436	0.99890	0.95430	0.95408	0.99590

Figures

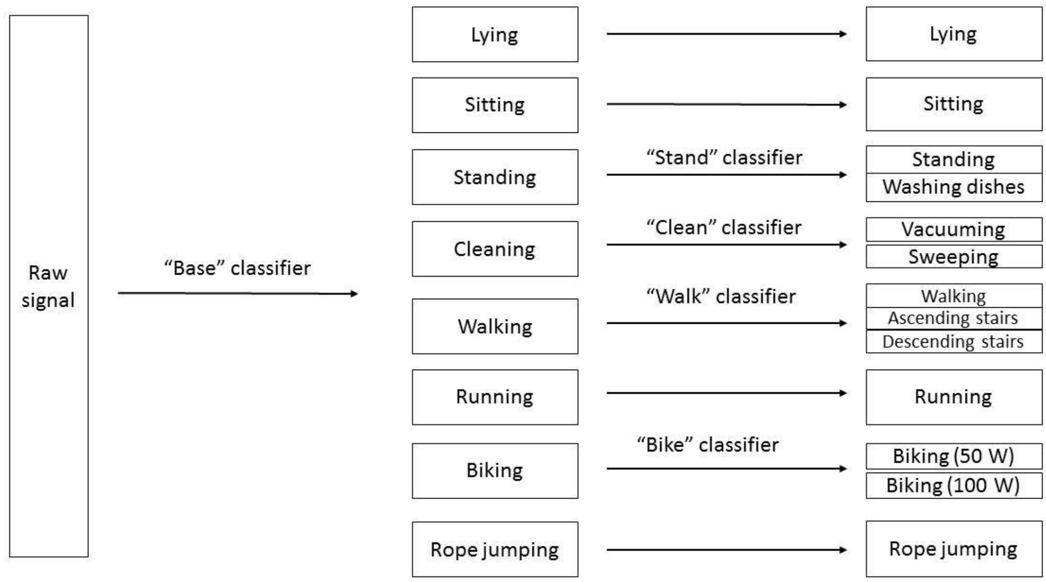

Figure 1: ***Illustration of our hierarchical classification system***

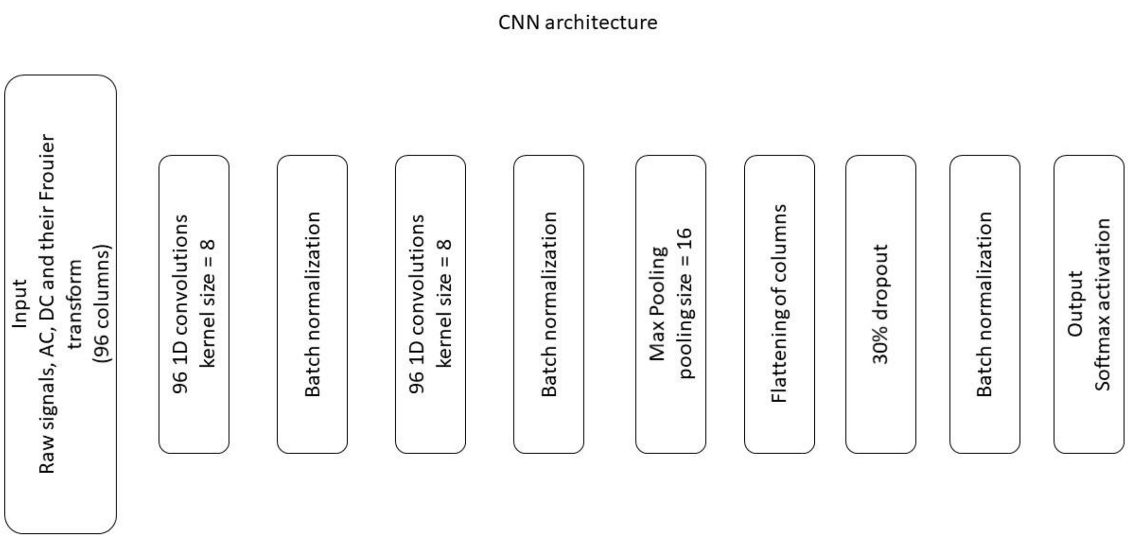

Figure 2: The ***Convolution Neural Network*** *tested here. Except for the output, all layers were activation with the RELU function.*

CHAPTER IV

Effects of attributes of the urban environment on physical activity

Chapitre IV: Effets de l'environnement urbain sur l'activité physique (résumé français)

Dans les chapitres précédents, nous avons abordé la question du lien entre l'activité physique et la santé et de l'utilisation d'accéléromètres déployés sur une population afin de mesurer son niveau d'activité physique en conditions de vie libres. Le caractère pratiquement continu du monitoring par accéléromètre devient particulièrement utile lorsque ces mesures d'activité physique à haute résolution temporelle sont croisées avec des informations sur des stimuli concomitants auxquels les individus sont exposés. Un tel croisement d'information permet d'identifier, en aval, des facteurs déterminants de l'activité physique et, dans la mesure où ces facteurs sont modifiables, d'élaborer des plans d'action visant à encourager l'activité physique et améliorer indirectement l'état de santé dans la population.

L'environnement bâti, surtout dans les villes, comme facteur déterminant de l'activité physique des riverains a beaucoup été étudié en santé publique. Ainsi, le lien entre de nombreux attributs de l'environnement et le niveau d'activité physique a été démontré dans la littérature. Par exemple, de nombreuses études ont montré que la présence d'espaces verts ou une forte densité de destinations (magasins, monuments, lieux de loisirs) d'un certain quartier encourageaient la marche chez les riverains. Malgré de très nombreuses études réalisées au cours des dernières décennies, les progrès technologiques et méthodologiques récents ont ouvert la voie à de nouvelles perspectives, qui seront abordés dans cette étude.

Les études traditionnelles sur le lien entre l'environnement et l'activité physique, par manque de moyens techniques et informatiques, se sont principalement concentrées sur l'étude de l'environnement autour d'un point d'ancrage dans la vie des sujets étudiés, notamment le domicile ou le lieu de travail. Ces études s'intéressaient donc, pour reprendre notre exemple, aux espaces verts autour du domicile et à leur lien avec l'activité physique cumulée sur la journée. Or, les individus pouvant être mobiles tout le long de la journée, ces études ignoraient *où* cette activité physique observée était effectuée et ne pouvaient établir des liens *directs* entre l'environnement et l'activité. De plus le caractère transversal de la plus grande partie des études ne permettait aucune conclusion pertinente sur la causalité du lien observé. Bien au contraire : autant que l'environnement affecte l'intention d'être actif (je me promène parce que la rue est pleine de verdure), l'intention d'être actif influence le choix de l'environnement (je veux me promener donc je vais dans la rue pleine de verdure). Enfin, jusqu'à présent, les études précédentes ont étudié des mesures d'activité très sommaires ; l'influence de l'environnement sur l'allocation posturale a été pratiquement ignorée.

Notre étude utilise les données RECORD présentées au **chapitre II**. En plus des accéléromètres, les individus suivis ont été équipés d'un appareil GPS permettant de les localiser en continu. Ainsi, nous avons pu modéliser l'activité physique en fonction de la position concomitante de l'individu dans l'espace. En croisant les données de localisation avec des sources d'information géographiques provenant des recensements et d'agences de l'aménagement du territoire, nous avons pu obtenir une image claire de l'exposition continue de l'individu à divers attributs de l'environnement urbain et de son activité physique concomitante. Pour remédier au problème de l'inférence causale, nous nous sommes limités à l'analyse de « tranches de vie » spatio-temporelles dans lesquelles l'individu n'a très probablement pas le choix du lieu dans lequel il se trouve, mais il a le choix de l'activité physique qu'il va effectuer. Notamment, dans les trajets domicile-travail, l'individu ne choisit pas l'endroit où il se trouve (on peut dire que ce trajet lui est quasiment imposé) mais il a le choix entre divers moyens de transports, ou une combinaison de ceux-ci, qui vont déterminer la composition de son budget-temps d'activité durant le trajet. A un trajet effectué en voiture correspondra un budget-temps avec une composante sédentaire très élevée (à moins que l'individu combine de la marche avec le déplacement en voiture), alors qu'un trajet effectué à bicyclette ou à pied aura une composante d'activité modérée-à-vigoureuse élevée. Suite aux résultats présentés au **chapitre I**, qui ont mis en évidence les bienfaits du temps passé debout, la posture « debout » est distinguée des postures « assis » ou « couchés ».

Quatre caractéristiques de l'environnement urbain ont été étudiées : le niveau de verdure (arbres, espaces verts), la densité de destinations (magasins, monuments, cafés, restaurants, lieux de divertissement, services publics…), le niveau socio-économique moyen (proportion de la population ayant atteint un niveau d'éducation post-bac), et l'efficacité des transports en commun (ratio du temps de trajet en transports en commun contre le temps en véhicule motorisé privé). Le niveau d'exposition aux attributs d'intérêt a été mesuré tout le long du chemin le plus court reliant le lieu de travail et le domicile, en agrégeant les mesures obtenues (proportion d'espaces vert, densité de destination moyenne, niveau d'éducation moyen) dans une zone tampon d'un rayon de 200 mètres alentour.

Sur les quatre variables environnementales étudiées, nous avons constaté que trois variables influençaient l'activité pendant les trajets de façon significative. Sur un trajet plus « vert », le temps actif était plus important. De même, jusqu'à un certain niveau, le niveau d'éducation moyenne aux alentours du chemin parcouru influençait positivement le volume d'activité physique modérée-à-vigoureuse. Enfin, plus les transports en commun reliaient le lieu de travail et le domicile rapidement (par rapport à un véhicule motorisé privé), plus la composante « debout » et « activité physique modéré à vigoureuse » était importante.

Ces résultats indiquent que des attributs esthétiques de l'environnement urbain (niveau de verdure) ou sociaux (niveau d'éducation) peuvent affecter des choix portant sur le niveau d'activité physique sur des trajets utilitaires. Alors que nos résultats sur l'effet de la verdure peuvent s'expliquer simplement par le fait que les individus favorisent les transports actifs lorsque le trajet est plus agréable, le niveau socio-économique est probablement une variable *proxy* qui en cache d'autres. Des études en sciences sociales ont montré que les attributs de l'environnement urbains bénéfiques à la santé étaient mal répartis dans l'espace – plutôt dans les quartiers riches que les quartiers pauvres – créant ainsi une *injustice environnementale* envers les classes défavorisées. Enfin, une bonne infrastructure des transports en commun peut également augmenter l'activité physique et réduire le comportement sédentaire. Le **chapitre V** de cette thèse examine plus avant la répartition des activités physiques en fonction du mode de transports choisi.

Malgré son cadre limité, cette étude illustre de façon claire comment, à l'ère des *big data,* des sources de données complexes peuvent être exploitées pour ouvrir de nouvelles perspectives. Dans cette étude, nous avons combiné des données de géolocalisation, d'accélérométrie et des systèmes d'informations géographiques afin de surmonter les limites d'inférence statistique existantes et mettre en évidence l'effet immédiat de l'environnement sur le comportement physique précis de ceux qui y sont exposés.

Effects of built and social environmental factors on distribution of physical activity and body postures during daily commutes

Isaac Debache[1]*, Cédric Sueur[1], Julie Vallée[2], Audrey Bergouignan[1,3], Basile Chaix [4].

[1] Institut Pluridisciplinaire Hubert Curien (IPHC) UMR 7178 Centre National de la Recherche Scientifique (CNRS), Université de Strasbourg 67000 Strasbourg, France.

[2] UMR 8504 Centre National de la Recherche Scientifique (CNRS) 75005 Paris, France.

[3] Division of Endocrinology, Metabolism, and Diabetes and Anschutz Health and Wellness Center, University of Colorado, School of Medicine, Aurora, Colorado, USA

[4] Institut Pierre Louis d'Epidémiologie et de Santé Publique (IPLESP), Institut National de la Santé et de la Recherche Médicale (INSERM), Nemesis Team, 75012 Paris, France.

*Corresponding author: isaac.debache@iphc.cnrs.fr; Tel.: +33 388106931

Abstract

Assessing the effects of social and built environments on physical activity is important for promoting healthy life style in cities. Yet, very few studies use objective localization and measures of physical activity while considering causality. In addition, the effect of environment on body postures, albeit of physiological importance, is rarely addressed. Using mixed models for compositional data on sensor-derived data, we estimated the effects of greenery, destination density, neighborhood average educational level and public transports efficiency along 692 home-work journeys made by 121 healthy adult patients (80 men, 41 women). Higher levels of greenery, average education and public transports time efficiency in the areas crossed during commutes were found to reduce contemporaneous sedentary behaviors and increase physical activity. These causal, observed relationships suggest that deciders should consider greening, as well as increasing environmental justice and public transports efficiency, as an effective way to fight the pandemic of sedentary behaviors.

Keywords

Built environment, urban environment, environmental factors, greenery, destination density, public transport, socio-economic status, education, GPS, accelerometer.

Introduction

Because of its well-known effects on reducing health hazards, such as cardio-vascular diseases, diabetes, cancer and depression, physical activity (PA) is widely promoted by public health policies [1, 2]. Despite these efforts, the prevalence of physical inactivity has not declined over the last decade, while the time spent in sedentary behaviors (SB, i.e. sitting or reclining with an energy expenditure below 1.5 METs [3]) increased [4]. As a potential field of intervention, research has studied various features of social and physical environment and their link to PA. To date, there is some evidence to the positive effect of environmental features, such as the presence of green spaces and paths, population and destination density, access to transit and walkability on both leisure PA and active transports [5–7]. The goal of the present study is to bring novel evidence to the relationship between four features of the physical environments (greenery level, destination density, average neighborhood education and access of the public transportation network) and PA. The literature addressing these issues is abundant, however, the present study innovates by proposing a new study design a) using objective, contemporaneous measures of environmental features and physical activity, b) aiming at causal inference c) and distinguishing between objective measures PA and SB, i.e. between body motion and body posture.

A contemporaneous design

The largest part of the evidence as to the effect of environment on PA was obtained by linking environmental attributes of a fixed location, such as the subjects' home or workplace, and PA levels aggregated over a certain period (e.g. a day). However, newer research has argued that, individuals being mobile over the day, their PA at a certain moment should be regarded as a function of their contemporaneous exposure to environment [8–10]. Following recent studies, we implemented this design (hereafter 'contemporaneous design') by using environmental features of subjects' GPS-derived locations and their simultaneous accelerometer-derived PA performed at the same time

[11, 12]. With this method, the simultaneousness of the exposure and the PA yields accurate, valid estimates of the associations between the two.

Selection biases and causal inference

Studies on the association between environment and PA are subject to two well-identified methodological biases. The residential self-selection bias refers to the preference of individuals inclined to engage in PA for neighborhoods favorable to PA. This self-selection makes it difficult to determine the causal direction of any observed relationship between residential environment and PA [5]. The instantaneous selective mobility bias refers to the fact that as much as locations can affect the intention to perform PA, the intention to perform PA can affect the choice of locations to which one goes [13–16]. Despite the growing number of studies with a contemporaneous design, the selective mobility bias has rarely been addressed in previous studies.

To address the residential self-selection bias, we took into account subjects' reported motivation in choosing their residential neighborhood. Furthermore, to address the selective mobility bias, we have limited our frame of analysis to the sole utilitarian journeys between home and workplace. As individuals arguably do not have any latitude in the choice of their work location over the study period (given that their workplace is likely determined before the 7-day observation period), these home-work journeys can be regarded as free of instantaneous selective mobility bias. Thus, the present study proposes a research strategy that acknowledges the advantage of a contemporaneous design while minimizing biases that compromise causal inference. In addition, since home-work journeys are a necessary activity, making up a considerable share of the wake time in the active population, they represent an interesting opportunity to achieve the daily PA recommended by physicians.

Physical activity and body postures

Past studies using objective measurements focused only on PA (typically walking or cycling), but not on SB (reclining/sitting postures). Yet, with the new findings regarding the deleterious effects of prolonged SB [17–20], the focus of research has shifted towards a refined definition of physical activity, including not only the body motion, but its postures [21, 22]. In the present study, we investigate the effect on environment on both PA and SB, by using different categories of physical behaviors, such as sitting/reclining, standing, or walking.

Environmental outcome studied

In a literature review, Kondo [23] identified 12 articles using objective localization and PA that found a positive association between exposure to greenery and contemporaneous levels of moderate-to-vigorous PA. Yet, as we pointed out, the validity of such studies is compromised by selective mobility bias [10]. Notwithstanding this limitation, we rely on these findings to posit that this observed human preference for green environments when engaging in PA should result in more active commuting along routes featuring greenery.

Likewise, accessibility to destinations required for daily living, such as shops, is as a factor promoting healthy behaviors, including walking for transports [24]. Yet, among extant evidence [25–28], causal evidence to contemporaneous effects is still missing. Here, we hypothesized that high destination density represents a rich, stimulating environment, which encourages active commuting [10].

It has been alleged that the spatial distributions of health-related environmental features was unequal across neighborhoods of different socio-demographics [29]. Among the few addressing the neighborhood socio-economic status (SES) as a potential determinant of PA [30], a study by Riva et al. [31] showed an increase in utilitarian walking in high education neighborhood, independent of individual SES. Here, we rely on the theory of environmental justice to posit that work-home routes crossing high SES neighborhoods, as captured by the residents' average areal educational attainment, would feature more PA and less SB.

All in all, better access to public transit system is thought to have a positive impact on PA levels, as travelling public transports typically implies more walking episodes than travelling with a private car [32–34]. Thus, we hypothesized that greater incentives of commuting with public transports, such as better access and higher time efficiency, would also be associated with more PA and less SB over the journey. SB is of particular interest here, as public transports often require commuters to be standing.

There is no doubt that environment affects PA through the choice of transportation mode. For our population, the associations between environment and transportation mode, as well as between transportation mode and physical activity, will investigated in a separate study. The goal is the present study is to assess and quantify the direct effect of environmental features on objective physical activity, derived from accelerometers worn by participants. Thus, we will address the question of transportation mode as linking mechanism between only incidentally.

Material and Methods

Population

The present study uses data collected in the RECORD (Residential Environment and COronary heart Disease) study, which investigated spatial disparities in health. From February 2007 to March 2008, individuals that came to one of the four IPC (Investigation Préventive et Clinique) Medical Centers for a free medical examination offered by the French National Insurance System for Employees were invited to enter the RECORD study. Eligibility criteria were age 30-79 years and residence in ten given districts of Paris (out of 20) or in 111 other municipalities of the Ile-de-France region, as well as sufficient cognitive and linguistic abilities to comply with the instructions. During the second wave of the study (between September 2013 and June 2015), former and newly recruited participants underwent a medical examination, after which they were invited to enter the RECORD MultiSensor ancillary study whenever sensors were available. In this study, they were asked to wear body-mounted sensors, including a pair of accelerometers and a GPS tracker. Participants in this ancillary study were instructed to wear the sensors for 7 consecutive days, as they carried out their usual activities in free-living conditions, and to keep a logbook with the places that they were visiting and the transportation modes used in journeys between those places. In addition, they were requested to answer a questionnaire regarding their health and dietary habits, neighborhood, demographics and SES. The study protocol was approved by the French Data Protection Authority (Decision No. DR-2013-568 on 2/12/2013).

Physical behaviors during travel

Direct measures of PA and SB were derived from the data of two tri-axial Vitamove Research-V1000® devices, worn at the right upper leg and on the chest during wake time (except for water-based activities). We regrouped the behavioral categories provided by the software into three broader categories: sedentary behaviors (lying or sitting, SB), standing still and light movements (ST) and physical activity (walking, running and bicycling, PA).

Journeys

The environmental attributes of the journey were calculated along the shortest route (walking) between subjects' homes and workplaces, which was determined using GoogleMaps. The shortest route was preferred over the actual route taken by participants, as any observed correlation between the features of the latter and PA is likely to suffer from selective mobility bias [23]. In fact, the actual route taken by individuals is also a function of their preferred commute mode, and therefore endogenous to our research question. In contrast, assuming that individuals generally prefer the shortest route for their commute, and that this preference is at least partly independent of the individual's *a priori* preferences for commute modes, our hypotheses as to the effects of environment on PA levels could be verified along the shortest route.

Journeys that included any stop (for example stopping at shops or at friends' house) were not included in the analysis. Likewise, we removed journeys that included segments outside of Ile-de-France (region of Paris). The final dataset included 692 journeys recorded for 121 participants.

Measures of environmental attributes

Greenery index

We used two methods to measure greenery level. First, we sampled a set of equidistant points (each 200 meters) along the shortest routes and calculated the mean of the shortest distances between the points and the network of green paths. This variable represents the opportunity cost, in terms of distance covered, of using the green network. Second, we calculated the proportion of green spaces for a buffer zone of 100m radius around the shortest routes, representing a proxy for the direct proximity to greenery during the journey, as illustrated in Figure 1. Greenery level best predicts health when measured in a radius of 1-2 km around individuals' homes [36], but we argue that the area around the route that may affect the traveler's behavior is much smaller. We calculated these measures using the 2008 open data of the Ile-de-France Institute for Urbanism [37].

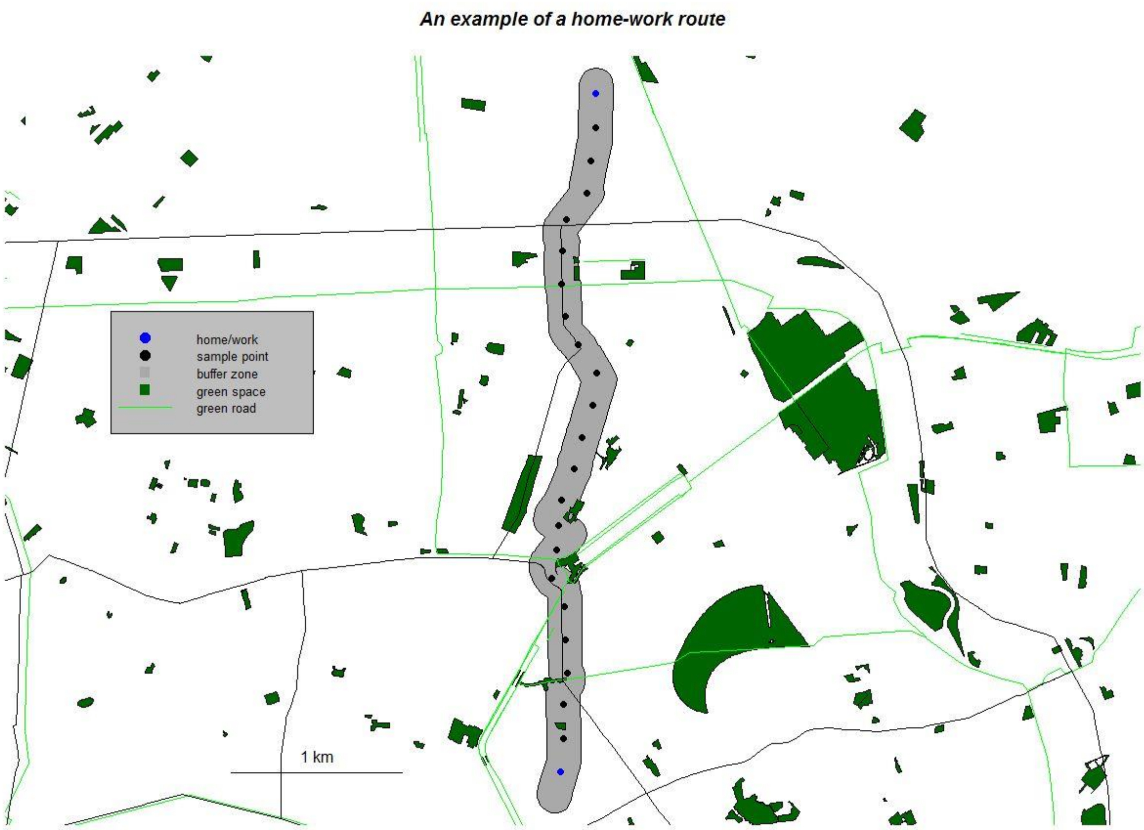

*Figure 13 : **Examples of the shortest route (walking) of a work-home journey** in the North of Paris and the neighboring Aubervilliers. The green ratio is ratio of the green areas intersecting with the buffer along the route to the total buffer area. The distance to the network of green roads is the average of the distances between the points sample along the route.*

Destination density

Destination density was calculated as the number of destinations per km2, including public services, shops, entertainment facilities etc., in a buffer zone of 100m radius around the shortest routes (French National Institute for Statistics and Economic Studies INSEE, 2011 [38]).

Education

To estimate the SES of the areas crossed by the participants, we created buffers of 100m radius around the shortest route for the journey. In each buffer, the educational attainment of the population was defined as the share of people aged 20 and over holding a university degree (INSEE 2010 census [39]).

Accessibility to and time cost of public transportation

Two methods were used to estimate the incentive to use public transports. For the residence and workplace, the Euclidian distance to the nearest station belonging to the bus, metro, tramway or railway networks (Île-de-France Mobilités, 2012 [40]) was calculated, and the larger distance of the two was defined as the accessibility variable. The time cost variable was defined as the ratio of the travel time using public transports to the travel time using a car, as estimated by GoogleMaps.

Explanatory covariables

A number of potentially confounding factors known to influence physical activity have been measured. At the journey level, the length (in km) of the shortest route was taken into account as it probably plays a role in the choice of the travel mode. At the individual level, we considered sex, age, being in couple, having children under 14 years at home, household income (by tertiles) and individual educational attainment (No higher education, undergraduate, graduate). To control for neighborhood selection bias, we accounted for the answers given by the participants in the questionnaires regarding neighborhood selection [15]. Participants were asked to score, on a scale varying from 'not at all' to 'very much' (coded 0-3), how important the greenery level, presence of shopping facilities, SES and accessibility of public transports were in the choice of neighborhood to which they moved.

Statistical analysis

The goal of the analysis was to estimate the effect of exposure to environmental route attributes on the time budget of physical behaviour during the journey. Thus, the physical behaviors were regarded as compositional data adding up to 1, where each of the three parts corresponds to the share of the time spent in sedentary postures, standing, and PA. Compositions add up to a constant sum and are hence interdependent and constrained between 0 and 1. To model compositions of physical behaviors ($\boldsymbol{y} = [y_{SB}, y_{ST}, y_{PA}]$) as dependent variables explained by environmental measures, we applied the additive log ratio (alr) transformation to the data [41], taking sedentary time as reference :

$$\boldsymbol{y}' = [0, \ln\frac{y_{ST}}{y_{SB}}, \ln\frac{y_{PA}}{y_{SB}}].$$

With this linear transformation, the log-ratios can take any real value and be modelled using usual tools of statistical analysis, while preserving their relative nature.

When a part of the composition was zero, we added an epsilon (0.01) and divided all the parts by the sum, to avoid infinite values when applying the alr transformation. The two log-ratios $\widehat{y'_{st}}$, $\widehat{y'_{PA}}$ were modelled as linear functions of the environmental and social factors and co-variables, using mixed linear regression with patients as random intercepts. The predicted log-ratio vector $\widehat{\boldsymbol{y}'}$ could be then back-transformed to predict a composition $\widehat{\boldsymbol{y}}$ for any set of values of environmental features.

The environmental factors (greenery ratio, distance to greenery, destination density, average areal education, time cost of public transports and distance to the nearest station) were included with the following co-variables: age, sex, being in couple, having children at home, individual educational level, income level, length of route, squared length, importance of variable of interest in neighborhood selection. A squared term was added for route length as we assumed that it has a non-linear effect on the probability to be active. In all models, average areal education was included as a proxy for the SES of the areas crossed, as we suspected it to be a confounding variable correlated with both the outcome and the explanatory variable of interest. As we had no reason to assume the relationship between the environmental factors and physical behaviors were linear, we tested a squared term and retain it in the models if it significantly improved the model (p-value<0.05 for a Chi-squared test).

All analyses were run using R [42], with libraries 'rgeos' [43] 'sf' and 'sp' [44] for spatial analysis and 'lme4' for mixed models [45].

Results

Description of journeys

The average length of the 692 journeys (shortest route) was 9.8±8.5 km (mean ± standard deviation), and the average duration 31.6±19.3 minutes. On average, participants spent 47% ± 36% of the travel time in SB (typically sitting or reclining), 25% ± 26% standing or performing light body movements (ST), and 27% ± 27% performing PA (typically walking and bicycling). Full descriptive statistics of all variables used in the models are shown in Table 1.

