Patterns of sedentary and active time accumulation are associated with mortality in US adults: The NHANES study

Junrui Di, Andrew Leroux, Jacek Urbanek, Ravi Varadhan, Adam P. Spira, Jennifer Schrack, Vadim Zipunnikov

1 Department of Biostatistics, Johns Hopkins Bloomberg School of Public Health; 2 Department of Oncology, Johns Hopkins School of Medicine; 3 Johns Hopkins Center on Aging and Health; 4 Department of Mental Health, Johns Hopkins Bloomberg School of Public Health; 5 Department of Psychiatry and Behavioral Sciences, Johns Hopkins School of Medicine; 6 Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health

Corresponding author:
Junrui Di
Department of Biostatistics
Johns Hopkins Bloomberg School of Public Health
615 N Wolfe St. E3039
Baltimore, MD, 21205
Tel: (410)955-4394
Email: jdi2@jhu.edu
Abstract

**Purpose:** Sedentary behavior has become a public health pandemic and has been associated with a variety of comorbidities including cardiovascular disease, type 2 diabetes, and some cancers. Previous studies have also shown that excessive amount of sedentary behavior is associated with all-cause mortality. However, no studies investigated whether patterns of sedentary and active time accumulation are associated with mortality independently of total sedentary and total active times. This study addresses this question by i) comparing several analytical ways to quantify patterns of both sedentary and active time accumulation through metrics of fragmentation of objectively-measured physical activity and ii) exploring the association of these metrics with all-cause mortality in a nationally representative US sample of elderly adults.

**Methods:** The accelerometry data of 3400 participants aged 50 to 84 in the National Health and Nutrition Examination Survey 2003-2006 cohorts were analyzed. Ten fragmentation metrics were calculated to quantify the duration of sedentary and active bouts: average bout duration, Gini index, average hazard, between-state transition probability, and the parameter of power law distribution. The association of these fragmentation metrics with all-cause mortality followed through December 31, 2011 was assessed with survey-weighted Cox proportional hazard models.

**Results:** In models adjusted for age, sex, race/ethnicity, education, body mass index, common comorbidities, and total sedentary/active time, four fragmentation metrics were associated with lower mortality risk: average active bout duration (HR=0.72 for 1SD increase, 95% CI = 0.59-0.88), Gini index for active bouts (HR = 0.75, 95% CI = 0.64-0.86), the parameter of power law distribution for sedentary bouts (HR = 0.75, 95% CI = 0.63-0.90), and sedentary-to-active
transition probability (HR = 0.77, 95% CI = 0.61-0.96), and four fragmentation metrics were associated with higher mortality risk: the active-to-sedentary transition probability (HR = 1.40, 95% CI=1.23-1.58), the parameter of power law distribution for active bouts (HR = 1.33, 95% CI = 1.16-1.52), average hazard for durations of active bouts (HR = 1.32, 95% CI = 1.18-1.48), and average sedentary bout duration (HR =1.07, 95% CI = 1.01-1.13). After sensitivity analysis, average sedentary bout duration and sedentary-to-active transition probability became insignificant.

**Conclusion:** Longer average duration of active bouts, a lower probability of transitioning from active to sedentary behavior, and a higher normalized variability of active bout durations were strongly negatively associated with all-cause mortality independently of total active time. A larger proportion of longer sedentary bouts were positively associated with all-cause mortality independently of total sedentary time. The results also suggested a nonlinear association of average active bout duration with mortality that corresponded to the largest risk increase in subjects with average active bout duration less than 3 minutes.

**Key Words:** objectively-measured physical activity, accelerometry, sedentary behavior, fragmentation, mortality, NHANES.
Introduction

Sedentary behavior is a significant risk factor for a wide range of chronic diseases, comorbidities, and mortality [1–10]. As people age, sedentary behaviors increase [7,11,12]. However, little is known about the detailed changes in daily sedentary patterns that accompany these shifts. A better understanding of temporal patterns of sedentary behavior in the characteristics of daily activities may clarify the role sedentary behaviors play in the progression of disability and disease, and thus provide more relevant public health recommendations regarding aging and diseases. Recently, the proliferation of wearable accelerometers has provided researchers with high-resolution, continuous activity data. As a result, accelerometer-measured physical activity (PA) offers the potential for both exploring detailed patterns of sedentary behaviors and providing more accurate estimates of overall sedentary behaviors, which have traditionally been underestimated by subjective methods [13,14].