Entirely active travels (N=112) were on average composed of large shares of PA (65% ± 29%), some ST (28% ± 25%), and very little SB (7% ± 19%); in journeys including public transports (N=307), compositions of physical behaviors were more balanced (PA: 32% ± 18%, ST: 28% ± 21%, SB: 40% ± 26%). Other journeys (N=273), i.e. those made with a private motorized vehicle, were mostly sedentary (PA: 6% ± 11%, ST: 22% ± 29%, SB: 73% ± 31%). The compositions of physical behaviors over the journeys by the commute mode are shown in a ternary plot (Figure 2) and in Table 2 (including the center and covariance matrix [150]).

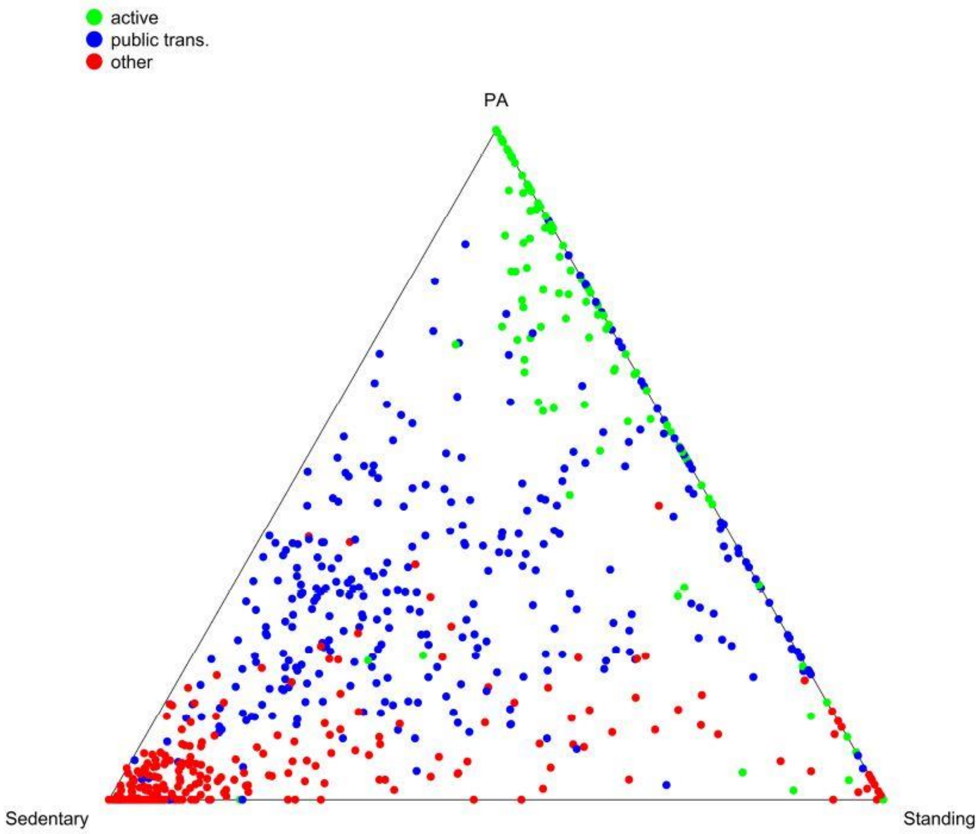

*Figure 2: **Distribution of compositions of behaviors over the journeys.** Colors indicate the commute mode: active - all the journey was made using physically active commute mean (typically walking or bicycling); public transports - public transports were used at least partly during the journey, other - otherwise, typically commuting by car.*

Model results

Results of the models are presented as ratios ST/SB or PA/SB (followed by the 95%-confidence interval in square brackets) throughout this section, and in Table 3 and Figure 3. As ratios are less intuitive, we added the predicted compositions for the range of values taken by the variables of interest. These are reported in Table 4 and illustrated in Figure 4.

Greenery

Green ratio along the shortest route significantly predicted the composition of physical behaviors over the journey. A 0.01 increase in green ratio led to a 11% [+2%, +22%] increase in the ST/SB ratio, and a 13% [3%, 23%] increase in PA/SB. By comparison, the green ratio for the journeys was typically bound between 0.01 and 0.07 (1st and 9th deciles). Regarding the average distance from the green road network, 1 SD increase in

(+198m) led to a decrease in ST/SB and PA/SB ratios by respectively -19% [-37%, +5%] and -20% [-37%, +5%]. The predicted decrease for this variable was, however, not statistically significant.

Destination density

Our hypothesis regarding the effect of destination density on behavior could not be clearly verified by our data. Models predicted a practically unchanged ST/SB ratio (-4% [-33%, +37%]) and a positive, but statistically insignificant rise in PA/SB ratio (+18% [-17%, 70%]) for a 1 SD increase in destination density (+695 destinations per km²).

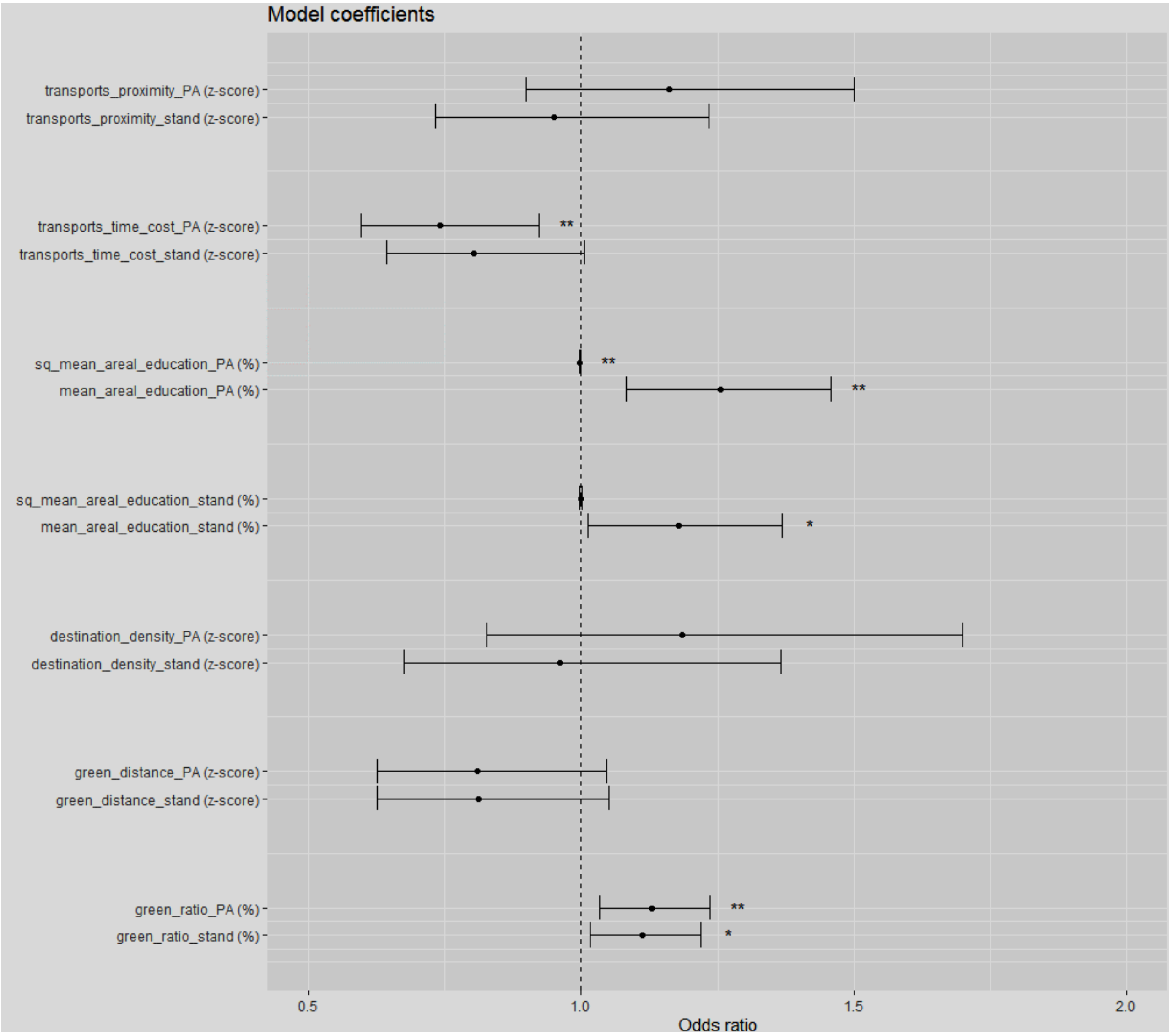

Figure 3: ***Model coeffcients*** *for the effects of environmental variables on the composition ratios standing/SB and PA/SB. For example, the predicted average effect of a 1%-increase in green ratio on the PA/SB ratio is 1.13 [1.03-1.23]. * : p-value < 0.05; ** : p.value < 0.01.*

Average areal education

Among the variables of interests tested, average areal education (share of residents holding a university degree) along the shortest route was the only variable to have a significant non-linear effect on behavior. At baseline, an additional 1% in the areal education score multiplies the ST/SB ratio by a coefficient of 1.18 [1.01, 1.37], and the PA/ST ratio by 1.26 [1.08, 1.46]. These coefficients are modified by negative quadratic coefficients ($\exp(-0.0015x^2)$ and $\exp(-0.0022x^2)$), respectively. This points out to an inverted U-shaped effect of areal education on ST and PA, reaching its maximum around the mean score (52%). As the distribution of the educational score is right-skewed (1^{st} decile = 30%, 9^{th} decile = 62%), the positive marginal effect on ST and PA for low levels of areal education is larger than the negative marginal effect for high levels. The relation between the variables is illustrated in Figure 4.

Public transports efficiency

Maximum distance to a public transports was not found to have an important effect on behavior during journeys. The effects of an increase of 1 SD in distance (70m) were not statistically significant, and amounted to -5% [-27%, 23%] and +16% [-10%, 50%] for the ST/SB and PA/SB ratios, respectively. However, the public transit to car time cost had a positive effect on SB. One SD increase in the time cost (+80% travel duration using public transports compared to car) had an estimated effect of -20% [-36%, +0%] on the ST/SB ratio and –26% [-40%, -8%] on the PA/SB ratio.

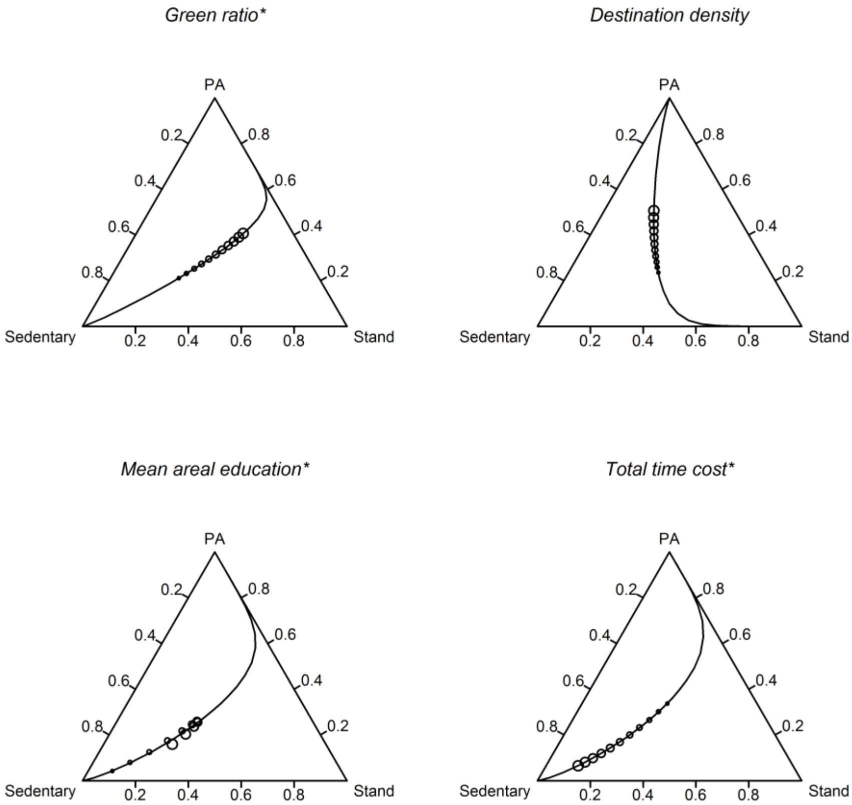

Figure 4: ***Composition of physical behaviors predicted by the models*** *for various values of the variables of interest while all covariables take their mean values. The circles, from the smallest to the biggest, represent the predicted compositions for the lowest to the highest value in the population.*

Discussion

Interpretation of findings

In this study, we investigated the effects of environmental attributes of home-work routes on physical behaviors during the journeys in urban adults. It yielded three important insights. First, high proportions of green area along the route increase the share of time spent in non-sedentary behaviors (standing or being physically active). Second, higher time cost of public transport compared to private car results in lower shares of time spent in standing and PA. The fact is all the more interesting as it remains when individual income and education were controlled for, suggesting that the use of public transports may be more a choice than a need. Third, the residents' level of education along the route positively affects standing and PA in low education areas, but the effect is reversed (although somewhat weaker) in areas with high areal education score.

Our findings about the positive effects of greenery on PA agree with previous work on this topic [23]. However, to our knowledge, they are the first to assess this effect in a causal, contemporaneous design with sensor-derived data. This effect was better captured considering all green spaces intersecting the buffers around the routes than when considering only green paths. This hints to a sensitivity to greenery present in a broader area (100m), rather than to the sole immediate surroundings.

Time cost of public transportation was significantly associated with active behaviors, in accordance with previous literature [32–34]. Distance to public transport station, however, did not predict any behavior significantly. It should be noted that, for our population, 90% of the trips started and ended within a 255m radius around a station, meaning that the distance to the station is unlikely to play a major role in the choice of commute mode. In addition, distance to station might have conflicting effects, as it is an incentive to use car and reduce activity, but it increases activity when actually walking to station.

Areal education, up to a certain level, had a negative effect on SB and a positive one on PA, independent of personal SES, safety and other environmental variables tested here. This finding concurs with the findings for an adult, urban French Canadian population [31]. It is plausible that areal education is a proxy for an ensemble of environmental features, both social and physical, which favor physical activity. However, it is difficult to understand why this relation is inverted in the highest quantiles of areal education. Although the mechanisms of this observed relationship remain uncertain, this finding clearly points to lower incentives to active commuting in low-SES areas in the Paris region. It suggests that future studies will need to investigate the role neighborhood SES in the relationship between environment and PA.

Our hypothesis regarding the positive effect of destination density on activity, supported by several extant studies [25–28], could not be verified. A plausible explanation is that our study omits journeys including stops that are not travel-related.

Yet, destination-rich routes can be an incentive to walk or bicycle because they offer good opportunities to visit places, such as shops, along the way. In fact, considering all sorts of travels in the same population, Chaix et al. found a positive association between destination density and activity [35]. Interestingly, the models suggest that people who report giving importance to the presence of shops in their neighborhood are much likelier to be non-sedentary in journeys.

The present article investigated the effect of environment on both PA and SB. Interestingly, the effect magnitudes on ST and PA in all significant models were close to each other (although slightly higher for the latter). Thus, the decrease in SB was made up by nearly equally proportionate increases in ST and PA. Insofar as this finding is generalizable to the study of physical behaviors during commuting, it suggests that there is no need to use these refined categories of physical behaviors, and that classical measures of body motions are sufficient.

Strengths and limitations

By looking at the home-to-work and work-to-home journeys, the present study addressed the major issues that undermined causal inference in most past studies. More than assessing the co-occurrence of environmental features and physical activity, it aimed at determining the causal effects of environment on physical behavior, using precise, sensor-derived data. In addition, this study is one of the very few to address the question of posture during commuting using objective measurements of body posture.

Yet, the issue of neighborhood selection is not fully addressed in this study: we still do not know whether a specific workplace was chosen by an individual because it fitted her/his preferred commute mode. Thus, it would have been better to ask participants whether the agreement level between the environment around their workplace and their preferred commute mode played a role in the choice of a workplace, as we did for the choice of the residence. However, we argue that such considerations are very unlikely to be critical in the choice of a workplace, especially after we controlled for factors related the choice of residence.

In this study, the shortest path was determined using GoogleMaps. With the Google Maps API, it was impossible to calculate itineraries back at the time of the study. We therefore had to assume that changes that occurred in the topography and traffic infrastructures of the Paris region over the last 4-6 years did not greatly alter the routes taken by individuals and that, if it was the case, that it would not substantially affect the levels of exposure to our variables of interest. The same caveat needs to be emphasized with regard to our environmental variables, which were calculated using databases that were a few years old at the time of the study.

Conclusion

Using an innovative design oriented towards causal inference and objective, precise measures, this study generated new insights bridging issues of urban planning and public health. Individual variables and journey length being equal, individuals were less likely to be sedentary during journeys across areas featuring more greenery. Likewise, they were less likely to be sedentary during journeys that could be made with time-efficient public transports. Results also pointed to important disparities in the levels of PA across neighborhood SES, with the highest level of PA achieved when the routes for commute crossed middle-to-high SES neighborhoods. We conclude that, prioritizing greenery and a time-efficient public transportation networks are urban planning steps that can effectively help fight the pandemic of physical inactivity and sedentary behavior, and that attention should be paid to equal access to environmental features promoting PA. Future research should build upon the framework proposed here, both qualitatively and quantitatively, by examining other environmental attributes, and with datasets large enough to improve statistical inference.

Tables

Table 1: Descriptive statistics of variables in the study										
		Quantiles								
		0%	10%	50%	90%	100%	Mean	Standard deviation	No. of journeys	No. of participants
Personal variables	age (years)	0	49	60	69	82	59	8	692	121
	sex (male=1)						0.66		692	121
	couple						0.71		692	121
	Children<14 yo						0.45		692	121
	income category	1	1	2	3	3	1.83	0.86	692	121
	educational category	1	1	3	3	3	2.27	0.91	692	121
	Safety in nghb.	1	2	2	3	3	2.30	0.59	692	121
Importance of variable in choice of neighborhood	greenery	0	0	2	3	3	2.06	1.00	692	121
	shops	0	1	2	3	3	2.24	0.86	692	121
	SES	0	0	2	3	3	1.50	0.97	692	121
	Public transports	0	0	3	3	3	2.26	1.07	692	121
Trip attributes	Length (km)	0.56	1.69	7.38	21.11	51.20	9.84	8.49	692	121
	actual length (km)	0.04	1.20	7.64	23.33	64.34	10.66	9.95	692	121
	Duration (min)	4.00	12.04	27.85	52.91	243.00	31.61	19.34	692	121
	green ratio	0	0.0095	0.0336	0.0715	0.1389	0.0382	0.0274	692	121
	green distance (m)	42	137	341	606	1919	360	198	692	121
	destination density (km^{-2})	4	155	487	1854	4476	777	695	692	121
	mean areal education	0.1785	0.2958	0.5149	0.6189	0.7026	0.4854	0.1261	692	121
	transp. proximity (m)	31	91	152	255	532	162	70	692	121
	trans. time cost	0.81	1.26	1.84	2.89	7.51	2.01	0.80	676	120
Composiiton of physical behaviors	sedentary	0	0	0.4988	0.9456	1	0.474	0.3572	692	121
	standing	0	0.0258	0.1654	0.6291	1	0.2554	0.252	692	121
	PA	0	0	0.1957	0.7244	1	0.2705	0.2718	692	121
Commute mode	active						0.16		692	121
	public transportation						0.53		580	107

Table 1 : Descriptive statistics of the variables used in this study

Table 2: Descriptive statistics of compositions of physical behaviors over the journeys				
ALL JOURNEYS				
		Sedentary behavior	Standing/light movements	Physical activity
Average (Center)		0.5310	0.2103	0.2587
Variation matrix	Sedentary behavior	1.6440		
	Standing/light movements	-0.7257	0.7814	
	Physical activity	-0.9183	-0.0557	0.9740
ACTIVE JOURNEYS				
Average (Center)		0.1440	0.2217	0.6343
Variation matrix	Sedentary behavior	0.9000		
	Standing/light movements	-0.2816	0.5534	
	Physical activity	-0.6184	-0.2718	0.8902
PUBLIC TRANSPORTS JOURNEYS				
Average (Center)		0.4345	0.2354	0.3301
Variation matrix	Sedentary behavior	0.8418		
	Standing/light movements	-0.5967	0.6954	
	Physical activity	-0.2451	-0.0987	0.3438
OTHER JOURNEYS				
Average (Center)		0.7817	0.1247	0.0936
Variation matrix	Sedentary behavior	1.0379		
	Standing/light movements	-0.7038	0.8705	
	Physical activity	-0.3341	-0.1667	0.5008

Table 2 : Descriptive statistics of compositions of physical behaviors over the journeys of different types. The center (compositional mean) is defined as $C[exp\ (\frac{1}{N}\sum_i ln\ x_i)]$*, where C[.] is the closure operator dividing the parts by their sum. The variation matrix is defined as* $var\left(ln\frac{x_i}{x_j}\right)$*. These are preferred to the arithmetic mean and variance because of the constrained character of a composition [36].*

Table 3: Model coefficients [95% confidence intervals]						
Variable of interest	Green ratio (%)		Distance from green (z-score)		Destination density (z-score)	
	Standing	PA	Standing	PA	Standing	PA
Intercept	0.0043 [0.0001, 0.1574]	0.0008 [0, 0.0298]	0.0075 [0.0002, 0.2553]	0.0015 [0, 0.052]	0.0016 [0, 0.0545]	0.0004 [0, 0.0127]
Variable of interest	1.1125 [1.016, 1.2181]	1.1292 [1.0324, 1.235]	0.8104 [0.6253, 1.0503]	0.8092 [0.6265, 1.0452]	0.9597 [0.6739, 1.3666]	1.1853 [0.8265, 1.6998]
Squared variable of interest						
Reported importance of var. of interest	0.9986 [0.6966, 1.4316]	0.9006 [0.6181, 1.3122]	0.9968 [0.6974, 1.4247]	0.9042 [0.6203, 1.3182]	1.7746 [1.1905, 2.6453]	1.7151 [1.124, 2.617]
Age (years)	0.9484 [0.655, 1.3733]	1.0357 [0.7033, 1.5252]	0.954 [0.6605, 1.3777]	1.0452 [0.709, 1.5408]	1.0046 [0.7025, 1.4367]	1.1001 [0.7526, 1.6081]
Sex (1 = male)	0.6786 [0.3143, 1.4651]	0.7731 [0.3461, 1.7266]	0.7113 [0.3317, 1.5252]	0.8203 [0.367, 1.8336]	0.8703 [0.4155, 1.8229]	1.0549 [0.4822, 2.308]
Income category = 2	0.5835 [0.2361, 1.4421]	0.5413 [0.2105, 1.3921]	0.5806 [0.2366, 1.4245]	0.5411 [0.21, 1.3939]	0.6637 [0.2709, 1.6256]	0.6705 [0.2601, 1.7288]
Income category = 3	0.6806 [0.26, 1.7814]	0.6083 [0.2227, 1.6615]	0.7201 [0.2779, 1.8656]	0.6509 [0.2384, 1.7769]	0.8683 [0.3329, 2.2645]	0.8712 [0.316, 2.4023]
Education category = 2	2.1851 [0.6586, 7.2498]	2.9633 [0.8465, 10.3738]	2.5532 [0.7724, 8.4397]	3.4898 [0.9892, 12.3114]	1.934 [0.5952, 6.2841]	2.4559 [0.705, 8.556]
Education category = 3	1.2159 [0.5198, 2.8442]	1.698 [0.6987, 4.1267]	1.2674 [0.5461, 2.941]	1.7772 [0.7309, 4.3213]	1.2471 [0.5461, 2.848]	1.7323 [0.722, 4.1564]
Couple (1 = yes)	0.7915 [0.3346, 1.8724]	0.5818 [0.2366, 1.431]	0.8215 [0.3501, 1.9278]	0.6074 [0.2468, 1.4947]	0.7326 [0.3171, 1.6928]	0.5682 [0.234, 1.3798]
Children <14 at home (1 = yes)	1.5211 [0.6845, 3.3804]	1.9074 [0.8277, 4.3955]	1.4945 [0.6774, 3.297]	1.8633 [0.808, 4.2969]	1.2129 [0.5604, 2.6251]	1.4405 [0.635, 3.2679]
Mean areal education (%)	1.2138 [1.0421, 1.4138]	1.2926 [1.1115, 1.5032]	1.2005 [1.0318, 1.3968]	1.2765 [1.0977, 1.4844]	1.2158 [1.0471, 1.4116]	1.2895 [1.1107, 1.497]
Squared mean areal education (%)	0.9981 [0.9964, 0.9998]	0.9975 [0.9958, 0.9991]	0.9982 [0.9966, 0.9999]	0.9976 [0.9959, 0.9993]	0.9981 [0.9964, 0.9997]	0.9974 [0.9957, 0.999]
Length (km)	0.2441 [0.1659, 0.3592]	0.1247 [0.085, 0.1828]	0.2473 [0.1683, 0.3634]	0.1249 [0.0851, 0.1835]	0.2444 [0.166, 0.36]	0.1293 [0.0878, 0.1905]
Squared length (km)	1.3434 [1.2008, 1.5031]	1.5524 [1.3918, 1.7315]	1.3563 [1.208, 1.5228]	1.5688 [1.4005, 1.7572]	1.327 [1.1864, 1.4842]	1.5194 [1.3618, 1.6954]

Table 3 : Exponentiated coefficients (β) of mixed linear regressions modelling the relationship between personal and environmental variables and the log ratios ST/SB and PA/SB. After a change of one unit in the independent variable, the ratio is estimated as β times the ratio before the change.

Table 3: Model coefficients [95% confidence intervals] - continued						
Variable of interest	Mean areal education (%)		Maximum distance to station (z-score)		Time cost of public transports (z-score)	
	Standing	PA	Standing	PA	Standing	PA
Intercept	0.0169 [0.0006, 0.5095]	0.003 [0.0004, 0.0889]	0.0055 [0.0002, 0.1667]	0.004 [0, 0.0137]	0.0068 [0.0002, 0.1991]	0.007 [0, 0.02]
Variable of interest	1.1775 [1.0127, 1.3692]	1.2561 [1.0816, 1.4588]	0.9501 [0.732, 1.2331]	1.1605 [0.8977, 1.5003]	0.8034 [0.6423, 1.0048]	0.7407 [0.5952, 0.9218]
Squared variable of interest	0.9985 [0.9968, 1.0001]	0.9978 [0.9961, 0.9994]				
Reported importance of var. of interest	0.7417 [0.5201, 1.0576]	0.6586 [0.4545, 0.9543]	1.278 [0.9136, 1.7878]	1.5187 [1.0684, 2.1589]	1.2792 [0.9146, 1.7891]	1.4618 [1.0308, 2.073]
Age (years)	1.0331 [0.7132, 1.4964]	1.1448 [0.7771, 1.6866]	0.9756 [0.6762, 1.4078]	1.0407 [0.7088, 1.5279]	0.9609 [0.6685, 1.3812]	1.0341 [0.7086, 1.509]
Sex (1 = male)	0.6408 [0.3039, 1.3512]	0.7545 [0.3464, 1.6431]	0.7305 [0.346, 1.5424]	0.975 [0.4463, 2.1299]	0.7805 [0.3729, 1.6335]	0.9885 [0.4585, 2.1312]
Income category = 2	0.5801 [0.2354, 1.4296]	0.5305 [0.207, 1.3595]	0.6384 [0.2566, 1.5882]	0.6682 [0.258, 1.7305]	0.6832 [0.2769, 1.6858]	0.66 [0.2581, 1.6878]
Income category = 3	0.6569 [0.2507, 1.7217]	0.574 [0.2099, 1.57]	0.7729 [0.2923, 2.0434]	0.8268 [0.2996, 2.282]	0.7828 [0.3006, 2.0385]	0.7781 [0.2876, 2.105]
Education category = 2	2.4498 [0.7408, 8.101]	3.4081 [0.9775, 11.8816]	2.124 [0.6412, 7.0355]	2.808 [0.8021, 9.8297]	2.0265 [0.6176, 6.6494]	2.609 [0.7578, 8.9824]
Education category = 3	1.3773 [0.5873, 3.2303]	1.9984 [0.8204, 4.8677]	1.1966 [0.5086, 2.8149]	1.4951 [0.6111, 3.6582]	1.1646 [0.5008, 2.7081]	1.5734 [0.6537, 3.7875]
Couple (1 = yes)	0.8862 [0.3751, 2.0935]	0.6897 [0.281, 1.693]	0.8467 [0.3586, 1.9993]	0.605 [0.2463, 1.4861]	0.8384 [0.3576, 1.9652]	0.6097 [0.2513, 1.4791]
Children <14 at home	1.5052 [0.6922, 3.2732]	1.7936 [0.7965, 4.0392]	1.3093 [0.5906, 2.9027]	1.4441 [0.6271, 3.3256]	1.2709 [0.5765, 2.8016]	1.3977 [0.6133, 3.1855]
Mean areal education (%)			1.1897 [1.0233, 1.3831]	1.2774 [1.1, 1.4833]	1.1818 [1.0181, 1.3718]	1.2751 [1.0989, 1.4795]
Squared mean areal education (%)			0.9983 [0.9967, 1]	0.9976 [0.996, 0.9993]	0.9983 [0.9967, 1]	0.9975 [0.9959, 0.9991]
Length (km)	0.2517 [0.1724, 0.3676]	0.1255 [0.0862, 0.1829]	0.2588 [0.1766, 0.3794]	0.126 [0.0861, 0.1843]	0.2401 [0.1626, 0.3546]	0.1227 [0.0831, 0.1811]
Squared length (km)	1.3212 [1.1827, 1.4759]	1.5313 [1.3747, 1.7056]	1.321 [1.1815, 1.477]	1.5152 [1.3595, 1.6888]	1.399 [1.2023, 1.6279]	1.5636 [1.3478, 1.814]

Table 4 : Predicted compositions												
	Green ratio			*Destination density*			*Areal education*			*Public transports time cost*		
Percentile of variable of interest	*Sedentaty*	*Stand*	*PA*	*Sedentaty*	*Stand*	*PA*	*Sedentaty*	*Stand*	*PA*	*Sedentaty*	*Stand*	*PA*
0	53.03	25.86	21.11	42.49	33.98	23.53	86.6	9.14	4.26	33.98	32.25	33.77
10	49.09	27.77	23.14	41.73	32.5	25.77	78.15	13.95	7.89	39.04	30.82	30.14
20	45.17	29.63	25.2	40.86	30.99	28.15	68.45	18.99	12.56	44.32	29.1	26.58
30	41.3	31.42	27.28	39.88	29.46	30.66	59.12	23.42	17.46	49.7	27.14	23.16
40	37.54	33.11	29.35	38.81	27.91	33.28	51.53	26.78	21.69	55.06	25.01	19.93
50	33.92	34.69	31.39	37.63	26.36	36.01	46.36	29.04	24.61	60.28	22.77	16.95
60	30.47	36.14	33.39	36.37	24.81	38.82	43.73	30.34	25.93	65.25	20.5	14.25
70	27.23	37.45	35.32	35.03	23.27	41.71	43.63	30.8	25.57	69.88	18.26	11.86
80	24.21	38.61	37.18	33.61	21.74	44.65	46.02	30.37	23.61	74.12	16.11	9.77
90	21.42	39.62	38.95	32.14	20.24	47.62	50.9	28.86	20.25	77.93	14.09	7.98
100	18.88	40.49	40.63	30.61	18.78	50.61	58.13	26.01	15.87	81.31	12.23	6.47

Table 4 : Predicted compositions of physical behaviors for various levels of our variables of interest, all other variables taking their mean values otherwise.

Ethics statement

All participants signed an informed consent form. The study protocol was approved by the French Data Protection Authority (Decision No. DR-2013-568 on 2/12/2013).

Data availability

Data is available upon request

Funding

The RECORD study was funded by the National Research Agency (ANR), Institute for Public Health Research (IReSP), National Institute for Prevention and Health Education (INPES), National Institute of Public Health Surveillance (InVS), French Ministry of Research, French Ministry of Health, National Health Insurance Office for Salaried Workers (CNAM-TS), Regional Direction of Health and Social Affairs in Ile-de-France (DRASSIF), Regional Group of Public Health in Ile-de-France (GRSP), and Regional Direction for Youth and Sports in Ile-de-France (DRDJS).

Acknowledgments

The authors thank the Center for Preventive and Clinical Intervention (IPC) in Paris, the participants in the study and the collaborators who made this study possible.

Conflicts of Interest: The authors declare no conflicts of interest.

CHAPTER V

Commuting with public transports: How does it influence physical activity levels?

Chapitre V: Le comportement physique selon le mode de transport utilisé (résumé français)

Dans le chapitre précédent, nous avons montré que diverses caractéristiques de l'environnement urbain pouvaient influencer les comportements physiques de ceux qui s'y trouvaient. Entre autres caractéristiques, nous avons étudié l'infrastructure des transports en communs et son effet sur le budget des comportements physiques adoptés lors des déplacements. Ce chapitre vise à mieux comprendre cet effet en explorant de façon systématique les budgets-temps lors des déplacements en fonction du mode de transports choisi.

Le temps passé en déplacement, quel que soit le mode de transport, représente en moyenne une fraction importante de la journée (1h45 dans notre population RECORD). Ce temps peut donc être considéré comme une ressource potentielle de premier ordre pour la réalisation du temps d'activité physique recommandé pour la santé. L'intérêt scientifique pour ce sujet n'est donc pas nouveau, mais cette étude, par la précision des méthodes employées, innove par rapport aux quelques études réalisées précédemment.

Étudier empiriquement les budgets-temps comportementaux dans une population en fonction du mode de transports présuppose une connaissance du comportement physique des individus en continu ainsi que le mode et le temps précis des transports qu'ils empruntent. D'une part, s'appuyer sur des données déclaratives concernant le comportement physique et le temps de transports peut être problématique, puisque ces données souffrent souvent d'un biais de mémoire. D'autre part, les rares études ayant collecté à la fois des données de localisation et d'activité physique se sont montrées peu précises dans la prédiction du moyen de transport utilisé par les individus monitorés. En revanche, notre étude, toujours avec les mêmes sujets RECORD présentés plus haut, a croisé les données de localisation (capteur GPS) et d'activité (accéléromètre) avec un journal de bord tenu par les sujets de l'étude contenant des informations sur les moyens de transports utilisés. Il en résulte un aperçu détaillé et fiable des déplacements des individus tout le long de la journée. Ces données ont été vérifiées manuellement et, en cas de doute, les sujets ont été contactés pour préciser des déplacements qui semblaient incertains. Ce degré de précision nous a permis de diviser les déplacements, quand cela était applicable, en plusieurs *étapes* effectuées avec des modes de transports différents et d'obtenir non seulement une estimation précise des budgets-temps pour des catégories de transports « pures », mais aussi une vue d'ensemble sur les déplacements consistants d'étapes réalisées avec des modes de transports différents et séparées par des « stations » (arrêt de bus, parking…), pour lesquelles aucun déplacement n'est enregistré.

Nous avons classé les déplacements en quatre catégories : marche uniquement, autres transports actifs (bicyclette, trottinette…), transports en commun et véhicule privé. Une cinquième catégorie « mixte » désignait un déplacement comportant deux étapes de catégories différentes autres que la marche. Ainsi, notre analyse contient des données d'activité physique pour 4683 déplacements réalisés en 7692 étapes. En contrôlant l'autocorrélation temporelle entre les étapes successives et en standardisant par la durée des déplacements, nous avons constaté que, par rapport à 10 minutes de déplacement en véhicule privé, le déplacement à pied entraînait une baisse de 5,26 minutes du temps sédentaire, et une hausse de 4,83 minutes du temps d'activité modérée-à-vigoureuse. Les valeurs obtenues pour les autres transports actifs étaient quasiment identiques à celles obtenues pour la marche. Les transports en commun, quant à eux, entraînaient une baisse de 3,10 minutes du temps sédentaire et une hausse de 2,44 minutes d'activité modérée-à-vigoureuse par jour comparé à l'utilisation d'un véhicule personnel motorisé.

Une comparaison entre véhicule privé et des transports actifs, surtout sur en ville, doit tenir compte du fait que même en véhicule privé, il existe toujours une composante d'activité correspondant au temps de marche entre l'endroit où la voiture est garée et la destination, et que mêmes dans les déplacements actifs, des temps d'arrêt peuvent survenir (par exemple arrêt au feu, repos sur un banc). Il n'est donc pas surprenant de constater que, sur 10 minutes, les transports actifs accusaient un gain de moins de 5 minutes en termes d'activité modérée-à-vigoureuse. Avec environ 2,5 minutes d'activité modérée-à-vigoureuse, les transports en commun s'avèrent, quant à eux, un moyen également intéressant de réduire l'inactivité physique sur la journée, surtout en tenant compte du fait que les déplacements en transports en commun sont souvent assez longs.

L'étude présentée dans ce chapitre illustre et complète le chapitre précédent. Dans le chapitre précédent, nous avons montré qu'une bonne infrastructure des transports en commun, *de manière générale*, encourageait l'activité au niveau de la population. Les résultats de ce chapitre nous montrent que l'utilisation des transports en commun entraînait une hausse importante de l'activité physique sur l'ensemble du déplacement. De ces deux chapitres, on peut conclure que, dans des conditions similaires à celles de l'étude, une infrastructure de transports efficace augmente la probabilité d'utiliser ce mode de transports et donc le temps d'activité physique et le temps non-sédentaire.

Cette étude, à cause de son petit effectif, n'est certainement pas tout à fait représentative du très grand nombre de formes de transports et de leurs combinaisons qu'on pourra trouver dans une mégapole comme la région parisienne. De même, les résultats observés en région parisienne ne sont pas forcément généralisables à des villes plus petites et à des modèles d'urbanismes différents. Malgré ces limites et comparant ce qui est comparable, cette étude indique qu'un réseau de transports en commun rapide et efficace peut représenter un moyen efficace de lutter contre la pandémie de l'inactivité physique.

Physical activity and sedentary behavior related to transport activity assessed from multiple body-worn accelerometers: the RECORD MultiSensor Study

Sanjeev Bista[1], Isaac Debache[2], Basile Chaix[1]

[1]Sorbonne Université, INSERM, Institut Pierre Louis d'Épidémiologie et de Santé Publique, Nemesis research team, F75012, Paris, France
[2]Institut Pluridisciplinaire Hubert Curien (IPHC) UMR 7178, Centre National de la Recherche Scientifique (CNRS), Université de Strasbourg 67000 Strasbourg, France.

Abstract

The study explored the physical activity and sedentary behaviour related to transport activity, to support public health and transport policies aiming to encourage people to reach daily recommendation of physical activity. In 2013-2015, the RECORD MultiSensor study collected data from 155 participants using two accelerometers worn on the thigh and trunk of participants and Global Positioning System (GPS) receivers complemented with a GPS-based mobility survey. Relationships between transport modes and the durations and partition patterns of physical behaviors were established at the trip stage (n=7692) and trip levels (n=4683) using multilevel linear models with a random effect at the individual level and taking into account temporal autocorrelation. Participants travelled for a median of 1 hour 45 minutes per day. Trip stages and trips involving walking, other active modes, or public transport were associated with a lower sitting duration and a higher MVPA duration than those with a personal motorized vehicle. Using public transport was associated with a lower number of transitions between sedentary behaviors and non-sedentary behaviors but with a larger number of transitions between non-sedentary behaviors and moderate to vigorous physical activity than relying on a private motorized vehicle. Our study is the first to assess the association of transport mode used with physical activity and sedentary behaviors captured with thigh- and trunk-worn accelerometers at both the trip stage and trip levels. Our results demonstrate that in addition to active transport modes, encouraging people to use public transport increases physical activity and reduces sedentary time.

Introduction

A report from the United Kingdom reported that lack of physical activity is ranked as the 4th leading cause of mortality (1). Similarly, a surveillance article concluded that the level of physical activity is low throughout the world, where only about 69% of adults meet the recommended physical activity level (2). Apart from physical activity itself, studies have highlighted that prolonged sedentary behaviour is associated with the risk of diabetes, obesity, and some cancers (3–5).

Active transport, as a component of physical activity, has positive health effects (6). Adults have been shown to be active when they are travelling back and forth to work, and particularly in trips made with public transport (7). However, measuring physical activity in trips is particularly challenging. For example, a study used Global Positioning System (GPS) receivers, movement sensors, and heart rate monitoring for measuring the physical activity in a limited number of trips, only home–work trips, and therefore had a limited generalizability (8). Many studies that used GPS receiver, accelerometers, and advanced algorithm to predict transport mode (9,10) have not verified the predicted transport mode with participants, which likely results in prediction error in their data. On the opposite, our study validated transport modes trip by trip with all the participants involved in the study through phone calls to participants (mobility survey) (7,11).

Another pitfall of previous research is that studies have often aggregated the information on physical behaviors into daily averages of physical activity or have aggregated information into overall indicators of energy expenditure or percentage of cumulative time spent in physical behaviors (e.g., sitting, standing, walking, etc.). However, such analyses ignore the continuous sequence physical behaviors performed by individuals over the day, which are known from sensors, e.g., at the minute level (12). Our study, using two accelerometers worn on the participant's trunk and thigh, attempted to accurately measure the body postures and movements per second, which can be thereafter summed up over any interval of time (12). Relatedly, a secondary aim of our paper was to compare hip-worn accelerometers and combined thigh-trunk accelerometers in their ability to assess physical behaviors such as sitting or standing at the trip and trip stage levels.

Several studies investigated patterns of physical activity at the trip level; nevertheless, they did not accurately classify the transport modes used in trips (13–15). In order to fill this gap, our study aimed to establish profiles of physical behaviors, in terms of duration, of partition (12,16), and of number of transitions between categories of behaviors, by type of transport mode used. Partition in our study, as opposed to the cumulative time of a behavior, relates to the number and lengths of continuous periods over which the behavior is detected. Contrary to studies that only examined sedentary behaviour (SB) (16,17), partition profiles in our study encompass other behaviors such as non-sedentary behaviour (NSB) and moderate to vigorous physical activity (MVPA).

Overall, the objective of our study was to analyze the relationships between transport mode and the duration and partition profile of physical behaviors, at both the trip stage and trip levels, using linear mixed models.

Methods

Population

The data used come from the RECORD MultiSensor Study (18), of the Record Cohort (19). From February 2007 to March 2008, 7290 participants were recruited without a priori sampling during preventive health cheakups conducted in four sites of the Centre d'Investigations Préventives et Cliniques (IPC) funded by the National Insurance System for Employees and Salaried Workers. People had to be 30 to 79 year old, had to live in 10 districts (out of 20) of Paris and 111 other municipalities of the Ile-de-France region, and had to be free of cognitive and linguistic disabilities to be eligible for the study. In 2011–2015, these participants as well as new participants from the IPC medical center were invited to take part in the second wave of the RECORD Study. From September

2013 to June 2015, participants from the second wave of RECORD were invited to the RECORD MultiSensor Study whenever sensor devices were available (i.e., brought back by previous participants). The study has been approved by the French Data Protection Authority (Decision No. DR-2013-568 on 2/12/2013).

Among the 286 final participants to the RECORD Multisensor Study, 157 were included in the substudy where they had to carry a BT-Q1000XT GPS receiver, a wGT3X+ waist-worn accelerometer, and two combined accelerometers at the trunk and thigh (VitaMove, Temec Instruments, The Netherlands) over a period of 7 days. For 2 participants, the trunk and thigh accelerometers did not function properly; these participants were excluded from the analysis, leaving 155 participants for the analyses. Out of 8085 stages of trips made by these 155 participants, 393 (4.9%) were not included due to a failure of the VitaMove devices or because these devices were not worn. Therefore, 7692 trip stages from 4683 trips were included in our analysis.

Classification of trip stages and trips

Trip stages are portions of trips with a unique mode. Within a trip, two trip stages are necessarily separated by an episode of transfer between the two assigned to a punctual location, which also count as a trip stage.

The data extracted from the BT-Q1000XT GPS receiver were pre-processed after the 7-day data collection in order to identify the visited places as well as the start and end times of each trip stage, defined as a segment of a trip using a unique transport mode (20). Using a web mapping application, these data were then consolidated during a phone mobility survey with the participants, producing in the end a detailed timetable covering the 7-day observation period (see Web Appendix 1 for details). This timetable consisted of a time-stamped list of the visited places and trip stages between them.

Each trip comprises one or several trip stages. In trips with several stages, the whole trip also includes the transfer time between several trip stages. Our crude classification of trip stages was as follows: entirely walked, biking/rollers/skateboard (other active modes); public transport; privately owned vehicles; and other non-local trips involving long-distance trains and planes. Our detailed classification of modes further distinguished driving own personal vehicle from travelling through a private vehicle as a passenger (including taxi); and subdivided public transport into (i) bus/coach, (ii) metro, (iii) suburban train including RER (trains travelling within Paris and suburban cities), standard suburban trains, and TER (trains for joining Paris to suburbs or nearby regions), and (iv) trams.

At the trip level, trips were classified into the same categories. Trips that comprised two or more non-walking modes were assigned to a separate multi-mode trip category.

Additionnally, each trip or trip stage was coded as on a weekday vs. weekend day and performed in Spring or Summer vs. Automn or Winter.

Processing of accelerometer data

The VitaScore software was used to process the VitaMove trunk and thigh accelerometer data, and classify each second into 5 groups: sitting, lying, standing, light physical activity (LPA, including slow walking), and moderate-to-vigorous physical activity (MVPA).

ActiLife 6.11.9 was used to process the waist-worn accelerometer data. The standard inclinometer data indicated the number of seconds of sitting, lying, and standing for each of the 5-second epochs. For comparing the VitaMove posure data to the Actigraph inclinometer data, the VitaMove standing, LPA, and MVPA were considered as standing.

For each trip or each trip stage, we calculated the cumulated duration in each physical behavior. We also calculated a version of these variables standardized per units of 10 minutes of trip or trip stage.

For a simplified partition analysis based on the VitaMove data, physical behaviours were categorized those into 3 broad groups: SB (combining lying and sitting), NSB (including standing and LPA), and MVPA. Identifying uninterrupted segments of SB, NSB, and MVPA within each trip / trip stage, we determined the median length of such segments separately for SB, NSB, and MVPA within each trip and trip stage. This indicator was standardized per 1 minute of trip or trip stage. We also calculated the number of transitions between SB and NSB, NSB and MVPA, and SB and MVPA within each trip and each trip stage. The latter partition indicator was standardized per units of 10 minutes of trip / trip stage.

Statistical Analysis

Unstandardized and standardized durations of physical behaviors and partition indicators were tabulated by transport modes at the trip and trip stage levels. Relationships between transport modes and unstandardized and standardized durations of physical behaviors were estimated separately at the trip level and trip stage level (one observation per trip and trip stage) using multilevel linear models with a random effect at the individual level. Time autocorrelation was also accounted for, using an AR(1) continuous autoregressive structure.

As discussed previously, confounding by individual characteristics is unlikely for the relationship of interest (7). Weekday/weekend day and season of the trip were associated with durations of certain physical behaviors, but not with durations standardized by 10 minutes of travel time. Therefore, they were only introduced in models for unstandardized outcomes.

We also estimated multilevel models with interaction terms between transport mode and weekdays/weekend days and/or between transport mode and season. Since the ranking of transport modes was similar in the models with only one or with the two interaction terms, we used the models with a single interaction term for plotting the predicted durations of physical behaviors by weekday/weekend day and season. Models with two interaction terms are reported in Web Appendix 2.

In order to compare the associations between transport modes and the duration of physical behaviors as estimated from waist-worn accelerometrers (Actigraph) and trunk- and thigh-worn accelerometers (VitaMove), we re-estimated the regression models among 4008 trips and 6901 trip stages (154 participants) with information for both sensors.

All statistical analyses were conducted using R software (version 3.4.4) and R Studio (version 1.1.463) (21).

Results

Sample description

The 155 participants had an average age of 50 years (range: 34–82 years). Among them, 98 were males. Fifty-five of them were from Paris, 30 lived in the close suburb, and 69 in the far suburb. Thirty-eight participants had no formal education or primary education or lower secondary education; 33 had a higher secondary education or lower tertiary education; 30 had an intermediate tertiary education; and 54 participants had an upper tertiary education. Among participants, 122 had a stable job, 4 had fixed-term contracts, and 4 participants were unemployed.

Descriptive information on trips

The median follow up time of participants in our study was 7 days (interdecile range: 5–7 days). Participants had a median number of trips per day of 5 (interdecile range: 3–7), corresponding to a median number of trip stages per day of 8 (interdecile range: 4–13). Participants were travelling (as opposed to being at a place) for a median of 1 hour 45 minutes per day (interdecile range: 56 minutes – 3 hours 2 minutes). Following Web Appendix 3, the most frequently used mode of transport was walking, corresponding to 53.8% of trip stages 39.6% of trips, followed by private motorized vehicles, corresponding to 23.1% (19.9% as a driver and 3.2% as a passenger) of trip stages and 39.6% (31.8% as a driver and 4.5% as a passenger) of trips.

Association between transport mode and physical activity

In models for both unstandardized and standardized outcomes (Tables 1 and 2), not only trips or trip stages by walking or with other active modes but also (although to a lesser extent) those with public transport were associated with a lower sitting time than when using a personal motorized vehicle. In the models with standardized outcomes (Table 2), the coefficient showing that there was less sitting time in public transport was stronger in the model at the trip level than at the trip stage level (as opposed to public transport trip stages, public transport trips also include the walking episodes). This is not the case in the models with unstandardized outcomes (Table 1) that are difficult to interpret due to the fact that trips and trip stages with different modes have different durations.

Regarding MVPA, when durations of trips and trip stages were accounted for (standardized outcome, Table 2, third and fourth columns), trips and trip stages by walking, other active modes, and as expected to a lesser extent with public transport were all associated with more minutes of MVPA than those with a personal motorized vehicle. The coefficient showing more minutes of MVPA associated with public transport was stronger in the model at the trip level (Table 2, column 4) than in the model at the trip stage level (column 3), as public transport trips also include typically include walked trip stages.

Figure 1 reports average durations of sitting and MVPA in a trip by transport modes (predicted from separate models with an interaction of transport modes with either the weekend / weekday variable or the season variable, with unstandardized outcome). MVPA duration in a trip was higher during weekends than on weekdays in trips with all modes, although the difference was particularly sharp only for multi-modes trips (this finding is based, however, on 9 and 52 multi-mode trips on the weekend and on weekdays, respectively, and is attributable to the fact that these weekend trips had an average duration of 188 minutes vs. 81 minutes for the weekday trips). Regarding the interaction with seasons, spring or summer was associated with a longer duration of MVPA per trip for all transport modes (except perhaps multi-mode trips), with a non-overlapping confidence intervals only for trips with other active modes.

Partition profile: transition rates

As shown in Table 3, transport modes differ in the number of transitions among SB, NSB, and MVPA. For example, both at the trip stage and trip level, using public transport was related to a lower number of transitions between SB and NSB (or the other way round) than driving or being the passenger of a private motorized vehicle, but it was related to a larger number of transitions between NSB and MVPA. Walking or relying on other active modes had the largest number of transitions from NSB to MVPA. Compared to other two types of transitions, those between SB and MVPA were particularly rare.

Statistics on the length of uninterrupted episodes of physical behaviors (SB, NSB, and MVPA within trips and trip stages are reported in Web Appendix 4.

Comparison of waist-worn to and thigh- and trunk-worn accelerometers

Considering time periods with both waist-worn and thigh- and trunk-worn accelerometers, the standing duration per individual per eight hours of device wear time had a median of 290.4 minutes (interdecile range: 123.6, 434.1) when assessed with the thigh- and trunk-worn accelerometers, as compared to 271.6 minutes (interdecile range: 138.3, 388.5) when assessed with the single waist-worn accelerometer. The corresponding figures for sitting time were 183.8 minutes (interdecile range: 42.1, 352.2) and 208.6 minutes (interdecile range: 92.9, 330.0).

Table 4 shows that the contrast in sitting duration between using a personal motorized vehicle and the other modes (public transport, walking, and other active modes) was substantially overestimated by the waist-worn accelerometer compared to the the thigh- and trunk-worn accelerometers, in both duration-unstandardized and standardized models (trip stage model). The corresponding models at the trip level are reported in Web Appendix 5.

Discussion

Strengths and limitations compared to previous literature

Regarding strength of our approach, this paper is one of the few published studies to explore the association of transport mode with physical activity at the trip level using objective sensor-based measures measured outcomes (7,22). And it is the first to conduct such a detailed analysis with two complementary body-worn accelerometers that permit a more accurate assessment of body posture, including sitting. Two accelero-sensors placed on the trunk and thigh that provide information on the orientation of the body compared to the gravitation field are useful to infer body posture.

Another strength of this paper is that it performed this analysis comparatively at the trip stage level and trip level. Investigating the relationship between transport mode and physical behaviors is of interest both at the trip stage level, for a description of each transport mode, and at the trip level, to investigate how the different non-walking modes generate walking and physical activity. Previous studies did not reach this level of precision, for example those which modeled the relationship between transport mode and physical activity at the individual level rather than trip stage and trip levels (23). A study that analyzed trip-level information used self-reported rather than accelerometer-derived physical activity, which makes the findings less trustworthy (24). Another study investigated the association between transport mode and physical activity using a linear mixed model (7); however, trip level data but not trip stage level data were considered and temporal autocorrelation was not taken into account, which is important when analyzing repeated observations (25). To overcome these limitations, we collected trip data at the trip stage level, and timestamps were available for all transitions between modes within trips over 7 days, and had been pre-identified with algorithms and then verified on the phone with participants.

Regarding limitations, first, the recruitment of participants was not at random (convenience sample). Beyond non-randonmess, findings from a small sample of 155 participants cannot be generalized to the complex transport habits of a population of more than 24 million inhabitants (Paris and close and far suburbs). For instance, if the odds of participating in the study were lower for those public transport users living in municipalities far from recruitment area, then longer public transport trip stages would be underrepresented in the study. Since a larger segment of a public transport trip is related to walking when the trip is short than when the trip is long, such a hypothetical recruitment bias would influence the comparison of physical activity between private motorized vehicle trips and public transport trips. Second, the estimated time of physical behaviors assigned to transport modes was based on the accelerometer wear time. If specific trips in terms of physical behaviors were more frequently excluded due to nonwear of the accelerometer, then it would bias our comparisons.

Interpretation of findings

Trips and trip stages by walking or other active modes, but also (although to a lesser extent) with public transport, were associated with longer walking durations and shorter sitting durations than trips based on a personal motorized vehicle, and these findings hold whether sitting or MVPA time were standardized or not by trip or trip stage durations. This finding supports previous studies quantifying the physical activity gains of biking (26,27) and walking (7). Regarding public transport, our findings are in accordance with previous research; for example, it has been found that public transport users had 24.3 minutes of physical activity per day while travelling, which is a substantial portion of recommended physical activity levels in
guidelines (28). The health benefits gained from the physical activity associated with the use of public transport have been investigated in previous literature (23).

In our study, in models with standardized outcomes, the coefficient showing that there was less sitting time in public transport and the coefficient showing more minutes of MVPA with public transport were stronger in the models at the trip level than at the trip stage level. This is because, in addition to the potential active movements within public transport vehicles, trips also typically include walked trip stages to and from public transport stations (7). Thus our study comparing analyses at the trip level and trip stage level was useful to distinguish between these two sources of physical activity. Walked distances to and from public transport stations may thus help people achieve physical activity recommendations, especially people who do not have time for other kinds of physical activity (23,28). However, it is critical to keep in mind that it may not be possible for everyone to increase their level of physical activity by transport mode, due to various types of health, environmental, or time constraints. It should also be emphasized that the physical activity gains from choosing public transport instead of a private motorized vehicle as a transport mode is likely to differ from one city to the other, because of variations in the configuration of transport systems and travel habits of people.

Conclusion

In conclusion, our study is the first to assess the relationship between various transport modes and physical behaviors based on GPS, mobility survey, and waist, thigh, and trunk accelerometer data, with a comparative analysis at the trip stage and trip levels. This pioneering approach allowed us to accurately measure differences in physical behaviors between transport modes.

Even if future research will have to rely on larger and more representative study samples to yield more generalizable findings, our study shows that promoting walking and biking but also public transport in daily routines may have a significant impact at the population level in terms of increasing the share of people reaching the physical activity recommandation.

final disclosure: Authors have no conflict of interest to disclose and there were no source of funding for this research.

1

Tables

Table 1. Association between transport mode used and physical behaviors (trip stage level n =7692, trip level n = 4683, N = 155 participants)[a], unstandardized outcome

	Sitting duration in minutes		MVPA duration in minutes	
Transport mode	**Trip stage level β (95%CI)**	**Trip level β (95%CI)**	**Trip stage level β (95%CI)**	**Trip level β (95%CI)**
Detailed classification				
Private motorized (driver)	Ref.	Ref.	Ref.	Ref.
Private motorized (passenger)	-2.19 (-3.70, -0.68)	-1.26 (-4.19, 1.67)	-1.25 (-2.06, -0.44)	1.39 (-0.01, 2.78)
Bus/coach	-10.69 (-11.8, -9.58)	-6.13 (-9.62, -2.64)	1.29 (0.76, 1.82)	7.88 (6.17, 9.60)
Metro	-10.55 (-11.51, -9.59)	-3.73 (-6.28, -1.18)	0.93 (0.46, 1.41)	10.92 (9.68, 12.16)
Tram	-10.61 (-12.63, -8.59)	-10.1 (-17.12, -3.08)	-0.09 (-1.03, 0.84)	8.48 (5.09, 11.87)
Suburban train	-5.99 (-7.19, -4.79)	-2.38 (-6.55, 1.79)	-0.21 (-0.81, 0.38)	17.15 (15.18, 19.13)
Biking and other active	-13.66 (-15.16, -12.16)	-13.26 (-16.46, -10.06)	5.58 (4.81, 6.36)	6.96 (5.48, 8.44)
Entirely walking	-14.04 (-14.74, -13.34)	-15.26 (-16.8, -13.72)	2.04 (1.67, 2.40)	4.39 (3.67, 5.12)
Multi-mode	NA	12.42 (9.92, 14.92)	NA	15.31 (14.09, 16.53)
Other[b]	3.13 (0.48, 5.78)	1.26 (-4.06, 6.58)	2.52 (0.91, 4.13)	3.00 (0.58, 5.41)
Crude classification				
Private motorized	Ref.	Ref.	Ref.	Ref.
Public transport	-9.27 (-10.06, -8.48)	-1.10 (-2.95, 0.75)	1.02 (0.62, 1.42)	11.54 (10.65, 12.43)
Other active mode	-13.36 (-14.85, -11.87)	-13.12 (-16.3, -9.94)	5.82 (5.06, 6.58)	6.80 (5.32, 8.28)
Entirely walking	-13.62 (-14.27, -12.97)	-15.29 (-16.74, -13.84)	2.27 (1.93, 2.61)	4.14 (3.45, 4.83)
Multi-mode	NA	24.58 (19.97, 29.19)	NA	20.90 (18.58, 23.22)
Other[b]	3.48 (0.84, 6.12)	1.23 (-4.1, 6.56)	2.74 (1.13, 4.35)	2.79 (0.36, 5.22)

CI: Confidence interval, MVPA: Moderate to vigorous physical activity, NA: Not applicable at the trip stage level.

[a]The multilevel linear models included a random effect at the individual level. The crude and the detailed transport mode variables were introduced in separate models. The models were adjusted for day of week and season, and took account of temporal autocorrelation.

[b]Long-distance train and plane.

Table 2. Association between transport mode used and physical behaviors (trip stage level n = 7692, trip level n = 4683, N = 155 participants)[a], standardized outcome

	Sitting time per 10 minutes of trip (minutes)		MVPA time per 10 minutes of trip (minutes)	
Transport mode	**Trip stage level**	**Trip level**	**Trip stage level**	**Trip level**
	β (95%CI)	**β (95%CI)**	**β (95%CI)**	**β (95%CI)**
Detailed classification				
Private motorized (driver)	Ref.	Ref.	Ref.	Ref.
Private motorized (passenger)	-0.59 (-1.03, -0.15)	-1.17 (-1.60, -0.74)	0.26 (-0.14, 0.67)	0.45 (0.08, 0.82)
Bus/coach	-2.79 (-3.20, -2.38)	-3.57 (-4.10, -3.04)	0.54 (0.15, 0.92)	1.88 (1.42, 2.33)
Metro	-2.88 (-3.20, -2.56)	-3.43 (-3.82, -3.04)	1.54 (1.25, 1.83)	2.95 (2.63, 3.27)
Tram	-3.37 (-4.16, -2.58)	-3.54 (-4.59, -2.49)	0.22 (-0.53, 0.96)	2.98 (2.06, 3.9)
Suburban train	-1.67 (-2.06, -1.28)	-3.83 (-4.44, -3.22)	0.54 (0.19, 0.9)	3.25 (2.72, 3.78)
Biking and other active	-5.58 (-6.04, -5.12)	-5.67 (-6.14, -5.20)	3.96 (3.53, 4.38)	4.04 (3.69, 4.4)
Entirely walking	-5.37 (-5.59, -5.15)	-5.40 (-5.63, -5.17)	4.78 (4.59, 4.98)	4.89 (4.71, 5.06)
Multi-mode	NA	-2.89 (-3.27, -2.51)	NA	1.73 (1.41, 2.06)
Other[b]	-2.00 (-2.67, -1.33)	-1.83 (-2.58, -1.08)	1.18 (0.57, 1.78)	1.08 (0.53, 1.63)
Crude classification				
Private motorized	Ref.	Ref.	Ref.	Ref.
Public transport	-2.46 (-2.71, -2.21)	-3.10 (-3.38, -2.82)	0.96 (0.73, 1.19)	2.44 (2.22, 2.66)
Other active mode	-5.49 (-5.95, -5.03)	-5.47 (-5.93, -5.01)	3.91 (3.49, 4.33)	3.99 (3.63, 4.35)
Entirely walking	-5.26 (-5.46, -5.06)	-5.17 (-5.39, -4.95)	4.73 (4.55, 4.91)	4.83 (4.66, 5.00)
Multi-mode	NA	-2.88 (-3.59, -2.17)	NA	1.25 (0.60, 1.90)
Other[b]	-1.89 (-2.55, -1.23)	-1.63 (-2.38, -0.88)	1.11 (0.51, 1.71)	1.02 (0.47, 1.57)

CI: Confidence interval, MVPA: Moderate to vigorous physical activity, NA: Not applicable at the trip stage level.

[a]The multilevel linear models included a random effect at the individual level. The crude and the detailed transport mode variables were introduced in separate models. The models were adjusted for day of week and season, and took account of temporal autocorrelation.

[b]Long-distance train and plane.

Table 3. Number of transitions between physical behaviors (SB, NSB, and MVPA) standardized by 10 minutes of trip, by transport mode, at the trip stage and trip levels: median (10th and 90th percentiles)

	Trip stage level			Trip level		
Transport mode	**SB and NSB transitions**	**NSB and MVPA transitions**	**SB and MVPA transitions**	**SB and NSB transitions**	**NSB and MVPA transitions**	**SB and MVPA transitions**
Detailed classification						
Private motorized (driver)	3.07 (0.00, 14.58)	0.61 (0.00, 6.41)	0.00 (0.00, 1.15)	2.95 (0.00, 13.73)	1.25 (0.00, 6.79)	0.00 (0.00, 1.16)
Private motorized (passenger)	3.25 (0.00, 14.42)	0.76 (0.00, 6.00)	0.00 (0.00, 0.00)	2.69 (0.00, 12.64)	1.24 (0.00, 8.09)	0.00 (0.00, 0.50)
Bus/coach	1.54 (0.00, 10.83)	1.35 (0.00, 7.72)	0.00 (0.00, 0.44)	1.23 (0.00, 6.23)	4.00 (1.38, 8.12)	0.00 (0.00, 0.45)
Metro	0.51 (0.00, 4.72)	2.73 (0.00, 12.61)	0.00 (0.00, 1.25)	0.81 (0.00, 3.13)	4.73 (2.23, 9.23)	0.00 (0.00, 0.78)
Tram	0.00 (0.00, 4.03)	1.27 (0.00, 4.4)	0.00 (0.00, 0.00)	0.45 (0.00, 2.60)	3.73 (0.00, 7.51)	0.00 (0.00, 0.33)
Suburban train	0.53 (0.00, 3.90)	1.69 (0.00, 7.85)	0.00 (0.00, 0.83)	0.64 (0.00, 2.21)	4.84 (2.04, 8.48)	0.00 (0.00, 0.71)
Biking and other active	1.48 (0.00, 9.99)	8.00 (1.41, 16.47)	0.30 (0.00, 5.33)	1.30 (0.00, 8.48)	7.96 (1.95, 16.15)	0.45 (0.00, 4.80)
Entirely walking	0.00 (0.00, 3.33)	5.03 (0.00, 16.56)	0.00 (0.00, 0.00)	0.00 (0.00, 2.88)	5.28 (0.00, 15.32)	0.00 (0.00, 0.00)
Multi-mode	NA	NA	NA	1.15 (0.05, 4.27)	4.00 (1.99, 7.02)	0.21 (0.00, 0.81)
Other[a]	1.20 (0.00, 5.99)	1.25 (0.00, 14.08)	0.00 (0.00, 2.01)	1.23 (0.00, 6.99)	5.08 (0.00, 16.52)	0.32 (0.00, 3.44)
Kruskal wallis test (p-value)	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
Crude classification						
Private motorized	3.07 (0.00, 14.48)	0.64 (0.00, 6.38)	0.00 (0.00, 1.07)	2.86 (0.00, 13.57)	1.24 (0.00, 6.96)	0.00 (0.00, 1.08)
Public transport	0.53 (0.00, 6.32)	2.13 (0.00, 9.96)	0.00 (0.00, 0.95)	0.93 (0.00, 3.81)	4.42 (2.08, 8.32)	0.00 (0.00, 0.69)
Other active mode	1.48 (0.00, 9.99)	8.00 (1.41, 16.47)	0.3 (0.00, 5.33)	1.30 (0.00, 8.48)	7.96 (1.95, 16.15)	0.45 (0.00, 4.80)
Entirely walking	0.00 (0.00, 3.33)	5.03 (0.00, 16.56)	0.00 (0.00, 0.00)	0.00 (0.00, 2.88)	5.28 (0.00, 15.32)	0.00 (0.00, 0.00)
Multi-mode	NA	NA	NA	1.75 (0.43, 10.1)	3.62 (1.63, 7.47)	0.15 (0.00, 1.32)
Other[a]	1.20 (0.00, 5.99)	1.25 (0.00, 14.08)	0.00 (0.00, 2.01)	1.23 (0.00, 6.99)	5.08 (0.00, 16.52)	0.32 (0.00, 3.44)
Kruskal wallis test (p-value)	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001

RECORD MultiSensor Study, 155 participants, 7692 trip stages and 4683 trips

SB: Sedentary behaviour, NSB: Non-sedentary behaviour, MVPA: Moderate to vigorous physical activity, NA: Not applicable at the trip stage level.

[a]Long distance train and plane.

Table 4. Association between transport mode and sitting duration, comparing a waist-worn accelerometer with two thigh- and trunk-worn accelerometers (analyzed at the trip stage level, n = 6901, N = 154 participants)[a]

Transport mode	**Sitting duration, waist-worn accelerometer (minutes) β (95%CI)**	**Sitting duration, thigh- and trunk-worn accelerometers (minutes) β (95%CI)**	**Sitting duration per 10 minutes of trip stage, waist-worn accelerometer (minutes) β (95%CI)**	**Sitting duration per 10 minutes of trip stage, thigh- and trunk-worn accelerometers (minutes) β (95%CI)**
Detailed classification				
Private motorized (driver)	Ref.	Ref.	Ref.	Ref.
Private motorized (passenger)	-1.45 (-2.83, -0.07)	2.58 (1.26, 3.90)	-0.42 (-0.89, 0.05)	0.89 (0.41, 1.37)
Bus/coach	-10.27 (-11.27, -9.27)	-5.79 (-6.8, -4.78)	-3.12 (-3.54, -2.70)	-0.49 (-0.92, -0.06)
Metro	-9.96 (-10.82, -9.10)	-7.2 (-8.05, -6.35)	-3.13 (-3.45, -2.81)	-1.44 (-1.77, -1.11)
Tram	-10.13 (-11.93, -8.33)	-5.80 (-7.65, -3.95)	-3.59 (-4.38, -2.80)	-0.33 (-1.14, 0.48)
Suburban train	-6.18 (-7.27, -5.09)	-3.7 (-4.78, -2.62)	-2.10 (-2.50, -1.70)	-0.50 (-0.91, -0.09)
Biking and other active	-13.73 (-15.05, -12.41)	-5.9 (-7.18, -4.62)	-6 (-6.48, -5.52)	-1.09 (-1.58, -0.60)
Entirely walking	-13.67 (-14.31, -13.03)	-10.88 (-11.49, -10.27)	-5.67 (-5.9, -5.44)	-4.06 (-4.29, -3.83)
Other[b]	15.69 (12.34, 19.04)	5.89 (2.86, 8.92)	1.3 (0.31, 2.29)	-2.96 (-3.96, -1.96)
Crude classification				
Private motorized	Ref.	Ref.	Ref.	Ref.
Public transport	-9.04 (-9.75, -8.33)	-6.49 (-7.18, -5.80)	-2.82 (-3.08, -2.56)	-1.12 (-1.39, -0.85)
Other active mode	-13.54 (-14.85, -12.23)	-6.35 (-7.62, -5.08)	-5.95 (-6.42, -5.48)	-1.24 (-1.72, -0.76)
Entirely walking	-13.39 (-13.98, -12.8)	-11.31 (-11.88, -10.74)	-5.59 (-5.80, -5.38)	-4.22 (-4.43, -4.01)
Other[b]	15.89 (12.55, 19.23)	5.50 (2.46, 8.54)	1.37 (0.39, 2.35)	-3.10 (-4.10, -2.10)

CI: Confidence interval.

[a]The multilevel linear models included a random effect at the individual level. The crude and the detailed transport mode variables were introduced in separate models. The models were adjusted for day of week and season, and took account of temporal autocorrelation.

[b]Long-distance train and plane.

DISCUSSION

Physical Activity: New paradigms, new methods

Over the last centuries, physical inactivity in the population has progressed hand in hand with technological development. Progresses in automation and the growth of the third activity sector have reduced overall levels of occupational physical activity [5]. Meanwhile, leisure became, to a large extent, screen-based and massive motorization has further reduced physical activity associated with mobility [2, 6, 7]. These trends are no longer restricted to the technologically developed Western world: with the rapid development of rising economies, they have been observed in increasingly larger fractions of the World's population [6]. Thus, physical inactivity is now viewed as a galloping pandemic, which requires appropriate public health action by the policy makers at the local and global level.

From the mid-20th century to these days, considerable advances were made in assessing and understanding the health hazards associated with inactivity. Recent paradigmatic shifts have heightened awareness to the fact that several of these hazards may be reduced by avoiding sedentary behaviors as much as possible and cumulating any short sequence of activity all along the day [9]. To the extent that this new paradigm is supported by empirical evidence, small behavioral changes can be undertaken all along the day and make a considerable difference. The determinants of physical activity need to be understood as any factor provoking even a small burst of activity or a change of posture in any context of the daily routine. Public health decisions must therefore focus not only on promotion of structured sessions of exercise, but on the multitude of factors that can impact physical behaviors in everyday life.

Studying physical activity within this new paradigm requires new scientific tools and methods. In population studies, monitoring of physical activity is to be implemented using precise tools capable of measuring a large array of behaviors at a high time resolution. Likewise, upstream research about the environmental determinants of physical behaviors must consider continuous momentary exposure to environmental stimuli and link them to contemporaneous physical behaviors. As data collected becomes more complex, methods of analysis must evolve to treat them appropriately.

This implies new methods for deriving physical behavior from motion sensors, methods for capturing the complex patterns of activity segmentation and methods for treating behavioral time-budgets with multiple components. At the determinants level, this new complexity implies merging various geographical, spatial and temporal layers in a meaningful way.

This thesis aimed to contribute to a better understanding of potential determinants of physical behaviors and associated health consequences within this new paradigm. We relied chiefly on the RECORD study cohort, a notable attempt to observe the detailed activity and mobility patterns of an urban French population of adults in their daily life. We have thus reached several new insights, which will be discussed in the three following sections, each devoted to one of the three main objectives of this thesis: (i) understanding the relationship between physical activity and health in order to help refine future physical activity recommendations, (ii) identifying ways to promote physical activity in the population, especially within the context of urban environment, (iii) improving measurement of physical activity in free-living conditions using accelerometers. This discussion concludes with the key messages and directions for future research.

Physical activity and health

We started by investigating, in **chapters I and II**, relationships between advanced measures of physical behaviors and health outcomes. The main innovations of our studies were the use of a relatively detailed nomenclatures of behaviors, the use of compositional methods, and a systematic approach to behavior segmentation.

The importance of posture allocation The first important result in the study presented in **chapter I** was the associations between standing volume and lipid profile. Standing volume positively associated with plasma HDL and negatively with triglycerides level. With the compositional models, we were able to predict health outcomes for any time-budget. Thus, we showed that budgets with sufficient time reallocation from sitting to standing (for instance comparing occupations requiring standing vs. sitting) could offset the detrimental associations observed in the absence of MVPA.

Previous research summarized by Miles-Chan and Dulloo showed that quiet standing alone, in average, does not induce a considerable change in energy expenditure [24]. Do these associations between standing and lipid profile point to physiological mechanisms that influence metabolic health irrespective of energy expenditure? These authors suggest that this assumption is not necessary, given that standing time could be only a proxy for the number of sit-stand transitions, which increases energy expenditure overtime. Our data do not support this assumption because the relationship observed between standing and lipid variables persisted even when we controlled for segmentation patterns at resolution of one second, i.e. even when transitions were accounted for.

Another possible explanation to the observed health consequences of standing that excludes the intervention of mechanisms that are independent from energy expenditure consists in assuming that the time recorded by the accelerometers as quiet standing actually comprises hidden energy expenditure. This can be many small vibrations of the body, frequent shifting of the body weight between the feet, etc., which can add up to an increased energy expenditure over the total standing time. An inspection of the accelerometer records of our patients made this explanation implausible, as the motion intensity, if we view it as a proxy for energy expenditure, even when cumulated over relatively long periods of time, remains negligible compared to MVPA. Nevertheless, this open question clearly points to the fact that the field is missing a precise, universal system of behavior classification addressing such questions. **Chapter III** proposes a new classification model of physical behaviors based on accelerometers; although it does not address the issue of very light activity while standing still, the model shows that using advanced accelerometers and processing methods, extremely subtle behaviors could be identified at a high degree of accuracy.

Our findings about the importance of standing volumes corroborates previous research by Healy and colleagues on Australian individuals in free-living conditions, and several other studies using an experimental design [23]. However, while Healy and colleagues used iso-temporal substitution to model the effects of replacing sitting with standing, we complemented our analysis with compositional models. These have been advocated by mathematicians [31] and have the advantage of considering the components of the compositions as a whole, allowing convenient predictions for different budgets that differ from each other with regard to several components. In addition, by controlling for segmentation at a high time resolution, our estimates reflect the “pure” associations of volume, irrespective of segmentation patterns.

Sitting and lying are two behaviors that are typically classified in a single category of sedentary behavior in research. Our study distinguished between them and revealed a strong correlation with adiposity. Thus, we observed higher BMI and waist circumference values when sitting was replaced with lying. Here, having a pair of accelerometers placed on the subjects’ chest and thigh proved useful, since the former helped determine the inclination degree of the trunk. Evidence for associations between postural allocation and health outcomes based on a sensor placed on the trunk is very scarce, and, to my knowledge, the literature has not thoroughly discussed similar results.

Behavior segmentation and health

Controlling for segmentation patterns proved useful in better understanding the correlations of health outcomes with *volumes*. However, segmentation patterns themselves also exhibited associations with health outcomes. Our study presented in **chapter I** run segmentation analysis on two types of behavior time-series. In one time-series, we applied a bout detection algorithm at a 5-minute resolution to identify homogenized prolonged behavior periods while ignoring short interruptions. In another time-series, we kept the time-series unchanged at a 1-second resolution, in order to keep track of any micro-sequence of activity and interruption. Bout segmentation analysis yielded inconclusive results and could not be linked to previous literature. As I explain in the **Annex**, although reducing time resolution of the behavior series by using bouts follows a justified scientific rationale, its use in the literature is highly problematic, as results greatly vary depending on preliminary data processing and the parameters of the bout detection algorithm applied. Likewise, our study presented in **chapter II**, comparing two conditions in which the *same* activity volume was performed in one 45-minute bout versus nine 5-minute bouts, could not detect any significant difference in insulin and glucose level.

Nevertheless, in **chapter I,** applying segmentation analysis on the unchanged behavior time series (i.e. at a 1-second time resolution) offered a valuable insight. Irrespective of volume, a high degree of partitioning of the total sedentary time associated with reduced fasting plasma glucose level. This result is in line with other studies that observed a reduction of postprandial glycaemia following interruptions of sedentary behavior with standing or physical activity episodes [23], but it is new as far as fasting glycaemia is concerned. Moreover, whereas the major studies that identified associations or effects between segmentation of sedentary time and glycaemia considered only relatively long behavior bouts (1 minute or more), our study arrived to this result by considering the distribution of sequences of any duration, thus pointing to the relevance of segmentation at a high time resolution (i.e. including micro-sequences and micro-breaks).

Whether there is an additional value to MVPA time when it is performed in long bouts (e.g. 10 or 20 minutes) is still a matter of debate in the literature [34]. Our results could not highlight any evidence for the alleged negative associations of MVPA segmentation with the health outcomes examined. Our findings are even more relevant considering that our approach rigorously separates volume and segmentation, and studies segmentation pattern both at different time resolutions. To this extent, our results are in line with the cautious position adopted by the American Physical Activity Guidelines (2018), which removed the recommendation to perform MPVA in bouts.

Limitations

The main shortcoming of the study presented in chapter I is its cross-sectional design and our inability to draw conclusions about the causal link between physical behavior and health. Despite the keen interest of epidemiological research in the effect of various aspects of physical behavior on health, causal evidence gathered in free-living conditions is very scarce due to the difficulty in implementing longitudinal studies. The causality of the observed relationship between lying and adiposity is particularly questionable, since being prone to lying can be as well the consequence of body composition. Nevertheless, in this context, it is worth mentioning the small-scale study by Levine and colleagues [21], which showed that time spend sitting and standing varied across lean and obese groups both before and after supervised weight gain, suggesting that a preference for a certain posture could be biologically determined and rather the cause than the consequence of adiposity. This causal direction, however, needs to be confirmed by studies in larger sample sizes.

Causality can be determined, of course, with strictly controlled experimental designs or longitudinal studies. Yet, they do not necessarily reflect the 'chaotic' character of unrestrained, free-living conditions. To this extent, longitudinal studies with activity monitoring, despite the challenge that they represent, are still necessary.

What can impact individuals' physical behaviors?

The relationships between aspects of the urban environment and physical activity have been thoroughly investigated in past literature [91, 97-100]. Our studies addressed these questions with a new, rigorous approach, aiming at causal inference. Relying on an analytical framework merging geographical information, personal logbooks, activity and location data, we were able to isolate spatio-temporal life-segments in individuals' daily life that can be viewed as virtually free of the instantaneous selection bias. Thus, in **chapter IV**, we studied physical activity during commuting as function of the urban environment along the routes: as commuting is a necessary trip, it is very unlikely to suffer from biases compromising causal inference. Likewise, **chapter V** isolated all daily trips to study their distribution of physical behaviors. Thus, studies in this thesis benefited from the advantages of a high-resolution approach linking between momentary exposure and contemporaneous activity, while circumventing the self-selection pitfall.

Effects of urban attributes on physical activity

Although limited in our frame of analysis, our approach yielded important results. A home-work route passing through greener or high-education areas positively affected the activity level performed during commutes and reduced sedentary time. Likewise, when this route was well connected by the public transportation network, overall activity increased, and sedentary time decreased. These results are in line with some previous findings [93, 97, 99, 100, 102, 103], but our innovative design yield a unique insight into the mechanisms linking environment and activity. Previous research could usually not tell if one was active because of being in an environment favorable to activity or if one was in an environment because the wish to be active.

The effect of greenery on activity level shows that an aesthetically pleasant environment can promote commutes by including more stages of active travelling. However, the interpretation of our findings regarding the effects of mean areal education on activity should probably not be taken at face value. Rather, they should be understood as a proxy for a variety of activity-generating environmental characteristics that are not included in the models. These results remain important as they that low socio-economic areas suffer from a disadvantage in terms of activity-generating attributes. This disadvantage can be related to various attributes of the environment, such as different aesthetical aspects, pedestrian and cycling infrastructure, atmospheric or noise pollution, etc. Thus, when planning interventions aiming to create environments that are more friendly to active commuting, decision-makers should ensure a fair spatial distribution of such improvements.

Our result about the effect of a better public transportation on activity can be better understood with the results presented in **chapter V**. In this chapter, we accurately quantify the distribution of physical behaviors for all trips as a function of the transportation used. We show that travelling with public transports significantly reduces sedentary time and increases activity compared to the use of a private motorized vehicles. Our approach is interesting insofar as it distinguishes between a trip as a whole, which can include several stages and stations between them, and a trip stage, which is realized in a continuous way without interruption or change of transportation mode. Considering trips as a whole shows that apart from the reduction of sedentary time associated with the use of public transports, such trips often include high-activity stages, which makes travelling with public transports a factor able to significantly increase overall daily activity level. Taking altogether data presented in **chapter IV**, we can infer that an efficient network of public transports represents a real incentive to use them and through the incremented activity that it implies compared to private motorized vehicles, it results in an overall significant increase in activity.

Implementing physical activity programs in individuals' daily life

Last, an important action to ponder, although not central in this thesis, is activity programs implemented in individuals' daily life. The study presented in **chapter II** reports the results of a three arms crossover intervention study in which sedentary and inactive adults with overweight were asked to perform daily for 3 days either nine 5-minute brisk walking bouts every hour for 9 consecutive hours, or a single 45-minute brisk walking bout. In the third trial, they were asked to maintain their habitual sedentary lifestyle for three days (control condition).

Results showed that moderate-to-vigorous intensity activity was higher when participants asked to perform a single long bout than several short bouts of activity. However, adults with overweight self-reported to feel less tired at the end of the day when breaking up sedentary time with short, frequent bouts of activity than when being engaged in a long continuous bout of activity. This suggest that for sedentary, inactive people, implementing physical activity through short bouts of activity spread across the day may be a first step to bring movement back into their daily life, and prevent long sedentary periods, but that overall adherence levels are higher when subjects are solicited once (even for a longer period) over the day.

Thus, although breaking up sedentary time is thought to have a beneficial effect on health that is independent from the total level of activity over the day, implementing one single long bout turned out to be a better solution, as far as volume is concerned. The right trade-off between a smaller total volume that implies multiple breaks a larger volume involving only one break hinges on the question whether sufficient volumes of MVPA can offset the detrimental effects of sedentary behaviors, which has recently been discussed in the literature [124] (we could not prove it clearly in **chapter I** at this resolution). This trade-off could explain the fact that despite unequal total volumes of activity observed in the two conditions, effects on glucose and insulin were similar. Future research in physiology could further investigate this trade-off between segmentation and volume and new epidemiological studies are needed to corroborate our findings in the longer term.

Limitations

Regarding the effects of urban attributes on physical attributes that we observed, a caveat needs to be added. Our study in **chapter V** confirms the very intuitive conception that, regarding total physical activity levels over the journey, the three main travelling modes are ranked as follows: active means (including walking, bicycling etc.), public transports and private motorized vehicles. Yet, it follows from this ranking that as much as an efficient public transportation system can incite people to use public transports instead of cars, it can incite people to use public transports instead of active means, thus reducing overall activity level. Very recent events can testify to this phenomenon: following the strike in France during winter 2019-2020 and the drop in frequency of public transports, the use of active means (bicycles in particular) in Paris sharply rose [121], possibly resulting in an overall increase in activity. In addition, even for individuals that completely depend on public transportation, there is a threshold of improvement beyond which public transports are so efficient that physical activity levels during the journey starts *to decrease.* This is the case when time saved by improvement public transports is due to connections that are so close to the origin and the destination that the activity component in the time-budget of the journey starts shrinking. Thus, improvement of the public transportation system is beneficial only when it is an incentive to replace private motorized vehicles and when it reduces the travel time while maintaining the activity time over the travel duration constant. In practice, such a distinction is difficult to make in an empirical study and, in any case, it seems that the current condition of the transportation network is such that improved efficiency of public transports still generates a sizable increase in overall level of physical activity.

We should also limit ourselves by concluding that our results should be applied with caution to contexts other than the region of Paris in normal times. This limitation applies more generally to our other findings regarding urban attributes. Urban aspects greatly vary among each other, as do the cultures of the cities' residents. In addition, even within the region of Paris, our small sample size is not necessarily representative of all mobility patterns that could be observed in a population of several millions of inhabitants.

Regarding our result about implementation of activity programs, although adherence level was overall satisfactory, the validity of the results is limited by the short duration (only three days) of the experiment and by the fact subjects were monitored. From a psychological perspective, it can be argued that adherence to the instruction could drop over longer period, and that the fact that subjects knew that they were monitored up-biased estimated adherence levels without surveillance.

Deriving physical behaviors from accelerometer data

Accelerometers are a widely used tool for physical activity monitoring in observational studies. It was used in all studies included in this thesis. As a central issue in physical activity epidemiology, **chapter III** presents a new algorithm for derivation of physical behaviors from raw data from body-mounted accelerometers.

Recent developments

Proprietary algorithms have played a central role in defining the key notions that would subsequently be used in the scientific discourse. Most notably, the very popular Actigraph accelerometers has introduced the notions of *count* and *bout* that would become determining in physical activity epidemiology, as explained in the **Annex**. Nevertheless, the literature highlighted the opaqueness of these algorithms, and the difficulties encountered when trying to compare results across devices and algorithms. In addition, as research evolves, new behaviors need to be defined and detected. As access to the devices' raw acceleration data is becoming standard, research has oriented itself towards new, transparent behavior detection algorithms at a high time resolution.

A new algorithm

Chapter III presents a new classification algorithm of 13 detailed daily behaviors defined by the *DaLiAc* public dataset. It outperformed other algorithms tested against this dataset despite its speed and simplicity. We have shown that a careful feature extraction process combined with a hierarchical classification system based on an understanding of the tasks at hand can prove more useful than heavy, hard-to-tune models.

Our algorithm relies on logistic regression models and a signal collected from four accelerometers placed at the chest, hip, wrist and ankle. Unlike previous evidence that showed that the optimal model can vary depending on the position of the device, our results show that, for all practical purposes, logistic regression was optimal or nearoptimal across all combinations of device positions and all classification sub-tasks compared to other models. In addition, our results suggest that adding gyroscopes to accelerometers yielded only a marginal improvement.

Thus, although the algorithm proposed in **chapter III** focused on the classification of a specific set of behaviors, the plasticity of the system employed is promising for other applications. For instance, future research questions might raise the need to investigate new behavior categories. For instance, algorithms could be trained to detect fidgeting, which is usually not captured in the current frameworks of activity detection, but can be of physiological significance. More generally, using linear regression, our model could be trained to estimate continuous energy expenditure, which was not addressed directly in this chapter. This limitation is discussed in the next section.

Limitations

Our algorithm classifies signal samples by assigning them to qualitative categories (walking, bicycling, lying, sitting etc.). Energy expenditure can thus be estimated by assigning a factor (i.e. number of METs) to each behavior. This approach contrasts with traditional approaches widely used in physical activity epidemiology, which focus on motion intensity as a proxy to energy expenditure and mainly on motion intensity [123]. By looking both at intensity and detecting complex patterns of the signal, our approach has the advantage of being able to detect different behaviors corresponding to different energy expenditures, but whose motion intensity was nearly similar: for instance, cycling at the same pace but with different resistance levels produces practically the same motion intensity, but some nearly imperceptible swinging of the of the body weight allow advanced algorithms to discriminate these activities with an accuracy of about 95%. However, energy expenditure can sometimes vary within a single category, such as ‘walking’. This category, which is detected based on a specific pattern of the signal, can comprise a continuum of energy expenditure levels, which are not captured using categorical classification.

Thus, an additional step in the development of the discipline might consist in applying models that fully exploit the power of raw accelerometer signal, such as the one that we present here, to accurately estimate continuous energy expenditure along with categorical behaviors. This task, compared to the one we addressed here, is very likely less complex. The main challenge remains to create a public, labeled raw acceleration dataset with high variability in energy expenditure, which will make training of machine learning algorithms possible.

Likewise, models should ideally be trained with and tested against datasets that simulate real-life behaviors. Algorithms are typically developed to optimize accuracy and computational time in classification of data representing 'clean' behaviors. However, their external validity remains limited due to the complex, chaotic character of real-life conditions. In these conditions, transitions between behaviors are swift and behaviors themselves are performed at various levels of intensity.

Conclusions

Refining recommendations for activity guidelines

Increasing the standing component at the expense of sedentary postures in the time-budget is a strong correlate of healthy lipid profile. Likewise, large volumes of lying are associated with increased risks of adiposity. Regarding reducing some health hazards, there seems to be a gradient lying-sitting-standing, pointing to a physiology that is independent from energy expenditure.

Using bouts, our studies could not highlight any clear pattern linking segmentation of any behavior to the health outcomes under study, however, taking micro-sequences of a few seconds into account, we could see that segmented sedentary volume is associated with lower fasting glycaemia. From a public health perspective, the importance of micro-interruptions of sedentary time should therefore be acknowledged and substitution of sitting with standing time encouraged (e.g. sit-stand desks), especially in office desk workers who are highly vulnerable to the adverse health effects of sedentary behaviors.

Future research should strive to a consensual behavior classification encompassing detailed postures and activity categories. In particular, the classification should stem from the big research questions in the field, and not the opposite. A deeper investigation of segmentation patterns and their effect on health should be undertaken, considering segmentation at a high time resolution and letting go of preconceptions inherited from past methodology. Models should treat behavior volumes as compositions and include both metrics of volume and segmentation in order to assess their independent effects. Last, large sample sizes and longitudinal designs remain paramount to verify hypotheses with a strong statistical power and detect causal links.

How can we encourage physical activity and reduce sedentary behavior in the population?

Greenery level and better efficiency were shown to reduce sedentary time and increase physical activity during travel. Considering the important time fraction spent travelling over the day, policy makers should therefore consider greening interventions and improvement of public transportation system as a way to help people achieve the recommended activity guidelines. Special attention should be paid to a fair distribution of activity-generating attributes of the urban environment over the city, as we observed spatial disparities in actual physical activity across areas of different socioeconomic levels.

Future research should favor contemporaneous and objective study designs allowing causal inference, such as the one proposed here (i.e. looking at life-segments where selection biases are not likely to occur), while investigating new environmental attributes in a variety of societies and urban contexts.

Regarding implementation of physical activity programs in individuals' everyday life, our results suggest that it can increase overall activity level, and that, levels of compliance are higher when individuals are asked to perform one long bout than when asked to perform several smaller bouts.

Future research should study implementation programs over longer periods and further investigate the trade-off between segmentation and volume from a theoretical and practical perspective.

Improving measurement of physical behaviors

Physical behavior monitoring should take advantage of raw data and advances in machine learning to derive wider spectra of behaviors, in accordance with the development of physical activity epidemiology, and to ensure better comparability and transparence. We proposed a simple and adaptable algorithm based on logistic regression, domain knowledge and a good understanding of the classification tasks at hand. This algorithm outperforms heavy and complex algorithms developed in the past.

Future research needs to focus on creating more realistic datasets and train models that are more adapted to these data. Recent advances in behavior detection algorithms could also be leveraged to improve prediction of energy expenditure using various, high-resolution public datasets.

Final words

Around 1990, Mark Weiser formulated envisioned a future in which *ubiquitous computing* would transform people's everyday reality [122]. Laptops, tablets, smartphones, sensors and smart home appliances have indeed invaded our existences ever since, gathering information that are analyzed and exploited with increasingly performant technologies. While this technological revolution is making our life easier, it is at the same time, part of the Physical Activity Transition, during which physical activity is gradually engineered out of humans' life. Never in human history has everyday life demanded so little physical activity from us, and, despite improvements in health and life expectancy, never has the health burden associated with our sedentary lifestyle been so tangible as today.

My thesis aimed at leveraging tools from this same ubiquitous computing revolution to contribute to the fight the pandemic physical inactivity. By means of wearable sensors and big data analytics, I have tried, together with my collaborators, to improve our state of knowledge on the causes and effects of physical inactivity. I hope that my modest contribution will help decision makers take actions that will make people's life healthier and happier.

ANNEX

Reflections on bouts in physical activity research, their definition and parameters

A fundamental concept in physical activity research

Research on physical activity and its relationship with health traditionally distinguishes between activity *volume*, i.e. the total time devoted to an activity over the study time, and *accumulation patterns*, i.e. how activity episodes of different durations add up to the total activity time. Traditional activity guidelines recommended accumulating a certain volume of moderate-to-physical activity (MVPA) in 'bouts' longer than 10, 20 or 30 minutes [20, 32]. Whether these recommendations were justified is a matter of debate in the scientific community. Some studies suggested that moments of MVPA had an additional beneficial value for metabolic health when performed within bouts of 10 minutes or longer [32, 35, 36] while others could not find evidence supporting this hypothesis [34, 37, 38, 116]. Nevertheless, the 2018 American Physical Activity Guidelines did not include any recommendation regarding accumulation patterns, stating only that a total of 150-300 minutes per of MVPA should be reached [9]. In contrast, accumulating sedentary time in long bouts was shown to have adverse effects on postprandial glucose and insulin levels [41–43, 117] and to positively correlate with fasting plasma triglycerides and adiposity measures [118].

As we see, a *bout* is a widely used, fundamental notion in considering accumulation patterns. When looking at bouts of activity over the monitoring period under consideration, we do not merely sum activity time, but examine whether this time is accumulated in more or less continuous episodes rather than in brief sporadic bursts of activity. Nevertheless, activity bouts are not simply continuous episodes of activities. In fact, it is difficult to find long sequences of activity without the slightest interruption. Such short interruptions might conceal interesting information, but we might as well ignore them in order to look at the larger picture. Although bouts were not intentionally developed to have this essential research aspect, their centrality in analysis of activity accumulation is probably due to their ability to show to what extend the volume is accumulated in long, significant periods of activity, *while discarding insignificant interruptions*. Thus, although there exists no consensual definition, we can define a bout of activity x very broadly, based on the use of this term in previous work, as a period of a minimum duration during which the observed activity is *predominantly* x. We say predominantly, because all definitions of bout agree that some sort of smoothing of the activity recorded is necessary in order to discard negligible interruptions.

An approach based on Actigraph's count number

Bouts were a straightforward concept in the pioneering research that emerged with the 2003-2006 NHANES study cohort. In this cohort, activity was assessed by means of an uni-axial Actigraph accelerometer, and motion intensity was recorded as number of "counts" per 60-second epochs [35]. What counts exactly are and mean is unclear and its derivation remains proprietary of the accelerometer manufacturer, but they can safely be understood as a measure of motion intensity over the epoch considered [67]. Within this technical frame, a behavior bout of n epochs was simply n consecutive epochs for which the number of counts recorded was above (when looking at moderate-to-vigorous activity) or below (looking at resting or sedentary behavior) a cut-point. It is important to emphasize that count number per 60-second epoch, as an aggregate of activity over time, was a metric that *already* included a certain degree of smoothing. For instance, in most situation a sequence of 25 seconds walking, 10 seconds resting and 25 seconds walking again would have had resulted in a count number per epoch that is above the threshold, and the entire epoch been viewed as one unit of activity.

A further level of smoothing at *the epoch level* was also possible. When looking for 10-minute or 20-minute bouts, some authors allowed for interruptions of one or a few epochs, while other recommended granting no 'grace period' [34, 46]. But, as we said, the aggregating of the continuous activity into epochs represented per se a first level of smoothing. At this point, it is important to point out that although the first NHANES study specified an epoch length of 60 seconds, Actigraph epochs can be set to various lengths [35]. Consequently, the debate on whether grace periods at the epoch level should be allowed must account for the preliminary smoothing resulting from the discretization of continuous into epochs of various lengths. For instance, setting the epoch length to five seconds without allowing any grace period at the epoch level makes it very difficult to aggregate consecutive activity epochs into bouts, while it is much easier with 1- or 5-minute epochs.

Bouts with the new generation of accelerometer monitoring: a general definition

As accelerometer-based monitoring became more and more popular in research, comparability across devices and specifications (epoch length, cut-points, count derivation) became difficult and considerably hindered building up coherent evidence about health effects of different accumulation patterns [34]. With newer accelerometer devices allowing retrieval of raw acceleration data, derivation of activity for very short time units (e.g. 1 second) became possible [61]. However, derivation of activity at a high temporal resolution came at a cost: in most cases, the smoothing that used to be performed in the Actigraph activity derivation through the aggregation over epochs no longer existed. Hence, bout detection needed to be redefined.

I start by presenting a broad definition, which I believe would be accepted by a broad majority of authors, and to which I refer as the simple definition. In order to define it in precise terms, some formalization is needed. Say we have a time-series $X = \{x_t : t \in T\}$, where, for a time unit t, x takes a certain categorical value of a physical activity a from a set A containing all physical activities studied. Any series can be run-length encoded (rle) into in a series of tuples containing runs (sequences of the same values) l and their corresponding values a, $\{(l, a)_i\}$. For example, applying a run-length encode function to a time series X = {1,1,1,3,3,2,3,3}, we obtain $rle\,(X) = \Psi_x = \{(3,1),(2,3),(1,2),(2,3)\}$. We can further define a function that extracts from the run-length encoded time-series only those run lengths corresponding to a certain activity a, $\Psi_x(a)$. In our example, $\Psi_x(1) = \{3\}$, $\Psi_x(2) = \{1\}$, $\Psi_x(3) = \{2,2\}$ and $\Psi_x(4) = \{\}$.

When we are interested in activity detection for a particular activity, we binarize the time series X with regard to the activity of interest a, so that $x_t = 1$ when the value a is observed and $x_t = 0$ otherwise. In its broadest sense, a *bout of activity a* can be any series Y over a time span T' extracted from a binary behavior time-series X, $Y = \{y_t; t \in T'\}$, that fulfills the following conditions:

I. The length of the bout in epochs is greater than a certain *minimum bout length* m.

$$|Y| > m$$

II. Y does not contain any sequence of a behavior other than that of interest that is longer than a certain tolerance time, expressed as a fraction β of the bout length m. For $rle(Y) = \Psi_Y$:

$$\max(\Psi_Y(0)) < \beta m$$

III. The sum of elements taking the behavior of interest is longer than a 'purity' threshold α, expressed as a fraction of the bout length:

$$\sum_{t \in T} y_t > \alpha m$$

IV. The first and last element of Y must take the value 1.

Bouts with activpalProcessing

Of particular interest is the definition of bout as it appears from the bout detection algorithm of the R-package *acitvpalProcessing* by Lyden and Staudenmayer [119]. Recent ActivePAL accelerometers output detailed activity records for very short time units (0.1 seconds), and have become increasingly popular [120]. The main innovation in this algorithm consists in the fact that *the tolerance time is variable* and depends on the neighboring sequences. To detect bouts, we take the binarized series with regard to the activity of interest, X. A bout is any time-series extracted from X, $Y = \{y_t; t \in T'\}$, for which:

I. Any t in Y verifies:

$$\frac{1}{\min(m, t)} \sum_{\mathrm{i} = \min(1, t-m)}^{t} y_i > \alpha m$$

II. All tuples $(l, 1)_i$ followed by $(l, 0)_{i+1}$ in $\Psi_Y(0)$ verify:

$$\frac{\min(l_i - l_{i+1}, m - l_{i+1})}{\min(l_i + l_{i+1}, m)} > \alpha m$$

III. The length of the bout in epochs is greater than a certain *minimum bout length* m.

$$|Y| > m$$

Conditions are to be verified also for the time-series indexed in a reverse order, as short interruptions should be agglutinated to the bout not only when *followed* by long activity sequences, but also when *preceding* them.

To summarize this definition in words, sequences of behaviors other than that of interest (null sequences) are agglutinated to neighboring sequences of the behavior of interest (positive sequences) if the null sequences do not represent a fraction $(1 - \alpha)$ of the neighboring positive sequences *and* if they do not create a m-long episode in which the behavior of interest represents less than a fraction α. If by agglutinating null sequences a sequence of sequences longer than m is formed, we call it a bout. In the R-code, α equals 0.8, but the formulation was intended to be more general.

It should be noted that, to the best of my knowledge, this algorithm was never formalized and published in a paper; the explanations provided here are based on my own analysis of the R-code found in the *activpalProcessing* package.

How do different definitions of bouts compare to each other?

We have seen three main definitions of activity bout: one based on the traditional count number inherited from the Actigraph accelerometers, a simple definition based on a purity threshold and maximum interruption time, and the definition by Lyden, in which no interruption time exists, but null sequences are allowed into depending on neighboring sequences.

Determining whether these operational definitions and their parameters result in significantly different metrics of physical activity phenotyping would require comparing them on a heterogenous real-life dataset. In addition, it would be interesting to show whether they can affect estimated relationships with various health outcomes. Such a comparison was undertaken in a few studies, but they only account for varying epoch lengths within Actigraph's bout definition [46, 50]. Such an endeavor is beyond the scope of this short work.

Here, I consider a toy example based on simulated ambulation, with alternating sequences of walking and quiet standing whose lengths are random values drawn from Poisson distributions with various mean parameters λ. In Figure 1, we see 3 panels corresponding to different random scenarios. The color boxes represent the time detected as walking bouts according to the different definitions. In black, we have the real walking time, recorded at a resolution of the base time unit t (say 1 second). In blue, we have the bouts as in the Actigraph definition with 100-t epoch length, that is, motion intensity is summed over fixed intervals (epochs) of 100 t and assigned walking/resting depending on whether the sum of motion intensity exceeds a threshold. Here, the threshold was determined such that an epoch is considered walking if it amounts to over 80% of the epoch. In red, we have bouts as defined by Lyden, with a 80% purity threshold. In green, we have the simple definition proposed in section 3, with a 80% purity threshold and 20% maximum interruption time. The minimum bout length was set to 100 t in all definitions. Thus, parameters were kept the same across definitions to ensure better comparability.

Looking at Figure 1, a first noticeable fact is that many brief bursts of activity are not counted as bouts, which is the very purpose of working with bouts. Second, we notice that all Actigraph bouts (blue) are detected over fixed intervals, whereas they are dynamically detected in the two other algorithms.

Comparing the different algorithms, we see a high variability in size, number and location of the bouts across the definitions. Between t=480 and t=750 in the upper panel, we see that by the simple definition we detect one long bout, whereas Lyden detects two bouts, adding up to almost the same total time. The reason for this difference is that according to the simple definition, the criterion of purity is met for the entire bout (>0.8 of the bout is walking) and no interruption is longer than 20 (0.2 x min. bout length). However, according to Lyden, the interruption between t=600 and t=615 is too long *compared to the neighboring walking sequences* (520-599 and 616-698). Although the definition by Lyden seems stricter, we see that some regions in the middle and lower panels are detected by the Lyden definition but not when using the simple definition. This is because the Lyden definition fills up the remaining time needed to form a bout by agglutinating a short sequence of non-activity (even backward), while the simple definition proposed has the reasonable expectation that a bout must start with the activity under consideration. Compared to the other definitions, the Actigraph definition results in a much smaller number of bouts and shorter bout time. This is due the fact that the Actigraph definition has more difficulties finding bouts as it searches only over fixed intervals.

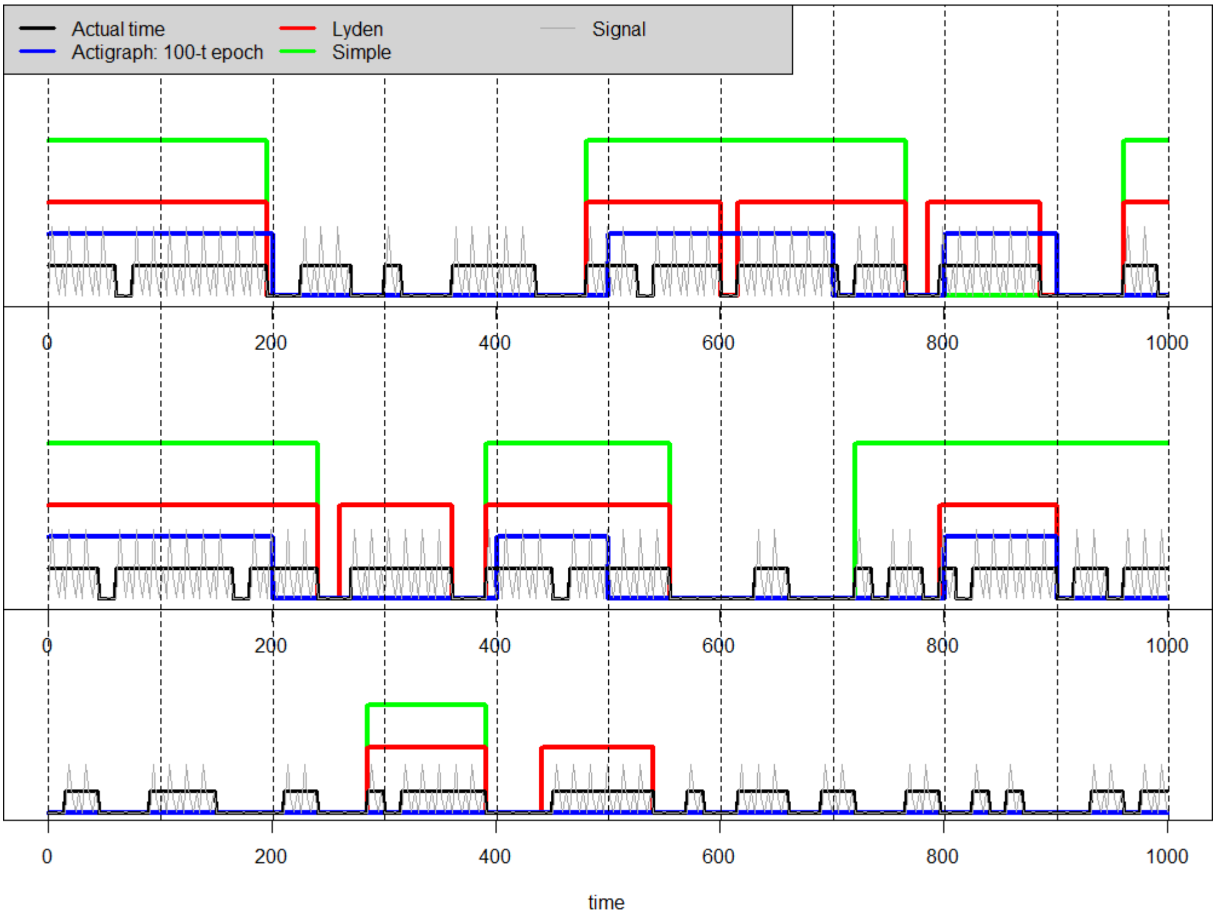

*Figure 14 : **Bout detection of walking activity according to the three definitions of bout** studied here. Detailed explanation are found in the text.*

Conclusion

In summary, bouts are low-resolution, homogenized sequences, which are assigned a single main behavior. Bout formation depends on the operational definition used and their parameters. A first inspection using a toy example showed a big variability in total time counted as bout, number and location of bouts over time. Despite the importance of these parameters for drawing conclusions, there is very little discussion in the literature about the bout detection definition used and the choice of parameters. Further work needs to investigate differences in a comprehensive manner, using real life data and preferably while examining consequences that they may have on assessing the relationships with health. In order to raise sensitivity to the methods of bout detection, and in order to enhance comparability across studies, a transparent, flexible definition accompanied by an algorithm could be a significant contribution to the epidemiology of inactivity.

BIBLIOGRAPHY

References in the Preamble, Introduction, Discussion, and Annex

[1] P. T. Katzmarzyk, « Physical Activity, Sedentary Behavior, and Health: Paradigm Paralysis or Paradigm Shift? », *Diabetes*, vol. 59, nº 11, p. 2717-2725, nov. 2010, doi: 10.2337/db10-0822.

[2] P. T. Katzmarzyk et C. Mason, « The Physical Activity Transition », *Journal of Physical Activity and Health*, vol. 6, nº 3, p. 269-280, mai 2009, doi: 10.1123/jpah.6.3.269.

[3] J.-P. Bocquet-Appel, « When the World's Population Took Off: The Springboard of the Neolithic Demographic Transition », *Science*, vol. 333, nº 6042, p. 560-561, juill. 2011, doi: 10.1126/science.1208880.

[4] M. Hayes, M. Chustek, S. Heshka, Z. Wang, A. Pietrobelli, et S. B. Heymsfield, « Low physical activity levels of modern Homo sapiens among free-ranging mammals », *International Journal of Obesity*, vol. 29, nº 1, p. 151-156, janv. 2005, doi: 10.1038/sj.ijo.0802842.

[5] T. S. Church *et al.*, « Trends over 5 decades in U.S. occupation-related physical activity and their associations with obesity », *PLoS ONE*, vol. 6, nº 5, p. e19657, 2011, doi: 10.1371/journal.pone.0019657.

[6] S. W. Ng et B. M. Popkin, « Time use and physical activity: a shift away from movement across the globe », *Obesity Reviews*, vol. 13, nº 8, p. 659-680, juin 2012, doi: 10.1111/j.1467-789X.2011.00982.x.

[7] E. Archer *et al.*, « 45-Year trends in women's use of time and household management energy expenditure », *PLoS ONE*, vol. 8, nº 2, p. e56620, 2013, doi: 10.1371/journal.pone.0056620.

[8] J. O. Hill et J. C. Peters, « Environmental contributions to the obesity epidemic », *Science*, vol. 280, nº 5368, p. 1371-1374, mai 1998, doi: 10.1126/science.280.5368.1371.

[9] « Scientific Report - 2018 Physical Activity Guidelines - health.gov ». [En ligne]. Disponible sur: https://health.gov/paguidelines/second-edition/report/. [Consulté le: 10-déc-2018].

[10] I.-M. Lee *et al.*, « Effect of physical inactivity on major non-communicable diseases worldwide: an analysis of burden of disease and life expectancy », *Lancet*, vol. 380, nº 9838, p. 219-229, juill. 2012, doi: 10.1016/S0140-6736(12)61031-9.

[11] C. P. Wen *et al.*, « Minimum amount of physical activity for reduced mortality and extended life expectancy: a prospective cohort study », *The Lancet*, vol. 378, nº 9798, p. 1244-1253, oct. 2011, doi: 10.1016/S0140-6736(11)60749-6.

[12] R. P. Troiano, D. Berrigan, K. W. Dodd, L. C. MâSse, T. Tilert, et M. Mcdowell, « Physical Activity in the United States Measured by Accelerometer »:, *Medicine & Science in Sports & Exercise*, vol. 40, nº 1, p. 181-188, janv. 2008, doi: 10.1249/mss.0b013e31815a51b3.

[13] D. Ding *et al.*, « The economic burden of physical inactivity : a global analysis of major non-communicable diseases », *The Lancet*, p. 1311-1324, 2016, doi: 10.1016/S0140-6736(16)30383-X.

[14] H. W. Kohl *et al.*, « The pandemic of physical inactivity: global action for public health », *The Lancet*, vol. 380, nº 9838, p. 294-305, juill. 2012, doi: 10.1016/S0140-6736(12)60898-8.

[15] D. E. R. Warburton, C. W. Nicol, et S. S. D. Bredin, « Health benefits of physical activity: the evidence », *CMAJ*, vol. 174, nº 6, p. 801-809, mars 2006, doi: 10.1503/cmaj.051351.

[16] A. Bergouignan, F. Rudwill, C. Simon, et S. Blanc, « Physical inactivity as the culprit of metabolic inflexibility: evidence from bed-rest studies », *Journal of Applied Physiology*, vol. 111, nº 4, p. 1201-1210, août 2011, doi: 10.1152/japplphysiol.00698.2011.

[17] M. T. Hamilton, G. N. Healy, D. W. Dunstan, T. W. Zderic, et N. Owen, « Too Little Exercise and Too Much Sitting: Inactivity Physiology and the Need for New Recommendations on Sedentary Behavior », *Curr Cardiovasc Risk Rep*, vol. 2, nº 4, p. 292-298, juill. 2008, doi: 10.1007/s12170-008-0054-8.

[18] G. N. Healy *et al.*, « Objectively measured light-intensity physical activity is independently associated with 2-h plasma glucose », *Diabetes Care*, vol. 30, nº 6, p. 1384-1389, juin 2007, doi: 10.2337/dc07-0114.

[19] A. Koster *et al.*, « Association of Sedentary Time with Mortality Independent of Moderate to Vigorous Physical Activity », *PLOS ONE*, vol. 7, nº 6, p. e37696, juin 2012, doi: 10.1371/journal.pone.0037696.

[20] W. L. Haskell *et al.*, « Physical activity and public health: updated recommendation for adults from the American College of Sports Medicine and the American Heart Association », *Med Sci Sports Exerc*, vol. 39, nº 8, p. 1423-1434, août 2007, doi: 10.1249/mss.0b013e3180616b27.

[21] J. A. Levine *et al.*, « Interindividual Variation in Posture Allocation: Possible Role in Human Obesity », *Science*, vol. 307, nº 5709, p. 584-586, janv. 2005, doi: 10.1126/science.1106561.

[22] J. P. Buckley, D. D. Mellor, M. Morris, et F. Joseph, « Standing-based office work shows encouraging signs of attenuating post-prandial glycaemic excursion », *Occup Environ Med*, vol. 71, nº 2, p. 109-111, févr. 2014, doi: 10.1136/oemed-2013-101823.

[23] G. N. Healy, E. A. H. Winkler, N. Owen, S. Anuradha, et D. W. Dunstan, « Replacing sitting time with standing or stepping: associations with cardio-metabolic risk biomarkers », *Eur Heart J*, vol. 36, nº 39, p. 2643-2649, oct. 2015, doi: 10.1093/eurheartj/ehv308.

[24] J. L. Miles-Chan et A. G. Dulloo, « Posture Allocation Revisited: Breaking the Sedentary Threshold of Energy Expenditure for Obesity Management », *Front Physiol*, vol. 8, juin 2017, doi: 10.3389/fphys.2017.00420.

[25] B. Ainsworth *et al.*, « 2011 Compendium of Physical Activities: A Second Update of Codes and MET Values », *Medicine & Science in Sports & Exercise*, vol. 43, nº 8, p. 1575-1581, août 2011, doi: 10.1249/MSS.0b013e31821ece12.

[26] C. Maher, T. Olds, E. Mire, et P. T. Katzmarzyk, « Reconsidering the Sedentary Behaviour Paradigm », *PLOS ONE*, vol. 9, nº 1, p. e86403, janv. 2014, doi: 10.1371/journal.pone.0086403.

[27] Young Deborah Rohm *et al.*, « Sedentary Behavior and Cardiovascular Morbidity and Mortality: A Science Advisory From the American Heart Association », *Circulation*, vol. 134, nº 13, p. e262-e279, sept. 2016, doi: 10.1161/CIR.0000000000000440.

[28] I. Debache, A. Bergouignan, B. Chaix, E. M. Sneekes, F. Thomas, et C. Sueur, « Associations of Sensor-Derived Physical Behavior with Metabolic Health: A Compositional Analysis in the Record Multisensor Study », *International Journal of Environmental Research and Public Health*, vol. 16, nº 5, p. 741, janv. 2019, doi: 10.3390/ijerph16050741.

[29] R. A. Mekary, W. C. Willett, F. B. Hu, et E. L. Ding, « Isotemporal substitution paradigm for physical activity epidemiology and weight change », *Am. J. Epidemiol.*, vol. 170, nº 4, p. 519-527, août 2009, doi: 10.1093/aje/kwp163.

[30] J. J. Egozcue, V. Pawlowsky-Glahn, G. Mateu-Figueras, et C. Barceló-Vidal, « Isometric Logratio Transformations for Compositional Data Analysis », *Mathematical Geology*, vol. 35, nº 3, p. 279–300, avril 2003, doi: 10.1023/A:1023818214614.

[31] Ž. Pedišić, D. Dumuid, et T. S. Olds, « 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*, vol. 49, nº 2, p. 252-269, sept. 2017.

[32] J. Clarke et I. Janssen, « Sporadic and bouted physical activity and the metabolic syndrome in adults », *Med Sci Sports Exerc*, vol. 46, nº 1, p. 76-83, janv. 2014, doi: 10.1249/MSS.0b013e31829f83a0.

[33] U. Ekelund, S. Brage, H. Besson, S. Sharp, et N. J. Wareham, « Time spent being sedentary and weight gain in healthy adults: reverse or bidirectional causality? », *Am J Clin Nutr*, vol. 88, nº 3, p. 612-617, sept. 2008, doi: 10.1093/ajcn/88.3.612.

[34] B. J. Jefferis *et al.*, « Does duration of physical activity bouts matter for adiposity and metabolic syndrome? A cross-sectional study of older British men », *International Journal of Behavioral Nutrition and Physical Activity*, vol. 13, nº 1, p. 36, mars 2016, doi: 10.1186/s12966-016-0361-2.

[35] S. J. Strath, R. G. Holleman, D. L. Ronis, A. M. Swartz, et C. R. Richardson, « Objective physical activity accumulation in bouts and nonbouts and relation to markers of obesity in US adults », *Prev Chronic Dis*, vol. 5, nº 4, p. A131, oct. 2008.

[36] D. L. Wolff-Hughes, E. C. Fitzhugh, D. R. Bassett, et J. R. Churilla, « Total Activity Counts and Bouted Minutes of Moderate-To-Vigorous Physical Activity: Relationships With Cardiometabolic Biomarkers Using 2003–2006 NHANES », *Journal of Physical Activity and Health*, vol. 12, nº 5, p. 694-700, mai 2015, doi: 10.1123/jpah.2013-0463.

[37] J. X. Fan, B. B. Brown, H. Hanson, L. Kowaleski-Jones, K. R. Smith, et C. D. Zick, « Moderate to Vigorous Physical Activity and Weight Outcomes: Does Every Minute Count? », *Am J Health Promot*, vol. 28, nº 1, p. 41-49, sept. 2013, doi: 10.4278/ajhp.120606-QUAL-286.

[38] P. D. Loprinzi et B. J. Cardinal, « Association between biologic outcomes and objectively measured physical activity accumulated in ≥ 10-minute bouts and <10-minute bouts », *Am J Health Promot*, vol. 27, nº 3, p. 143-151, févr. 2013, doi: 10.4278/ajhp.110916-QUAN-348.

[39] G. N. Healy *et al.*, « Breaks in Sedentary Time: Beneficial associations with metabolic risk », *Diabetes Care*, vol. 31, nº 4, p. 661-666, avr. 2008, doi: 10.2337/dc07-2046.

[40] D. W. Dunstan *et al.*, « Breaking Up Prolonged Sitting Reduces Postprandial Glucose and Insulin Responses », *Diabetes Care*, vol. 35, nº 5, p. 976-983, mai 2012, doi: 10.2337/dc11-1931.

[41] M. C. Peddie, J. L. Bone, N. J. Rehrer, C. M. Skeaff, A. R. Gray, et T. L. Perry, « Breaking prolonged sitting reduces postprandial glycemia in healthy, normal-weight adults: a randomized crossover trial », *Am J Clin Nutr*, vol. 98, nº 2, p. 358-366, août 2013, doi: 10.3945/ajcn.112.051763.

[42] M. Holmstrup, T. Fairchild, S. Keslacy, R. Weinstock, et J. Kanaley, « Multiple short bouts of exercise over 12-h period reduce glucose excursions more than an energy-matched single bout of exercise », *Metab. Clin. Exp.*, vol. 63, nº 4, p. 510-519, avr. 2014, doi: 10.1016/j.metabol.2013.12.006.

[43] K. Lyden, S. K. Keadle, J. Staudenmayer, B. Braun, et P. S. Freedson, « Discrete features of sedentary behavior impact cardiometabolic risk factors », *Med Sci Sports Exerc*, vol. 47, nº 5, p. 1079-1086, mai 2015, doi: 10.1249/MSS.0000000000000499.

[44] D. P. Bailey et C. D. Locke, « Breaking up prolonged sitting with light-intensity walking improves postprandial glycemia, but breaking up sitting with standing does not », *J Sci Med Sport*, vol. 18, nº 3, p. 294-298, mai 2015, doi: 10.1016/j.jsams.2014.03.008.

[45] A. A. Thorp, B. A. Kingwell, P. Sethi, L. Hammond, N. Owen, et D. W. Dunstan, « Alternating bouts of sitting and standing attenuate postprandial glucose responses », *Med Sci Sports Exerc*, vol. 46, nº 11, p. 2053-2061, nov. 2014, doi: 10.1249/MSS.0000000000000337.

[46] T. M. Altenburg *et al.*, « Occurrence and duration of various operational definitions of sedentary bouts and cross-sectional associations with cardiometabolic health indicators: The ENERGY-project », *Preventive Medicine*, vol. 71, p. 101-106, févr. 2015, doi: 10.1016/j.ypmed.2014.12.015.

[47] M. S. Tremblay *et al.*, « Sedentary Behavior Research Network (SBRN) – Terminology Consensus Project process and outcome », *International Journal of Behavioral Nutrition and Physical Activity*, vol. 14, nº 1, p. 75, juin 2017, doi: 10.1186/s12966-017-0525-8.

[48] X. Janssen et D. P. Cliff, « Issues Related to Measuring and Interpreting Objectively Measured Sedentary Behavior Data », *Measurement in Physical Education and Exercise Science*, vol. 19, nº 3, p. 116-124, juill. 2015, doi: 10.1080/1091367X.2015.1045908.

[49] Y. Kim, G. J. Welk, S. I. Braun, et M. Kang, « Extracting Objective Estimates of Sedentary Behavior from Accelerometer Data: Measurement Considerations for Surveillance and Research Applications », *PLOS ONE*, vol. 10, nº 2, p. e0118078, févr. 2015, doi: 10.1371/journal.pone.0118078.

[50] T. M. Altenburg et M. J. M. Chinapaw, « Bouts and breaks in children's sedentary time: currently used operational definitions and recommendations for future research », *Prev Med*, vol. 77, p. 1-3, août 2015, doi: 10.1016/j.ypmed.2015.04.019.

[51] S. F. M. Chastin et M. H. Granat, « Methods for objective measure, quantification and analysis of sedentary behaviour and inactivity », *Gait & Posture*, vol. 31, nº 1, p. 82–86, janv. 2010, doi: 10.1016/j.gaitpost.2009.09.002.

[52] S. F. M. Chastin, J. Palarea-Albaladejo, M. L. Dontje, et D. A. Skelton, « 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*, vol. 10, nº 10, p. e0139984, oct. 2015, doi: 10.1371/journal.pone.0139984.

[53] S. A. Prince, K. B. Adamo, M. E. Hamel, J. Hardt, S. C. Gorber, et M. Tremblay, « A comparison of direct versus self-report measures for assessing physical activity in adults: a systematic review », *Int J Behav Nutr Phys Act*, vol. 5, p. 56, nov. 2008, doi: 10.1186/1479-5868-5-56.

[54] R. P. Troiano, J. J. McClain, R. J. Brychta, et K. Y. Chen, « Evolution of accelerometer methods for physical activity research », *Br J Sports Med*, vol. 48, nº 13, p. 1019-1023, juill. 2014, doi: 10.1136/bjsports-2014-093546.

[55] C. E. Matthews *et al.*, « Amount of Time Spent in Sedentary Behaviors in the United States, 2003–2004 », *Am J Epidemiol*, vol. 167, nº 7, p. 875-881, avr. 2008, doi: 10.1093/aje/kwm390.

[56] C. R. Richardson, T. L. Newton, J. J. Abraham, A. Sen, M. Jimbo, et A. M. Swartz, « A meta-analysis of pedometer-based walking interventions and weight loss », *Ann Fam Med*, vol. 6, nº 1, p. 69-77, févr. 2008, doi: 10.1370/afm.761.

[57] C. V. C. Bouten, K. T. M. Koekkoek, M. Verduin, R. Kodde, et J. D. Janssen, « A triaxial accelerometer and portable data processing unit for the assessment of daily physical activity », *IEEE Transactions on Biomedical Engineering*, vol. 44, nº 3, p. 136-147, mars 1997, doi: 10.1109/10.554760.

[58] L. Bao et S. S. Intille, « Activity Recognition from User-Annotated Acceleration Data », in *Pervasive Computing*, 2004, p. 1-17.

[59] J. Fridolfsson, M. Börjesson, et D. Arvidsson, « A Biomechanical Re-Examination of Physical Activity Measurement with Accelerometers », *Sensors*, vol. 18, nº 10, p. 3399, oct. 2018, doi: 10.3390/s18103399.

[60] L. A. Kelly, D. G. McMillan, A. Anderson, M. Fippinger, G. Fillerup, et J. Rider, « Validity of actigraphs uniaxial and triaxial accelerometers for assessment of physical activity in adults in laboratory conditions », *BMC Medical Physics*, vol. 13, nº 1, p. 5, nov. 2013, doi: 10.1186/1756-6649-13-5.

[61] J. H. Migueles *et al.*, « Accelerometer Data Collection and Processing Criteria to Assess Physical Activity and Other Outcomes: A Systematic Review and Practical Considerations », *Sports Med*, vol. 47, nº 9, p. 1821-1845, sept. 2017, doi: 10.1007/s40279-017-0716-0.

[62] O. Banos, J.-M. Galvez, M. Damas, H. Pomares, et I. Rojas, « Window Size Impact in Human Activity Recognition », *Sensors (Basel)*, vol. 14, nº 4, p. 6474-6499, avr. 2014, doi: 10.3390/s140406474.

[63] T. Huynh et B. Schiele, « Analyzing Features for Activity Recognition », in *Proceedings of the 2005 Joint Conference on Smart Objects and Ambient Intelligence: Innovative Context-aware Services: Usages and Technologies*, New York, NY, USA, 2005, p. 159–163, doi: 10.1145/1107548.1107591.

[64] J. Ortiz Laguna, A. G. Olaya, et D. Borrajo, « A Dynamic Sliding Window Approach for Activity Recognition », in *User Modeling, Adaption and Personalization*, Berlin, Heidelberg, 2011, p. 219-230, doi: 10.1007/978-3-642-22362-4_19.

[65] V. T. van Hees *et al.*, « Separating Movement and Gravity Components in an Acceleration Signal and Implications for the Assessment of Human Daily Physical Activity », *PLOS ONE*, vol. 8, nº 4, p. e61691, avr. 2013, doi: 10.1371/journal.pone.0061691.

[66] J. Bai *et al.*, « An Activity Index for Raw Accelerometry Data and Its Comparison with Other Activity Metrics », *PLOS ONE*, vol. 11, nº 8, p. e0160644, août 2016, doi: 10.1371/journal.pone.0160644.

[67] K. Y. Chen et D. R. J. Bassett, « The Technology of Accelerometry-Based Activity Monitors: Current and Future », *Medicine & Science in Sports & Exercise*, vol. 37, nº 11, p. S490, nov. 2005, doi: 10.1249/01.mss.0000185571.49104.82.

[68] J. C. Brønd, L. B. Andersen, et D. Arvidsson, « Generating ActiGraph Counts from Raw Acceleration Recorded by an Alternative Monitor », *Med Sci Sports Exerc*, vol. 49, nº 11, p. 2351-2360, nov. 2017, doi: 10.1249/MSS.0000000000001344.

[69] P. Barralon, N. Vuillerme, et N. Noury, « Walk detection with a kinematic sensor: frequency and wavelet comparison », *Conf Proc IEEE Eng Med Biol Soc*, vol. 1, p. 1711-1714, 2006, doi: 10.1109/IEMBS.2006.260770.

[70] J.-H. Wang, J.-J. Ding, Y. Chen, et H.-H. Chen, « Real time accelerometer-based gait recognition using adaptive windowed wavelet transforms », in *2012 IEEE Asia Pacific Conference on Circuits and Systems*, 2012, p. 591-594, doi: 10.1109/APCCAS.2012.6419104.

[71] M. Awais, S. Mellone, et L. Chiari, « Physical activity classification meets daily life: Review on existing methodologies and open challenges », in *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2015, p. 5050-5053, doi: 10.1109/EMBC.2015.7319526.

[72] D. Figo, P. C. Diniz, D. R. Ferreira, et J. M. Cardoso, « Preprocessing Techniques for Context Recognition from Accelerometer Data », *Personal Ubiquitous Comput.*, vol. 14, nº 7, p. 645–662, oct. 2010, doi: 10.1007/s00779-010-0293-9.

[73] H. Leutheuser, D. Schuldhaus, et B. M. Eskofier, « Hierarchical, Multi-Sensor Based Classification of Daily Life Activities: Comparison with State-of-the-Art Algorithms Using a Benchmark Dataset », *PLOS ONE*, vol. 8, nº 10, p. e75196, oct. 2013, doi: 10.1371/journal.pone.0075196.

[74] T. Hastie, R. Tibshirani, et J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction, Second Edition*, 2nd edition. New York, NY: Springer, 2016.

[75] Y. Kwon, K. Kang, et C. Bae, « Unsupervised learning for human activity recognition using smartphone sensors », *Expert Systems with Applications*, vol. 41, nº 14, p. 6067-6074, oct. 2014, doi: 10.1016/j.eswa.2014.04.037.

[76] O. Yurur, C.-H. Liu, et W. Moreno, « Unsupervised posture detection by smartphone accelerometer », *Electronics Letters*, vol. 49, nº 8, p. 562-564, avr. 2013, doi: 10.1049/el.2013.0592.

[77] O. D. Lara et M. A. Labrador, « A Survey on Human Activity Recognition using Wearable Sensors », *IEEE Communications Surveys Tutorials*, vol. 15, nº 3, p. 1192-1209, Third 2013, doi: 10.1109/SURV.2012.110112.00192.

[78] J. Wang, Y. Chen, S. Hao, X. Peng, et L. Hu, « Deep learning for sensor-based activity recognition: A survey », *Pattern Recognition Letters*, vol. 119, p. 3-11, mars 2019, doi: 10.1016/j.patrec.2018.02.010.

[79] A. C. Müller et S. Guido, *Introduction to Machine Learning with Python: A Guide for Data Scientists*, 1 edition. Sebastopol, CA: O'Reilly Media, 2016.

[80] J.-B. Yang, N. Nhut, P. San, X. li, et P. Shonali, « Deep Convolutional Neural Networks on Multichannel Time Series for Human Activity Recognition », *IJCAI*, juill. 2015.

[81] E. Zdravevski *et al.*, « Improving Activity Recognition Accuracy in Ambient Assisted Living Systems by Automated Feature Engineering », *IEEE Access*, vol. PP, p. 1-1, mars 2017, doi: 10.1109/ACCESS.2017.2684913.

[82] N. Y. Hammerla, S. Halloran, et T. Plötz, « Deep, Convolutional, and Recurrent Models for Human Activity Recognition Using Wearables », in *IJCAI*, 2016.

[83] I. Cleland *et al.*, « Optimal Placement of Accelerometers for the Detection of Everyday Activities », *Sensors (Basel)*, vol. 13, nº 7, p. 9183-9200, juill. 2013, doi: 10.3390/s130709183.

[84] C. E. Matthews, M. Hagströmer, D. M. Pober, et H. R. Bowles, « Best Practices For Using Physical Activity Monitors In Population-Based Research », *Med Sci Sports Exerc*, vol. 44, nº 1 Suppl 1, p. S68-S76, janv. 2012, doi: 10.1249/MSS.0b013e3182399e5b.

[85] J. H. M. Bergmann, S. Graham, N. Howard, et A. McGregor, « Comparison of median frequency between traditional and functional sensor placements during activity monitoring », *Measurement (Lond)*, vol. 46, nº 7, p. 2193-2200, août 2013, doi: 10.1016/j.measurement.2013.03.004.

[86] C. C. T. Clark *et al.*, « Physical activity characterization: does one site fit all? », *Physiol. Meas.*, vol. 39, nº 9, p. 09TR02, sept. 2018, doi: 10.1088/1361-6579/aadad0.

[87] P. Tucker et J. Gilliland, « The effect of season and weather on physical activity: A systematic review », *Public Health*, vol. 121, nº 12, p. 909-922, déc. 2007, doi: 10.1016/j.puhe.2007.04.009.

[88] C. Cortis *et al.*, « Psychological determinants of physical activity across the life course: A "DEterminants of DIet and Physical ACtivity" (DEDIPAC) umbrella systematic literature review », *PloS One*, vol. 12, nº 8, p. e0182709, 2017, doi: 10.1371/journal.pone.0182709.

[89] G. Condello *et al.*, « Behavioral determinants of physical activity across the life course: a "DEterminants of DIet and Physical ACtivity" (DEDIPAC) umbrella systematic literature review », *Int J Behav Nutr Phys Act*, vol. 14, nº 1, p. 58, 02 2017, doi: 10.1186/s12966-017-0510-2.

[90] B. Giles-Corti *et al.*, « The influence of urban design on neighbourhood walking following residential relocation: Longitudinal results from the RESIDE study », *Social Science & Medicine*, vol. 77, p. 20-30, janv. 2013, doi: 10.1016/j.socscimed.2012.10.016.

[91] M. Smith *et al.*, « Systematic literature review of built environment effects on physical activity and active transport - an update and new findings on health equity », *Int J Behav Nutr Phys Act*, vol. 14, nº 1, p. 158, 16 2017, doi: 10.1186/s12966-017-0613-9.

[92] A. R. Abubakari, W. Lauder, M. C. Jones, A. Kirk, C. Agyemang, et R. S. Bhopal, « Prevalence and time trends in diabetes and physical inactivity among adult West African populations: The epidemic has arrived », *Public Health*, vol. 123, nº 9, p. 602-614, sept. 2009, doi: 10.1016/j.puhe.2009.07.009.

[93] A. Carlin *et al.*, « A life course examination of the physical environmental determinants of physical activity behaviour: A "Determinants of Diet and Physical Activity" (DEDIPAC) umbrella systematic literature review », *PLOS ONE*, vol. 12, nº 8, p. e0182083, août 2017, doi: 10.1371/journal.pone.0182083.

[94] K. L. Monda, P. Gordon-Larsen, J. Stevens, et B. M. Popkin, « China's transition: The effect of rapid urbanization on adult occupational physical activity », *Social Science & Medicine*, vol. 64, nº 4, p. 858-870, févr. 2007, doi: 10.1016/j.socscimed.2006.10.019.

[95] G. R. McCormack et A. Shiell, « In search of causality: a systematic review of the relationship between the built environment and physical activity among adults », *International Journal of Behavioral Nutrition and Physical Activity*, vol. 8, nº 1, p. 125, nov. 2011, doi: 10.1186/1479-5868-8-125.

[96] L. D. Frank *et al.*, « The development of a walkability index: application to the Neighborhood Quality of Life Study », *Br J Sports Med*, vol. 44, nº 13, p. 924-933, oct. 2010, doi: 10.1136/bjsm.2009.058701.

[97] D. W. Barnett, A. Barnett, A. Nathan, J. Van Cauwenberg, et E. Cerin, « Built environmental correlates of older adults' total physical activity and walking: a systematic review and meta-analysis », *International Journal of Behavioral Nutrition and Physical Activity*, vol. 14, p. 103, août 2017, doi: 10.1186/s12966-017-0558-z.

[98] V. Van Holle *et al.*, « Relationship between the physical environment and different domains of physical activity in European adults: a systematic review », *BMC Public Health*, vol. 12, nº 1, p. 807, sept. 2012, doi: 10.1186/1471-2458-12-807.

[99] E. Cerin, A. Nathan, J. van Cauwenberg, D. W. Barnett, A. Barnett, et on behalf of the Council on Environment and Physical Activity (CEPA) – Older Adults working group, « The neighbourhood physical environment and active travel in older adults: a systematic review and meta-analysis », *International Journal of Behavioral Nutrition and Physical Activity*, vol. 14, nº 1, p. 15, févr. 2017, doi: 10.1186/s12966-017-0471-5.

[100] J. Van Cauwenberg, A. Nathan, A. Barnett, D. W. Barnett, E. Cerin, et Council on Environment and Physical Activity (CEPA)-Older Adults Working Group, « Relationships Between Neighbourhood Physical Environmental Attributes and Older Adults' Leisure-Time Physical Activity: A Systematic Review and Meta-Analysis », *Sports Med*, vol. 48, nº 7, p. 1635-1660, 2018, doi: 10.1007/s40279-018-0917-1.

[101] B. E. Saelens et S. L. Handy, « Built Environment Correlates of Walking: A Review », *Med Sci Sports Exerc*, vol. 40, nº 7 Suppl, p. S550-S566, juill. 2008, doi: 10.1249/MSS.0b013e31817c67a4.

[102] B. B. Brown, C. M. Werner, C. P. Tribby, H. J. Miller, et K. R. Smith, « Transit Use, Physical Activity, and Body Mass Index Changes: Objective Measures Associated With Complete Street Light-Rail Construction », *Am J Public Health*, vol. 105, nº 7, p. 1468-1474, mai 2015, doi: 10.2105/AJPH.2015.302561.

[103] M. W. Knuiman *et al.*, « A longitudinal analysis of the influence of the neighborhood built environment on walking for transportation: the RESIDE study », *Am. J. Epidemiol.*, vol. 180, nº 5, p. 453-461, sept. 2014, doi: 10.1093/aje/kwu171.

[104] M. J. Duncan, H. M. Badland, et W. K. Mummery, « Applying GPS to enhance understanding of transport-related physical activity », *Journal of Science and Medicine in Sport*, vol. 12, nº 5, p. 549-556, sept. 2009, doi: 10.1016/j.jsams.2008.10.010.

[105] P. J. Krenn, S. Titze, P. Oja, A. Jones, et D. Ogilvie, « Use of global positioning systems to study physical activity and the environment: a systematic review », *Am J Prev Med*, vol. 41, nº 5, p. 508-515, nov. 2011, doi: 10.1016/j.amepre.2011.06.046.

[106] C. Perchoux, Y. Kestens, F. Thomas, A. V. Hulst, B. Thierry, et B. Chaix, « Assessing patterns of spatial behavior in health studies: Their socio-demographic determinants and associations with transportation modes (the RECORD Cohort Study) », *Social Science & Medicine*, vol. 119, p. 64-73, oct. 2014, doi: 10.1016/j.socscimed.2014.07.026.

[107] P. James *et al.*, « GPS-Based Exposure to Greenness and Walkability and Accelerometry-Based Physical Activity », *Cancer Epidemiol Biomarkers Prev*, févr. 2017, doi: 10.1158/1055-9965.EPI-16-0925.

[108] B. Chaix, « Mobile Sensing in Environmental Health and Neighborhood Research », *Annual Review of Public Health*, vol. 39, nº 1, p. 367-384, 2018, doi: 10.1146/annurev-publhealth-040617-013731.

[109] M. Browning et K. Lee, « Within What Distance Does "Greenness" Best Predict Physical Health? A Systematic Review of Articles with GIS Buffer Analyses across the Lifespan », *International Journal of Environmental Research and Public Health*, vol. 14, nº 7, p. 675, juill. 2017, doi: 10.3390/ijerph14070675.

[110] B. Chaix *et al.*, « Cohort profile: residential and non-residential environments, individual activity spaces and cardiovascular risk factors and diseases–the RECORD Cohort Study », *International Journal of Epidemiology*, vol. 41, nº 5, p. 1283–1292, oct. 2012, doi: 10.1093/ije/dyr107.

[111] A. Nathan, G. Pereira, S. Foster, P. Hooper, D. Saarloos, et B. Giles-Corti, « Access to commercial destinations within the neighbourhood and walking among Australian older adults », *Int J Behav Nutr Phys Act*, vol. 9, p. 133, nov. 2012, doi: 10.1186/1479-5868-9-133.

[112] K. Lachowycz, A. P. Jones, A. S. Page, B. W. Wheeler, et A. R. Cooper, « What can global positioning systems tell us about the contribution of different types of urban greenspace to children's physical activity? », *Health & Place*, vol. 18, nº 3, p. 586-594, mai 2012, doi: 10.1016/j.healthplace.2012.01.006.

[113] B. W. Wheeler, A. R. Cooper, A. S. Page, et R. Jago, « Greenspace and children's physical activity: A GPS/GIS analysis of the PEACH project », *Preventive Medicine*, vol. 51, nº 2, p. 148-152, août 2010, doi: 10.1016/j.ypmed.2010.06.001.

[114] B. Chaix *et al.*, « A GPS-Based Methodology to Analyze Environment-Health Associations at the Trip Level: Case-Crossover Analyses of Built Environments and Walking », *Am J Epidemiol*, vol. 184, nº 8, p. 579-589, oct. 2016, doi: 10.1093/aje/kww071.

[115] B. Chaix *et al.*, « GPS tracking in neighborhood and health studies: a step forward for environmental exposure assessment, a step backward for causal inference? », *Health Place*, vol. 21, p. 46-51, mai 2013, doi: 10.1016/j.healthplace.2013.01.003.

[116] U. Ekelund, S. Brage, H. Besson, S. Sharp, et N. J. Wareham, « Time spent being sedentary and weight gain in healthy adults: reverse or bidirectional causality? », *Am J Clin Nutr*, vol. 88, nº 3, p. 612-617, sept. 2008, doi: 10.1093/ajcn/88.3.612.

[117] D. W. Dunstan *et al.*, « Breaking Up Prolonged Sitting Reduces Postprandial Glucose and Insulin Responses », *Diabetes Care*, vol. 35, nº 5, p. 976–983, mai 2012, doi: 10.2337/dc11-1931.

[118] G. N. Healy *et al.*, « Breaks in Sedentary Time: Beneficial Associations with Metabolic Risk », *Diabetes Care*, févr. 2008, doi: 10.2337/dc07-2046.

[119] K. Lyden, *activpalProcessing: Process activPAL Events Files*. 2016.

[120] J. D. Bassett, D. John, S. A. Conger, B. C. Rider, R. M. Passmore, et J. M. Clark, « Detection of lying down, sitting, standing, and stepping using two activPAL monitors. », *Med Sci Sports Exerc*, vol. 46, nº 10, p. 2025-2029, oct. 2014, doi: 10.1249/MSS.0000000000000326.

[121] Le Monde. « Grève des transports : la fréquentation des voies cyclables à Paris continue d'augmenter », 16 dec. 2019. https://www.lemonde.fr/les-decodeurs/article/2019/12/16/greve-des-transports-la-frequentation-des-voies-cyclables-a-paris-continue-d-augmenter_6023096_4355770.html

[122] M. Weiser. « The computer of the 21st century ». *Scientific American.* Vol 265, no° 3, pp. 94-105. 1991.

[123] M. Garnotel, T. Bastian T, H.M. Romero-Ugalde, A. Maire , J. Dugas , A. Zahariev , *et al*. "Prior automatic posture and activity identification improves physical activity energy expenditure prediction from hip-worn triaxial accelerometry". *Journal of Applied Physiology.* 2017. Vol 124, p. 780–90.

[124] U. Ekelund *et al.* "Does physical activity attenuate, or even eliminate, the detrimental association of sitting time with mortality? A harmonised meta-analysis of data from more than 1 million men and women". *Lancet* 2016. Vol 388, p. 1302-1310.

References cited in chapter I

1. Warburton DER, Nicol CW, Bredin SSD. Health benefits of physical activity: the evidence. CMAJ. 2006;174:801–9.

2. Kohl HW, Craig CL, Lambert EV, Inoue S, Alkandari JR, Leetongin G, et al. The pandemic of physical inactivity: global action for public health. The Lancet. 2012;380:294–305.

3. Lee I-M, Shiroma EJ, Lobelo F, Puska P, Blair SN, Katzmarzyk PT, et al. Effect of physical inactivity on major non-communicable diseases worldwide: an analysis of burden of disease and life expectancy. Lancet (London, England). 2012;380:219–229.

4. Healy GN, Wijndaele K, Dunstan DW, Shaw JE, Salmon J, Zimmet PZ, et al. Objectively measured sedentary time, physical activity, and metabolic risk: the Australian Diabetes, Obesity and Lifestyle Study (AusDiab). Diabetes Care. 2008;31:369–371.

5. Thorp AA, Owen N, Neuhaus M, Dunstan DW. Sedentary behaviors and subsequent health outcomes in adults a systematic review of longitudinal studies, 1996-2011. American Journal of Preventive Medicine. 2011;41:207–215.

6. Wilmot EG, Edwardson CL, Achana FA, Davies MJ, Gorely T, Gray LJ, et al. Sedentary time in adults and the association with diabetes, cardiovascular disease and death: systematic review and meta-analysis. Diabetologia. 2012;55:2895–2905.

7. Andersen E, Ekelund U, Anderssen SA. Effects of Reducing Sedentary Time on Glucose Metabolism in Immigrant Pakistani Men. Medicine and Science in Sports and Exercise. 2015;47:775–781.

8. Howard RA, Freedman DM, Park Y, Hollenbeck A, Schatzkin A, Leitzmann MF. Physical activity, sedentary behavior, and the risk of colon and rectal cancer in the NIH-AARP Diet and Health Study. Cancer causes & control: CCC. 2008;19:939–953.

9. Gierach GL, Chang S-C, Brinton LA, Lacey JV, Hollenbeck AR, Schatzkin A, et al. Physical Activity, Sedentary Behavior, and Endometrial Cancer Risk in the NIH-AARP Diet and Health Study. International journal of cancer Journal international du cancer. 2009;124:2139–2147.

10. Hamilton MT, Hamilton DG, Zderic TW. Exercise physiology versus inactivity physiology: an essential concept for understanding lipoprotein lipase regulation. Exercise and Sport Sciences Reviews. 2004;32:161–166.

11. Owen N, Healy GN, Matthews CE, Dunstan DW. Too Much Sitting: The Population-Health Science of Sedentary Behavior. Exerc Sport Sci Rev. 2010;38:105–13.

12. Sedentary Behaviour Research Network. Letter to the editor: standardized use of the terms “sedentary” and “sedentary behaviours.” Applied Physiology, Nutrition, and Metabolism = Physiologie Appliquee, Nutrition Et Metabolisme. 2012;37:540–542.

13. Tremblay MS, Aubert S, Barnes JD, Saunders TJ, Carson V, Latimer-Cheung AE, et al. Sedentary Behavior Research Network (SBRN) – Terminology Consensus Project process and outcome. International Journal of Behavioral Nutrition and Physical Activity. 2017;14:75.

14. Gibbs BB, Hergenroeder AL, Katzmarzyk PT, Lee I-M, Jakicic JM. Definition, Measurement, and Health Risks Associated with Sedentary Behavior. Med Sci Sports Exerc. 2015;47:1295–300.

15. Alkhajah TA, Reeves MM, Eakin EG, Winkler EAH, Owen N, Healy GN. Sit-stand workstations: a pilot intervention to reduce office sitting time. Am J Prev Med. 2012;43:298–303.

16. Pulsford RM, Blackwell J, Hillsdon M, Kos K. 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. 2017;20:278–83.

17. Bellettiere J, Winkler EAH, Chastin SFM, Kerr J, Owen N, Dunstan DW, et al. Associations of sitting accumulation patterns with cardio-metabolic risk biomarkers in Australian adults. PLOS ONE. 2017;12:e0180119.

18. Biddle GJH, Edwardson CL, Henson J, Davies MJ, Khunti K, Rowlands AV, et al. Associations of Physical Behaviours and Behavioural Reallocations with Markers of Metabolic Health: A Compositional Data Analysis. Int J Environ Res Public Health. 2018;15.

19. Chastin SFM, Granat MH. Methods for objective measure, quantification and analysis of sedentary behaviour and inactivity. Gait & Posture. 2010;31:82–86.

20. Healy GN, Matthews CE, Dunstan DW, Winkler EAH, Owen N. Sedentary time and cardio-metabolic biomarkers in US adults: NHANES 2003–06. European Heart Journal. 2011;32:590–597.

21. Dunstan DW, Kingwell BA, Larsen R, Healy GN, Cerin E, Hamilton MT, et al. Breaking Up Prolonged Sitting Reduces Postprandial Glucose and Insulin Responses. Diabetes Care. 2012;35:976–983.

22. Aitchison J. The statistical analysis of compositional data. Caldwell, N.J: Blackburn Press; 2003.

23. van den Boogaart G, Tolosana-Delgado R. Analyzing compositional data with R. Berlin New York: Springer; 2013.

24. Pedišić Ž, Dumuid D, Olds TS. 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. 2017;49:252–269.

25. Mekary RA, Willett WC, Hu FB, Ding EL. Isotemporal substitution paradigm for physical activity epidemiology and weight change. American Journal of Epidemiology. 2009;170:519–527.

26. Healy GN, Winkler EAH, Owen N, Anuradha S, Dunstan DW. Replacing sitting time with standing or stepping: associations with cardio-metabolic risk biomarkers. European Heart Journal. 2015;36:2643–2649.

27. Fairclough SJ, Dumuid D, Taylor S, Curry W, McGrane B, Stratton G, et al. Fitness, fatness and the reallocation of time between children's daily movement behaviours: an analysis of compositional data. International Journal of Behavioral Nutrition and Physical Activity. 2017;14:64.

28. Talarico R, Janssen I. Compositional associations of time spent in sleep, sedentary behavior and physical activity with obesity measures in children. International Journal of Obesity. 2018;:1.

29. El Aarbaoui T, Méline J, Brondeel R, Chaix B. Short-term association between personal exposure to noise and heart rate variability: The RECORD MultiSensor Study. Environ Pollut. 2017;231 Pt 1:703–11.

30. Chaix B, Jouven X, Thomas F, Leal C, Billaudeau N, Bean K, et al. Why socially deprived populations have a faster resting heart rate: Impact of behaviour, life course anthropometry, and biology – the RECORD Cohort Study. Social Science & Medicine. 2011;73:1543–50.

31. Troiano RP, Berrigan D, Dodd KW, MâSse LC, Tilert T, Mcdowell M. Physical Activity in the United States Measured by Accelerometer: Medicine & Science in Sports & Exercise. 2008;40:181–8.

32. Lyden K. activpalProcessing: Process activPAL Events Files. 2016. https://CRAN.R-project.org/package=activpalProcessing. Accessed 17 Oct 2018.

33. Bowman SA. Television-viewing characteristics of adults: correlations to eating practices and overweight and health status. Preventing Chronic Disease. 2006;3:A38.

34. Chaix B, Kestens Y, Bean K, Leal C, Karusisi N, Meghiref K, et al. Cohort profile: residential and non-residential environments, individual activity spaces and cardiovascular risk factors and diseases–the RECORD Cohort Study. International Journal of Epidemiology. 2012;41:1283–1292.

35. Egozcue JJ, Pawlowsky-Glahn V, Mateu-Figueras G, Barceló-Vidal C. Isometric Logratio Transformations for Compositional Data Analysis. Mathematical Geology. 2003;35:279–300.

36. Boogaart KG van den, Tolosana-Delgado R, Bren M. compositions: Compositional Data Analysis. 2018. https://CRAN.R-project.org/package=compositions. Accessed 19 Oct 2018.

37. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

38. IDF Consensus Worldwide Definition of the Metabolic Syndrome. https://www.idf.org/our-activities/advocacy-awareness/resources-and-tools/60:idfconsensus-worldwide-definitionof-the-metabolic-syndrome.html. Accessed 31 Jan 2019.

39. Chaix B, Billaudeau N, Thomas F, Havard S, Evans D, Kestens Y, et al. Neighborhood effects on health: correcting bias from neighborhood effects on participation. Epidemiology. 2011;22:18–26.

40. Mann S, Beedie C, Jimenez A. Differential Effects of Aerobic Exercise, Resistance Training and Combined Exercise Modalities on Cholesterol and the Lipid Profile: Review, Synthesis and Recommendations. Sports Med. 2014;44:211–21.

41. Miles-Chan JL, Dulloo AG. Posture Allocation Revisited: Breaking the Sedentary Threshold of Energy Expenditure for Obesity Management. Front Physiol. 2017;8. doi:10.3389/fphys.2017.00420.

42. Tikkanen O, Haakana P, Pesola AJ, Häkkinen K, Rantalainen T, Havu M, et al. Muscle Activity and Inactivity Periods during Normal Daily Life. PLOS ONE. 2013;8:e52228.

43. Richter EA, Hargreaves M. Exercise, GLUT4, and skeletal muscle glucose uptake. Physiol Rev. 2013;93:993–1017.

44. Carson V, Wong SL, Winkler E, Healy GN, Colley RC, Tremblay MS. Patterns of sedentary time and cardiometabolic risk among Canadian adults. Preventive Medicine. 2014;65:23–7.

45. Jefferis BJ, Parsons TJ, Sartini C, Ash S, Lennon LT, Wannamethee SG, et al. Does duration of physical activity bouts matter for adiposity and metabolic syndrome? A cross-sectional study of older British men. International Journal of Behavioral Nutrition and Physical Activity. 2016;13:36.

46. Scientific Report - 2018 Physical Activity Guidelines - health.gov. https://health.gov/paguidelines/second-edition/report/. Accessed 10 Dec 2018.

47. Ekelund U, Brage S, Besson H, Sharp S, Wareham NJ. Time spent being sedentary and weight gain in healthy adults: reverse or bidirectional causality? Am J Clin Nutr. 2008;88:612–7

References cited in chapter II

1. Owen, N.; Healy, G.N.; Matthews, C.E.; Dunstan, D.W. Too Much Sitting: The Population-Health Science of Sedentary Behavior. *Exercise and Sport Sciences Reviews* **2010**, *38*, 105-113, doi:10.1097/JES.0b013e3181e373a2.
2. Bauman, A.E.; Chau, J.Y.; Ding, D.; Bennie, J. Too Much Sitting and Cardio-Metabolic Risk: An Update of Epidemiological Evidence. *Current Cardiovascular Risk Reports* **2013**, *7*, 293-298, doi:10.1007/s12170-013-0316-y.
3. Thorp, A.A.; Owen, N.; Neuhaus, M.; Dunstan, D.W. Sedentary behaviors and subsequent health outcomes in adults a systematic review of longitudinal studies, 1996-2011. *American journal of preventive medicine* **2011**, *41*, 207-215, doi:10.1016/j.amepre.2011.05.004.
4. Wilmot, E.G.; Edwardson, C.L.; Achana, F.A.; Davies, M.J.; Gorely, T.; Gray, L.J.; Khunti, K.; Yates, T.; Biddle, S.J. Sedentary time in adults and the association with diabetes, cardiovascular disease and death: systematic review and meta-analysis. *Diabetologia* **2012**, 55, 2895-2905, doi:10.1007/s00125-012-2677-z.
5. Edwardson, C.L.; Gorely, T.; Davies, M.J.; Gray, L.J.; Khunti, K.; Wilmot, E.G.; Yates, T.; Biddle, S.J. Association of sedentary behaviour with metabolic syndrome: a meta-analysis. *PloS one* **2012**, *7*, e34916, doi:10.1371/journal.pone.0034916.
6. de Rezende, L.F.; Rodrigues Lopes, M.; Rey-Lopez, J.P.; Matsudo, V.K.; Luiz Odo, C. Sedentary behavior and health outcomes: an overview of systematic reviews. *PloS one* **2014**, *9*, e105620, doi:10.1371/journal.pone.0105620.
7. Suchert, V.; Hanewinkel, R.; Isensee, B. Sedentary behavior and indicators of mental health in school-aged children and adolescents: A systematic review. *Preventive medicine* **2015**, *76*, 48-57, doi:https://doi.org/10.1016/j.ypmed.2015.03.026.
8. van Uffelen, J.G.; van Gellecum, Y.R.; Burton, N.W.; Peeters, G.; Heesch, K.C.; Brown, W.J. Sitting-time, physical activity, and depressive symptoms in mid-aged women. *American journal of preventive medicine* **2013**, *45*, 276-281, doi:10.1016/j.amepre.2013.04.009.
9. Rynders, C.A.; Blanc, S.; DeJong, N.; Bessesen, D.H.; Bergouignan, A. Sedentary behaviour is a key determinant of metabolic inflexibility. *The Journal of Physiology* **2017**, doi:10.1113/JP273282.
10. Judice, P.B.; Hamilton, M.T.; Sardinha, L.B.; Silva, A.M. Randomized controlled pilot of an intervention to reduce and break-up overweight/obese adults' overall sitting-time. *Trials* **2015**, *16*, 490, doi:10.1186/s13063-015-1015-4.
11. Healy, G.N.; Dunstan, D.W.; Salmon, J.; Cerin, E.; Shaw, J.E.; Zimmet, P.Z.; Owen, N. Breaks in sedentary time: beneficial associations with metabolic risk. *Diabetes Care* **2008**, *31*, 661-666, doi:10.2337/dc07-2046.
12. Healy, G.N.; Matthews, C.E.; Dunstan, D.W.; Winkler, E.A.; Owen, N. Sedentary time and cardio-metabolic biomarkers in US adults: NHANES 2003-06. *European heart journal* **2011**, *32*, 590-597, doi:10.1093/eurheartj/ehq451.
13. McCrady, S.K.; Levine, J.A. Sedentariness at Work: How Much Do We Really Sit? *Obesity* **2009**, *17*, 2103-2105, doi:doi:10.1038/oby.2009.117.
14. Ryan, C.G.; Dall, P.M.; Granat, M.H.; Grant, P.M. Sitting patterns at work: objective measurement of adherence to current recommendations. *Ergonomics* **2011**, *54*, 531-538, doi:10.1080/00140139.2011.570458.
15. Church, T.S.; Thomas, D.M.; Tudor-Locke, C.; Katzmarzyk, P.T.; Earnest, C.P.; Rodarte, R.Q.; Martin, C.K.; Blair, S.N.; Bouchard, C. Trends over 5 decades in U.S. occupation-related physical activity and their associations with obesity. *PloS one* **2011**, *6*, e19657, doi:10.1371/journal.pone.0019657.
16. Physical Activity Strategy for the WHO European Region. Availabe online: http://www.euro.who.int/__data/assets/pdf_file/0010/282961/65wd09e_PhysicalActivityStrategy_150474.pdf (accessed on 8/30).
17. Alkhajah, T.A.; Reeves, M.M.; Eakin, E.G.; Winkler, E.A.H.; Owen, N.; Healy, G.N. Sit–Stand Workstations: A Pilot Intervention to Reduce Office Sitting Time. *American journal of preventive medicine* **2012**, *43*, 298-303, doi:https://doi.org/10.1016/j.amepre.2012.05.027.

18. Carr, L.J.; Walaska, K.A.; Marcus, B.H. Feasibility of a portable pedal exercise machine for reducing sedentary time in the workplace. *British journal of sports medicine* **2012**, *46*, 430-435, doi:10.1136/bjsm.2010.079574.
19. Jones, R.A.; Hinkley, T.; Okely, A.D.; Salmon, J. Tracking physical activity and sedentary behavior in childhood: a systematic review. *American journal of preventive medicine* **2013**, *44*, 651-658, doi:10.1016/j.amepre.2013.03.001.
20. Gilson, N.D.; Suppini, A.; Ryde, G.C.; Brown, H.E.; Brown, W.J. Does the use of standing 'hot' desks change sedentary work time in an open plan office? *Preventive medicine* **2012**, *54*, 65-67, doi:10.1016/j.ypmed.2011.10.012.
21. Healy, G.N.; Eakin, E.G.; Owen, N.; Lamontagne, A.D.; Moodie, M.; Winkler, E.A.; Fjeldsoe, B.S.; Wiesner, G.; Willenberg, L.; Dunstan, D.W. A Cluster Randomized Controlled Trial to Reduce Office Workers' Sitting Time: Effect on Activity Outcomes. *Med Sci Sports Exerc* **2016**, *48*, 1787-1797, doi:10.1249/mss.0000000000000972.
22. Hutchinson, J.; Headley, S.; Matthews, T.; Spicer, G.; Dempsey, K.; Wooley, S.; Janssen, X. Changes in Sitting Time and Sitting Fragmentation after a Workplace Sedentary Behaviour Intervention. *International journal of environmental research and public health* **2018**, *15*, doi:10.3390/ijerph15061148.
23. Evans, R.E.; Fawole, H.O.; Sheriff, S.A.; Dall, P.M.; Grant, P.M.; Ryan, C.G. Point-of-choice prompts to reduce sitting time at work: a randomized trial. *American journal of preventive medicine* **2012**, *43*, 293-297, doi:10.1016/j.amepre.2012.05.010.
24. Carr, L.J.; Karvinen, K.; Peavler, M.; Smith, R.; Cangelosi, K. Multicomponent intervention to reduce daily sedentary time: a randomised controlled trial. *BMJ Open* **2013**, *3*, doi:10.1136/bmjopen-2013-003261.
25. Koepp, G.A.; Manohar, C.U.; McCrady-Spitzer, S.K.; Ben-Ner, A.; Hamann, D.J.; Runge, C.F.; Levine, J.A. Treadmill desks: A 1-year prospective trial. *Obesity (Silver Spring, Md.)* **2013**, *21*, 705-711, doi:10.1002/oby.20121.
26. Tudor-Locke, C.; Schuna, J.M., Jr.; Frensham, L.J.; Proenca, M. Changing the way we work: elevating energy expenditure with workstation alternatives. *International journal of obesity (2005)* **2014**, *38*, 755-765, doi:10.1038/ijo.2013.223.
27. Levine, J.A.; Miller, J.M. The energy expenditure of using a “walk-and-work” desk for office workers with obesity. *British journal of sports medicine* **2007**, *41*, 558.
28. Wennberg, P.; Boraxbekk, C.-J.; Wheeler, M.; Howard, B.; Dempsey, P.C.; Lambert, G.; Eikelis, N.; Larsen, R.; Sethi, P.; Occleston, J., et al. Acute effects of breaking up prolonged sitting on fatigue and cognition: a pilot study. *BMJ Open* **2016**, *6*.
29. Thorp, A.A.; Kingwell, B.A.; Owen, N.; Dunstan, D.W. Breaking up workplace sitting time with intermittent standing bouts improves fatigue and musculoskeletal discomfort in overweight/obese office workers. *Occupational and Environmental Medicine* **2014**, *71*, 765.
Bergouignan, A.; Legget, K.T.; De Jong, N.; Kealey, E.; Nikolovski, J.; Groppel, J.L.; Jordan, C.; O'Day, R.; Hill, J.O.; Bessesen, D.H. Effect of frequent interruptions of prolonged sitting on self-perceived levels of energy, mood, food cravings and cognitive function. *The international journal of behavioral nutrition and physical activity* **2016**, *13*, 113, doi:10.1186/s12966-016-0437-z.
30. Dunstan, D.W.; Kingwell, B.A.; Larsen, R.; Healy, G.N.; Cerin, E.; Hamilton, M.T.; Shaw, J.E.; Bertovic, D.A.; Zimmet, P.Z.; Salmon, J., et al. Breaking Up Prolonged Sitting Reduces Postprandial Glucose and Insulin Responses. *Diabetes Care* **2012**, *35*, 976-983, doi:10.2337/dc11-1931.
31. Peddie, M.C.; Bone, J.L.; Rehrer, N.J.; Skeaff, C.M.; Gray, A.R.; Perry, T.L. Breaking prolonged sitting reduces postprandial glycemia in healthy, normal-weight adults: a randomized crossover trial. *The American journal of clinical nutrition* **2013**, *98*, 358-366, doi:10.3945/ajcn.112.051763.
32. van Dijk, J.W.; Venema, M.; van Mechelen, W.; Stehouwer, C.D.; Hartgens, F.; van Loon, L.J. Effect of moderate-intensity exercise versus activities of daily living on 24-hour blood glucose homeostasis in male patients with type 2 diabetes. *Diabetes Care* **2013**, *36*, 3448-3453, doi:10.2337/dc12-2620.
33. Blankenship, J.M.; Granados, K.; Braun, B. Effects of subtracting sitting versus adding exercise on glycemic control and variability in sedentary office workers. *Applied physiology, nutrition, and metabolism = Physiologie appliquee, nutrition et metabolisme* **2014**, *39*, 1286-1293, doi:10.1139/apnm-2014-0157.
34. Buckley, J.P.; Mellor, D.D.; Morris, M.; Joseph, F. Standing-based office work shows encouraging signs of attenuating post-prandial glycaemic excursion. *Occup Environ Med* **2014**, *71*, 109-111, doi:10.1136/oemed-2013-101823.

35. Holmstrup, M.; Fairchild, T.; Keslacy, S.; Weinstock, R.; Kanaley, J. Multiple short bouts of exercise over 12-h period reduce glucose excursions more than an energy-matched single bout of exercise. *Metabolism - Clinical and Experimental 63*, 510-519, doi:10.1016/j.metabol.2013.12.006.
36. Bailey, D.P.; Locke, C.D. Breaking up prolonged sitting with light-intensity walking improves postprandial glycemia, but breaking up sitting with standing does not. *Journal of science and medicine in sport* **2015**, *18*, 294-298, doi:10.1016/j.jsams.2014.03.008.
37. Larsen, R.N.; Kingwell, B.A.; Robinson, C.; Hammond, L.; Cerin, E.; Shaw, J.E.; Healy, G.N.; Hamilton, M.T.; Owen, N.; Dunstan, D.W. Breaking up of prolonged sitting over three days sustains, but does not enhance, lowering of postprandial plasma glucose and insulin in overweight and obese adults. *Clinical science (London, England : 1979)* **2015**, *129*, 117-127, doi:10.1042/cs20140790.
38. Dempsey, P.C.; Blankenship, J.M.; Larsen, R.N.; Sacre, J.W.; Sethi, P.; Straznicky, N.E.; Cohen, N.D.; Cerin, E.; Lambert, G.W.; Owen, N., et al. Interrupting prolonged sitting in type 2 diabetes: nocturnal persistence of improved glycaemic control. *Diabetologia* **2017**, *60*, 499-507, doi:10.1007/s00125-016-4169-z.
39. Dempsey, P.C.; Larsen, R.N.; Sethi, P.; Sacre, J.W.; Straznicky, N.E.; Cohen, N.D.; Cerin, E.; Lambert, G.W.; Owen, N.; Kingwell, B.A., et al. Benefits for Type 2 Diabetes of Interrupting Prolonged Sitting With Brief Bouts of Light Walking or Simple Resistance Activities. *Diabetes Care* **2016**, *39*, 964-972, doi:10.2337/dc15-2336.
40. Hadgraft, N.T.; Lynch, B.M.; Clark, B.K.; Healy, G.N.; Owen, N.; Dunstan, D.W. Excessive sitting at work and at home: Correlates of occupational sitting and TV viewing time in working adults. *BMC Public Health* **2015**, *15*, 899, doi:10.1186/s12889-015-2243-y.
41. Kim, J.; Shin, W. How to Do Random Allocation (Randomization). *Clinics in Orthopedic Surgery* **2014**, *6*, 103-109, doi:10.4055/cios.2014.6.1.103.
42. Booth, M. Assessment of physical activity: an international perspective. *Res Q Exerc Sport* **2000**, *71*, doi:10.1080/02701367.2000.11082794.
43. Grant, P.M.; Ryan, C.G.; Tigbe, W.W.; Granat, M.H. The validation of a novel activity monitor in the measurement of posture and motion during everyday activities. *British journal of sports medicine* **2006**, *40*, 992-997, doi:10.1136/bjsm.2006.030262.
44. Godfrey, A.; Culhane, K.M.; Lyons, G.M. Comparison of the performance of the activPAL Professional physical activity logger to a discrete accelerometer-based activity monitor. *Medical engineering & physics* **2007**, *29*, 930-934, doi:10.1016/j.medengphy.2006.10.001.
45. Kozey-Keadle, S.; Libertine, A.; Lyden, K.; Staudenmayer, J.; Freedson, P.S. Validation of wearable monitors for assessing sedentary behavior. *Med Sci Sports Exerc* **2011**, *43*, 1561-1567, doi:10.1249/MSS.0b013e31820ce174.
46. Ryan, C.G.; Grant, P.M.; Tigbe, W.W.; Granat, M.H. The validity and reliability of a novel activity monitor as a measure of walking. *British journal of sports medicine* **2006**, *40*, 779-784, doi:10.1136/bjsm.2006.027276.
47. Zhang, Y.; Li, H.; Keadle, S.; Matthews, C.E.; Carroll, R. PAactivPAL: Summarize Daily Physical Activity from 'activPAL' Accelerometer Data. **2016**.
48. Actigraph Support Center. What is the difference among the Energy Expenditure Algorithms? https://actigraph.desk.com/customer/en/portal/articles/2515835-what-is-the-difference-among-the-energy-expenditure-algorithms-. Availabe online: (accessed on 05 may).
49. Sasaki, J.E.; John, D.; Freedson, P.S. Validation and comparison of ActiGraph activity monitors. *Journal of science and medicine in sport* **2011**, *14*, 411-416, doi:10.1016/j.jsams.2011.04.003.
50. Tremblay, M.S.; Aubert, S.; Barnes, J.D.; Saunders, T.J.; Carson, V.; Latimer-Cheung, A.E.; Chastin, S.F.M.; Altenburg, T.M.; Chinapaw, M.J.M.; Altenburg, T.M., et al. Sedentary Behavior Research Network (SBRN) – Terminology Consensus Project process and outcome. *International Journal of Behavioral Nutrition and Physical Activity* **2017**, *14*, 75, doi:10.1186/s12966-017-0525-8.
51. Reips, U.D.; Funke, F. Interval-level measurement with visual analogue scales in Internet-based research: VAS Generator. *Behavior research methods* **2008**, *40*, 699-704.
52. McNair, D.M. *Manual profile of mood states*; Educational & Industrial testing service: 1971.
53. Parry, S.; Straker, L. The contribution of office work to sedentary behaviour associated risk. *BMC Public Health* **2013**, *13*, doi:10.1186/1471-2458-13-296.
54. Haskell, W.L.; Lee, I.M.; Pate, R.R.; Powell, K.E.; Blair, S.N.; Franklin, B.A.; Macera, C.A.; Heath, G.W.; Thompson, P.D.; Bauman, A. Physical activity and public health: updated recommendation for adults from

the American College of Sports Medicine and the American Heart Association. *Med Sci Sports Exerc* **2007**, *39*, 1423-1434, doi:10.1249/mss.0b013e3180616b27.
55. Hill, J.O. Understanding and addressing the epidemic of obesity: an energy balance perspective. *Endocr Rev* **2006**, *27*, doi:10.1210/er.2006-0032.
56. Bergouignan, A.; Latouche, C.; Heywood, S.; Grace, M.S.; Reddy-Luthmoodoo, M.; Natoli, A.K.; Owen, N.; Dunstan, D.W.; Kingwell, B.A. Frequent interruptions of sedentary time modulates contraction- and insulin-stimulated glucose uptake pathways in muscle: Ancillary analysis from randomized clinical trials. *Scientific reports* **2016**, *6*, 32044, doi:10.1038/srep32044.
57. Black, A.E. Physical activity levels from a meta-analysis of doubly labeled water studies for validating energy intake as measured by dietary assessment. *Nutrition reviews* **1996**, *54*, 170-174.
58. Clemes, S.A.; O'Connell, S.E.; Edwardson, C.L. Office workers' objectively measured sedentary behavior and physical activity during and outside working hours. *J Occup Environ Med* **2014**, *56*, doi:10.1097/jom.0000000000000101.
59. Bessesen, D.; Bergouignan, A. Behavior Change Strategies for Increasing Exercise and Decreasing Sedentary Behaviors in Diabetes. In *Diabetes and Exercise: From Pathophysiology to Clinical Implementation*, Reusch, M.D.J.E.B., Regensteiner, P.M.A.B.A.J.G., Stewart, E.D.F.M.F.K.J., Veves, M.D.D.A., Eds. Springer International Publishing: Cham, 2018; pp. 201-219.
60. Clemes, S.A.; Parker, R.A. Increasing our understanding of reactivity to pedometers in adults. *Med Sci Sports Exerc* **2009**, *41*, 674-680, doi:10.1249/MSS.0b013e31818cae32.
61. Behrens, T.K.; Dinger, M.K. Motion sensor reactivity in physically active young adults. *Res Q Exerc Sport* **2007**, *78*, 1-8, doi:10.1080/02701367.2007.10762229.
62. Davis, R.E.; Loprinzi, P.D. Examination of Accelerometer Reactivity Among a Population Sample of Children, Adolescents, and Adults. *Journal of physical activity & health* **2016**, *13*, 1325-1332, doi:10.1123/jpah.2015-0703.

References cited in chapter III

[1] G. N. Healy, C. E. Matthews, D. W. Dunstan, E. A. H. Winkler, et N. Owen, « Sedentary time and cardio-metabolic biomarkers in US adults: NHANES 2003–06 », *Eur Heart J*, vol. 32, n° 5, p. 590-597, mars 2011, doi: 10.1093/eurheartj/ehq451.

[2] M. Kangas, A. Konttila, P. Lindgren, I. Winblad, et T. Jämsä, « Comparison of low-complexity fall detection algorithms for body attached accelerometers », *Gait & Posture*, vol. 28, n° 2, p. 285-291, août 2008, doi: 10.1016/j.gaitpost.2008.01.003.

[3] G. M. Weiss, J. L. Timko, C. M. Gallagher, K. Yoneda, et A. J. Schreiber, « Smartwatch-based activity recognition: A machine learning approach », in *2016 IEEE-EMBS International Conference on Biomedical and Health Informatics (BHI)*, 2016, p. 426-429, doi: 10.1109/BHI.2016.7455925.

[4] G. Plasqui, A. G. Bonomi, et K. R. Westerterp, « Daily physical activity assessment with accelerometers: new insights and validation studies », *Obesity Reviews*, vol. 14, n° 6, p. 451-462, 2013, doi: 10.1111/obr.12021.

[5] M. Garnotel *et al.*, « Prior automatic posture and activity identification improves physical activity energy expenditure prediction from hip-worn triaxial accelerometry », *Journal of Applied Physiology*, vol. 124, n° 3, p. 780-790, nov. 2017, doi: 10.1152/japplphysiol.00556.2017.

[6] B. Sun, Y. Wang, et J. Banda, « Gait Characteristic Analysis and Identification Based on the iPhone's Accelerometer and Gyrometer », *Sensors*, vol. 14, n° 9, p. 17037-17054, sept. 2014, doi: 10.3390/s140917037.

[7] M. Awais, S. Mellone, et L. Chiari, « Physical activity classification meets daily life: Review on existing methodologies and open challenges », in *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2015, p. 5050-5053, doi: 10.1109/EMBC.2015.7319526.

[8] D. Figo, P. C. Diniz, D. R. Ferreira, et J. M. Cardoso, « Preprocessing Techniques for Context Recognition from Accelerometer Data », *Personal Ubiquitous Comput.*, vol. 14, n° 7, p. 645–662, oct. 2010, doi: 10.1007/s00779-010-0293-9.

[9] L. Bao et S. S. Intille, « Activity Recognition from User-Annotated Acceleration Data », in *Pervasive Computing*, 2004, p. 1-17.

[10] J. Wang, Y. Chen, S. Hao, X. Peng, et L. Hu, « Deep learning for sensor-based activity recognition: A survey », *Pattern Recognition Letters*, vol. 119, p. 3-11, mars 2019, doi: 10.1016/j.patrec.2018.02.010.

[11] N. Y. Hammerla, S. Halloran, et T. Plötz, « Deep, Convolutional, and Recurrent Models for Human Activity Recognition Using Wearables », in *IJCAI*, 2016.

[12] T. Hur, J. Bang, T. Huynh-The, J. Lee, J.-I. Kim, et S. Lee, « Iss2Image: A Novel Signal-Encoding Technique for CNN-Based Human Activity Recognition », *Sensors*, vol. 18, n° 11, p. 3910, nov. 2018, doi: 10.3390/s18113910.

[13] T. Huynh-The, C.-H. Hua, et D.-S. Kim, « Visualizing Inertial Data For Wearable Sensor Based Daily Life Activity Recognition Using Convolutional Neural Network* », in *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2019, p. 2478-2481, doi: 10.1109/EMBC.2019.8857366.

[14] F. Chollet, *Deep Learning with Python*, 1st edition. Shelter Island, New York: Manning Publications, 2017.

[15] E. Zdravevski *et al.*, « Improving Activity Recognition Accuracy in Ambient Assisted Living Systems by Automated Feature Engineering », *IEEE Access*, vol. PP, p. 1-1, mars 2017, doi: 10.1109/ACCESS.2017.2684913.

[16] T. Hastie, R. Tibshirani, et J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction, Second Edition*, 2nd edition. New York, NY: Springer, 2016.

[17] H. Leutheuser, D. Schuldhaus, et B. M. Eskofier, « Hierarchical, Multi-Sensor Based Classification of Daily Life Activities: Comparison with State-of-the-Art Algorithms Using a Benchmark Dataset », *PLOS ONE*, vol. 8, n° 10, p. e75196, oct. 2013, doi: 10.1371/journal.pone.0075196.

[18] O. Banos, M. Damas, H. Pomares, F. Rojas, B. Delgado-Marquez, et O. Valenzuela, « Human activity recognition based on a sensor weighting hierarchical classifier », *Soft Comput*, vol. 17, n° 2, p. 333-343, févr. 2013, doi: 10.1007/s00500-012-0896-3.

[19] S. Zhang, P. Mccullagh, C. Nugent, et H. Zheng, « Activity Monitoring Using a Smart Phone's Accelerometer with Hierarchical Classification », présenté à Proceedings - 2010 6th International Conference on Intelligent Environments, IE 2010, 2010, p. 158-163, doi: 10.1109/IE.2010.36.

[20] V. T. van Hees *et al.*, « Separating Movement and Gravity Components in an Acceleration Signal and Implications for the Assessment of Human Daily Physical Activity », *PLOS ONE*, vol. 8, n° 4, p. e61691, avr. 2013, doi: 10.1371/journal.pone.0061691.

[21] O. Banos, J.-M. Galvez, M. Damas, H. Pomares, et I. Rojas, « Window Size Impact in Human Activity Recognition », *Sensors (Basel)*, vol. 14, n° 4, p. 6474-6499, avr. 2014, doi: 10.3390/s140406474.

[22] F. Pedregosa *et al.*, « Scikit-learn: Machine Learning in Python », *J. Mach. Learn. Res.*, vol. 12, n° null, p. 2825–2830, nov. 2011.

[23] M. Abadi *et al.*, « TensorFlow: a system for large-scale machine learning », in *Proceedings of the 12th USENIX conference on Operating Systems Design and Implementation*, Savannah, GA, USA, 2016, p. 265–283.

[24] G. C. Cawley et N. L. C. Talbot, « On Over-fitting in Model Selection and Subsequent Selection Bias in Performance Evaluation », *Journal of Machine Learning Research*, vol. 11, nº Jul, p. 2079-2107, 2010.

[25] Y. Chen, M. Guo, et Z. Wang, « An improved algorithm for human activity recognition using wearable sensors », in *2016 Eighth International Conference on Advanced Computational Intelligence (ICACI)*, 2016, p. 248-252, doi: 10.1109/ICACI.2016.7449833.

[26] R. P. Troiano, J. J. McClain, R. J. Brychta, et K. Y. Chen, « Evolution of accelerometer methods for physical activity research », *Br J Sports Med*, vol. 48, nº 13, p. 1019-1023, juill. 2014, doi: 10.1136/bjsports-2014-093546.

[27] J. Ortiz Laguna, A. G. Olaya, et D. Borrajo, « A Dynamic Sliding Window Approach for Activity Recognition », in *User Modeling, Adaption and Personalization*, Berlin, Heidelberg, 2011, p. 219-230, doi: 10.1007/978-3-642-22362-4_19.

[28] C. F. Martindale, S. Sprager, et B. M. Eskofier, « Hidden Markov Model-Based Smart Annotation for Benchmark Cyclic Activity Recognition Database Using Wearables », *Sensors (Basel)*, vol. 19, nº 8, avr. 2019, doi: 10.3390/s19081820.

[29] A. Nazábal, P. García-Moreno, A. Artés-Rodríguez, et Z. Ghahramani, « Human Activity Recognition by Combining a Small Number of Classifiers », *IEEE Journal of Biomedical and Health Informatics*, vol. 20, nº 5, p. 1342-1351, sept. 2016, doi: 10.1109/JBHI.2015.2458274.

[30] R. Jurca, T. Cioara, I. Anghel, M. Antal, C. Pop, et D. Moldovan, « Activities of Daily Living Classification using Recurrent Neural Networks », in *2018 17th RoEduNet Conference: Networking in Education and Research (RoEduNet)*, 2018, p. 1-4, doi: 10.1109/ROEDUNET.2018.8514124.

References cited in chapter IV

1. Kohl HW, Craig CL, Lambert EV, Inoue S, Alkandari JR, Leetongin G, et al. The pandemic of physical inactivity: global action for public health. The Lancet. 2012;380:294–305.

2. Warburton DER, Nicol CW, Bredin SSD. Health benefits of physical activity: the evidence. CMAJ. 2006;174:801–9.
3. Tremblay MS, Aubert S, Barnes JD, Saunders TJ, Carson V, Latimer-Cheung AE, et al. Sedentary Behavior Research Network (SBRN) – Terminology Consensus Project process and outcome. International Journal of Behavioral Nutrition and Physical Activity. 2017;14:75.
4. Du Y, Liu B, Sun Y, Snetselaar LG, Wallace RB, Bao W. Trends in Adherence to the Physical Activity Guidelines for Americans for Aerobic Activity and Time Spent on Sedentary Behavior Among US Adults, 2007 to 2016. JAMA Netw Open. 2019;2:e197597.
5. McCormack GR, Shiell A. In search of causality: a systematic review of the relationship between the built environment and physical activity among adults. International Journal of Behavioral Nutrition and Physical Activity. 2011;8:125.
6. Smith M, Hosking J, Woodward A, Witten K, MacMillan A, Field A, et al. Systematic literature review of built environment effects on physical activity and active transport - an update and new findings on health equity. Int J Behav Nutr Phys Act. 2017;14:158.
7. Bauman AE, Reis RS, Sallis JF, Wells JC, Loos RJ, Martin BW. Correlates of physical activity: why are some people physically active and others not? The Lancet. 2012;380:258–271.
8. Duncan MJ, Badland HM, Mummery WK. Applying GPS to enhance understanding of transport-related physical activity. Journal of Science and Medicine in Sport. 2009;12:549–56.
9. Krenn PJ, Titze S, Oja P, Jones A, Ogilvie D. Use of global positioning systems to study physical activity and the environment: a systematic review. Am J Prev Med. 2011;41:508–15.
10. Chaix B, Kestens Y, Duncan DT, Brondeel R, Méline J, El Aarbaoui T, et al. A GPS-Based Methodology to Analyze Environment-Health Associations at the Trip Level: Case-Crossover Analyses of Built Environments and Walking. Am J Epidemiol. 2016;184:579–89.
11. Wheeler BW, Cooper AR, Page AS, Jago R. Greenspace and children's physical activity: A GPS/GIS analysis of the PEACH project. Preventive Medicine. 2010;51:148–52.
12. Lachowycz K, Jones AP, Page AS, Wheeler BW, Cooper AR. What can global positioning systems tell us about the contribution of different types of urban greenspace to children's physical activity? Health & Place. 2012;18:586–94.
13. Chaix B, Kestens Y, Bean K, Leal C, Karusisi N, Meghiref K, et al. Cohort profile: residential and non-residential environments, individual activity spaces and cardiovascular risk factors and diseases–the RECORD Cohort Study. International Journal of Epidemiology. 2012;41:1283–1292.
14. Chaix B, Méline J, Duncan S, Merrien C, Karusisi N, Perchoux C, et al. GPS tracking in neighborhood and health studies: a step forward for environmental exposure assessment, a step backward for causal inference? Health Place. 2013;21:46–51.
15. Chaix B, Kestens Y, Perchoux C, Karusisi N, Merlo J, Labadi K. An interactive mapping tool to assess individual mobility patterns in neighborhood studies. Am J Prev Med. 2012;43:440–50.
16. Chaix B. Mobile Sensing in Environmental Health and Neighborhood Research. Annual Review of Public Health. 2018;39:367–84.
17. Healy GN, Wijndaele K, Dunstan DW, Shaw JE, Salmon J, Zimmet PZ, et al. Objectively measured sedentary time, physical activity, and metabolic risk: the Australian Diabetes, Obesity and Lifestyle Study (AusDiab). Diabetes Care. 2008;31:369–371.
18. Healy GN, Matthews CE, Dunstan DW, Winkler EAH, Owen N. Sedentary time and cardio-metabolic biomarkers in US adults: NHANES 2003–06. European Heart Journal. 2011;32:590–597.
19. Carson V, Tremblay MS, Chaput J-P, Chastin SFM. Associations between sleep duration, sedentary time, physical activity, and health indicators among Canadian children and youth using compositional analyses. Appl Physiol Nutr Metab. 2016;41 6 (Suppl. 3):S294–302.
20. Bellettiere J, Winkler EAH, Chastin SFM, Kerr J, Owen N, Dunstan DW, et al. Associations of sitting accumulation patterns with cardio-metabolic risk biomarkers in Australian adults. PLOS ONE. 2017;12:e0180119.
21. Chastin SFM, Palarea-Albaladejo J, Dontje ML, Skelton DA. Combined Effects of Time Spent in Physical Activity, Sedentary Behaviors and Sleep on Obesity and Cardio-Metabolic Health Markers: A Novel Compositional Data Analysis Approach. PLOS ONE. 2015;10:e0139984.
22. Debache I, Bergouignan A, Chaix B, Sneekes EM, Thomas F, Sueur C. Associations of Sensor-Derived Physical Behavior with Metabolic Health: A Compositional Analysis in the Record Multisensor Study. International Journal of Environmental Research and Public Health. 2019;16:741.

23. Kondo MC, Fluehr JM, McKeon T, Branas CC. Urban Green Space and Its Impact on Human Health. International Journal of Environmental Research and Public Health. 2018;15:445.
24. Giles-Corti B, Vernez-Moudon A, Reis R, Turrell G, Dannenberg AL, Badland H, et al. City planning and population health: a global challenge. Lancet. 2016;388:2912–24.
25. Saelens BE, Handy SL. Built Environment Correlates of Walking: A Review. Med Sci Sports Exerc. 2008;40 7 Suppl:S550–66.
26. Sugiyama T, Neuhaus M, Cole R, Giles-Corti B, Owen N. Destination and route attributes associated with adults' walking: a review. Med Sci Sports Exerc. 2012;44:1275–86.
27. Giles-Corti B, Bull F, Knuiman M, McCormack G, Van Niel K, Timperio A, et al. The influence of urban design on neighbourhood walking following residential relocation: Longitudinal results from the RESIDE study. Social Science & Medicine. 2013;77:20–30.
28. Knuiman MW, Christian HE, Divitini ML, Foster SA, Bull FC, Badland HM, et al. A Longitudinal Analysis of the Influence of the Neighborhood Built Environment on Walking for TransportationThe RESIDE Study. Am J Epidemiol. 2014;180:453–61.
29. Jennings V, Gaither CJ, Gragg RS. Promoting Environmental Justice Through Urban Green Space Access: A Synopsis. https://home.liebertpub.com/env. 2012. doi:10.1089/env.2011.0007.
30. Schüle SA, Bolte G. Interactive and Independent Associations between the Socioeconomic and Objective Built Environment on the Neighbourhood Level and Individual Health: A Systematic Review of Multilevel Studies. PLOS ONE. 2015;10:e0123456.
31. Riva M, Gauvin L, Apparicio P, Brodeur J-M. Disentangling the relative influence of built and socioeconomic environments on walking: The contribution of areas homogenous along exposures of interest. Social Science & Medicine. 2009;69:1296–305.
32. MacDonald JM, Stokes RJ, Cohen DA, Kofner A, Ridgeway GK. The Effect of Light Rail Transit on Body Mass Index and Physical Activity. American Journal of Preventive Medicine. 2010;39:105–12.
33. Chaix B, Kestens Y, Duncan S, Merrien C, Thierry B, Pannier B, et al. Active transportation and public transportation use to achieve physical activity recommendations? A combined GPS, accelerometer, and mobility survey study. International Journal of Behavioral Nutrition and Physical Activity. 2014;11:124.
34. Patterson R, Webb E, Millett C, Laverty AA. Physical activity accrued as part of public transport use in England. J Public Health (Oxf). 2019;41:222–30.
35. Chaix B, Kestens Y, Duncan DT, Brondeel R, Méline J, El Aarbaoui T, et al. A GPS-Based Methodology to Analyze Environment-Health Associations at the Trip Level: Case-Crossover Analyses of Built Environments and Walking. Am J Epidemiol. 2016;184:579–89.
36. Browning M, Lee K. Within What Distance Does "Greenness" Best Predict Physical Health? A Systematic Review of Articles with GIS Buffer Analyses across the Lifespan. International Journal of Environmental Research and Public Health. 2017;14:675.
37. Home | Portail open data de L'IAU île-de-France. http://data.iau-idf.fr/. Accessed 23 Sep 2019.
38. Base permanente des équipements | Insee. https://www.insee.fr/fr/metadonnees/source/serie/s1161. Accessed 23 Sep 2019.
39. Insee - Institut national de la statistique et des études économiques. https://www.insee.fr/fr/accueil. Accessed 23 Sep 2019.
40. Explore — Open data Île-de-France Mobilités. https://data.iledefrance-mobilites.fr/explore/?sort=modified. Accessed 23 Sep 2019.
41. Aitchison J. The Statistical Analysis of Compositional Data. London, UK, UK: Chapman & Hall, Ltd.; 1986.
42. R Core Team. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing; 2013. http://www.R-project.org/.
43. Bivand R, Rundel C. rgeos: Interface to Geometry Engine - Open Source ('GEOS'). 2017.
44. Pebesma E. sf: Simple Features for R}. 2018.
45. Bates D, Mächler M, Bolker B, Walker S. Fitting Linear Mixed-Effects Models Using lme4. Journal of Statistical Software. 2015;67:1–48.

References cited in chapter V

. A report on physical activity for health from the four home countries' Chief Medical Officers. Dep Heal [Internet]. 2011; Available from: https://www.sportengland.org/media/2928/dh_128210.pdf

2. Hallal PC, Andersen LB, Bull FC, Guthold R, Haskell W, Ekelund U. Global physical activity levels: surveillance progress, pitfalls, and prospects. Lancet [Internet]. 2012 Jul;380(9838):247–57. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0140673612606461

3. Owen N, Sugiyama T, Eakin EE, Gardiner PA, Tremblay MS, Sallis JF. Adults' Sedentary Behavior. AMEPRE. 2011;41(2):189–96.

4. Andersen E, Ekelund U, Anderssen SA. Effects of Reducing Sedentary Time on Glucose Metabolism in Immigrant Pakistani Men. Med Sci Sports Exerc. 2015;47(4):775–81.

5. Healy GN, Wijndaele K, Dunstan DW, Shaw JE, Salmon J, Zimmet PZ, et al. Objectively Measured Sedentary Time , Physical Activity , and Metabolic Risk. 2008;31(2).

6. Wanner M, Götschi T, Martin-Diener E, Kahlmeier S, Martin BW. Active transport, physical activity, and body weight in adults a systematic review. Am J Prev Med [Internet]. 2012;42(5):493–502. Available from: http://dx.doi.org/10.1016/j.amepre.2012.01.030

7. Chaix B, Kestens Y, Duncan S, Merrien C, Thierry B, Pannier B, et al. Active transportation and public transportation use to achieve physical activity recommendations? A combined GPS, accelerometer, and mobility survey study. Int J Obes Relat Metab Disord J Int Assoc Study Obesity. 2014;11(1):1–11.

8. Costa S, Ogilvie D, Dalton A, Westgate K, Brage S, Panter J. Quantifying the physical activity energy expenditure of commuters using a combination of global positioning system and combined heart rate and movement sensors. Prev Med (Baltim) [Internet]. 2015;81:339–44. Available from: http://dx.doi.org/10.1016/j.ypmed.2015.09.022

9. Wang S, Chen C, Ma J. Accelerometer based transportation mode recognition on mobile phones. APWCS 2010 - 2010 Asia-Pacific Conf Wearable Comput Syst. 2010;44–6.

10. Zhang Z, Poslad S. A new post correction algorithm (PoCoA) for improved transportation mode recognition. Proc - 2013 IEEE Int Conf Syst Man, Cybern SMC 2013. 2013;1512–8.

11. Chaix B. Mobile Sensing in Environmental Health and Neighborhood Research. Ssrn. 2018;

12. Wanigatunga AA, Ferrucci L, Schrack JA. Physical activity fragmentation as a potential phenotype of accelerated aging. Oncotarget. 2019;10(8):807–9.

13. Almanza E, Jerrett M, Dunton G, Seto E, Ann Pentz M. A study of community design, greenness, and physical activity in children using satellite, GPS and accelerometer data. Health Place [Internet]. 2012 Jan;18(1):46–54. Available from: https://linkinghub.elsevier.com/retrieve/pii/S135382921100164X

14. Rodríguez DA, Cho GH, Evenson KR, Conway TL, Cohen D, Ghosh-Dastidar B, et al. Out and about: Association of the built environment with physical activity behaviors of adolescent females. Heal Place. 2012;18(1):55–62.

15. Maher C, Olds T, Onywera V, Tudor-Locke C, Maia J, Hu G, et al. Inequality in physical activity, sedentary behaviour, sleep duration and risk of obesity in children: a 12-country study. Obes Sci Pract. 2018;4(3):229–37.

16. Chastin SFM, Granat MH. Methods for objective measure, quantification and analysis of sedentary behaviour and inactivity. Gait Posture. 2010;31(1):82–6.

17. Newton RL, Han H, Zderic T, Hamilton M. The Energy Expenditure of Sedentary Behavior: A Whole Room Calorimeter Study. PLoS One. 2013;8(5).

18. El Aarbaoui T, Méline J, Brondeel R, Chaix B. Short-term association between personal exposure to noise and heart rate variability: The RECORD MultiSensor Study. Environ Pollut. 2017;231:703–11.

19. Chaix B, Kestens Y, Bean K, Leal C, Karusisi N, Meghiref K, et al. Cohort profile: Residential and non-residential environments, individual activity spaces and cardiovascular risk factors and diseases-The RECORD cohort study. Int J Epidemiol. 2012;41(5):1283–92.

20. Wolf J, Schönfelder S, Samaga U, Oliveira M, Axhausen KW. Eighty weeks of global positioning system traces: Approaches to enriching trip information. Transp Res Rec. 2004;(1870):46–54.

21. R Core Team. A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2014.

22. Chaix B, Benmarhnia T, Kestens Y, Brondeel R, Perchoux C, Gerber P, et al. Combining sensor tracking with a GPS-based mobility survey to better measure physical activity in trips: Public transport generates walking. Int J Behav Nutr Phys Act. 2019;16(1):1–13.

23. Rissel C, Curac N, Greenaway M, Bauman A. Physical activity associated with public transport use-a review and modelling of potential benefits. Int J Environ Res Public Health. 2012;9(7):2454–78.

24. Morabia A, Mirer FE, Amstislavski TM, Eisl HM, Werbe-Fuentes J, Gorczynski J, et al. Potential health impact of switching from car to public transportation when commuting to work. Am J Public Health. 2010;100(12):2388–91.

25. Goldstein H, Healy MJR, Rasbash J. Multilevel time series models with applications to repeated measures data. Stat Med. 1994;13(16):1643–55.

26. Litman T. Transportation and Public Health. Annu Rev Public Health [Internet]. 2013 Mar 18;34(1):217–33. Available from: http://www.annualreviews.org/doi/10.1146/annurev-publhealth-031912-114502

27. Sahlqvist S, Song Y, Ogilvie D. Is active travel associated with greater physical activity? The contribution of commuting and non-commuting active travel to total physical activity in adults. Prev Med (Baltim) [Internet]. 2012;55(3):206–11. Available from: http://dx.doi.org/10.1016/j.ypmed.2012.06.028

28. Besser LM, Dannenberg AL. Walking to Public Transit. Int J Obes Relat Metab Disord J Int Assoc Study Obesity. 2005;29(4):273–80.

Université de Strasbourg

Isaac DEBACHE

Relationships between Urban Environment, Physical Activity, and Health

Résumé

Cette thèse vise à éclairer les liens complexes entre l'activité physique, l'environnement urbain et la santé. Elle propose des outils analytiques permettant une caractérisation détaillée de l'activité physique au sens large (posture, intensité, fractionnement) et des outils algorithmiques pour la dériver à partir de capteurs électroniques (accéléromètres, gyromètres) posés sur des patients en conditions de vie libre.

Une fois la mesure et la caractérisation de l'activité physique établies, le lien entre ses différents aspects et des indices de la santé métabolique sont explorés dans le cadre d'études conduites sur des populations de patients dont les comportements physiques ont été monitorés en conditions de vie libre. Les résultats indiquent que la réduction du temps d'inactivité par le cumul de "bouts" d'activité d'intensité légère ou de posture verticale (debout) tout le long de la journée est importante au maintien des individus en bonne santé, et ce indépendamment de la pratique régulière d'exercices physiques d'intensité élevée, qui avait été l'objet principal des recherches précédentes.

A la lumière de ces conclusions, les liens causaux entre des caractéristiques de l'environnement urbain (espaces verts, qualité du réseau de transport en commun) et la réduction du temps d'inactivité sous étudiés sur une population équipée des accéléromètres et des récepteurs GPS en conditions de vie libre. Les résultats suggèrent qu'un aménagement urbain approprié peut effectivement réduire le temps inactif et les risques sanitaires qu'il implique.

Résumé en anglais

The present thesis aims at shedding light on the complex relationships between physical activité, urban environment and health. It develops analytical tools for a detailed characterization of physical activity in broad sense (posture, intensity, segmentation) and algorithms for deriving it from sensors (accelerometers, gyroscopes) worn by monitored patients in free living conditions.

Once the measuring and characterization of physical activity have been established, the relationships between its different aspects and markers of metabolic health are explored in a population of patients, whose physical behavior has been monitored in free living conditions. Results suggest that reducing inactivity time through the cumulation of bouts of light-intensity activity and standing posture throughout the day is important for maintaining individuals in good health, regardless of the level of regular high-intensity exercising, which had drawn most of the focus of past research.

In light of these conclusions, the causal links between features of urban environment (greenness, quality of public transportation) and reducing inactivity time are inferred from a monitored population of patients equipped with accelerometers and GPS trackers in free living conditions. Results suggest that an adapted urban design can effectively reduce the time spent in health-damaging physical inactivity.