Frequently, studies quantify sedentary behavior via an absolute amount (total sedentary minutes per day) or proportion (percentage) of waking hours spent sedentary [7,11,15,16]. Isotemporal substitution (ITS) model has been recently proposed to examine effects of replacing sedentary behavior with light PA (LiPA) and moderate-to-vigorous PA (MVPA) on weight, cardiovascular disease biomarkers, and mortality [2,17–19]. Compositional data analysis (CoDA) has been applied recently [20–22] to study the combined effects of time spent in LiPA and MVPA, sedentary behaviors and sleep while taking into account the codependence between those behaviors due to the finite time during a day. CoDA considers the time budget composition of the day without encountering issues of spurious correlations and collinearity. Both ITS and CoDA study the effect of allocation of the 24-hour budget between time spent in sedentary behavior, LiPA, MVPA and sleep and do not take into account how those total times have been
accumulated. Several methods have been recently proposed to quantify the patterns of sedentary
time accumulation to better understand how these patterns affect health and functional status.
Conceptually, these methods segment objectively-measured daily activity into alternating bouts
of sedentary and active time and the patterns are quantified via summaries of duration of
frequency of switching between sedentary and active bouts.

From a statistical perspective, these methods can be grouped into two categories:
nonparametric and parametric. Nonparametric approaches do not impose any distributional
assumptions and summarize the distribution of bout durations via average duration, variability of
durations, or describe properties of durations via hazard function. Paraschiv-Ionescu [23] used
the area above the cumulative distribution function curve (AAC) of bout durations to study the
relationship between chronic pain and PA. Mathematically, AAC is equivalent to the average
bout duration. Healy et al. [5,24–26] and Chastin et al. [27] both proposed to use the reciprocal
of average sitting or sedentary bout duration, which is shown in this paper to be related to the
sedentary-to-active transition probability. The Gini index, a normalized measure of the
variability of durations of sedentary bouts, has been proposed and applied by Chastin [13]. Lim
[28] proposed to use two non-parametric summaries, $k_{ar}$ and $k_{ra}$, to quantify the frequency of
switching from active to resting (sedentary) and from resting (sedentary) to active state and
defined them as the constant level of corresponding hazard functions.

Parametric approaches assume that bout durations follow a specific probability
distribution, and then estimate the parameters that characterize the chosen distribution
[13,23,27,29]. The most popular distribution used to quantify patterns of sedentary time
accumulation is the power law distribution [13,23]. The parameter of power law has
demonstrated stronger associations with clinical outcomes the total sedentary time in many
clinical populations [13,23,27]. Another choice is the exponential distribution, a commonly used
distribution for time-to-event data [30]. Nakamura [29] modeled the duration of
active periods using the stretched exponential distribution. Paraschiv-Ionescu [23] compared the
performance of several heavy tail distributions to model sedentary and active duration, including
the lognormal and the double Pareto distribution (Pareto2).

This work studies patterns of sedentary and active time accumulation with the
fragmentation metrics outlined above using accelerometry data from 50 year and older
participants of 2003-2006 National Health and Nutrition Examination Survey (NHANES). It
compares the fragmentation metrics from both a practical and a statistical perspective and
investigates whether these metrics are associated with mortality independently of total sedentary
and total active time.

Methods

Study Populations and Measures

The National Health and Nutrition Examination Survey (NHANES) is a stratified,
multistage, probabilistic sample representative of the civilian non-institutionalized U.S.
population, described in detail elsewhere [31]. Fragmentation metrics were calculated from
NHANES 2003-2004 and 2005-2006 waves, in which NHANES recruited a representative sub-
sample aged 6 years and older to objectively evaluate PA using accelerometry. Since a survival
analysis was conducted, subjects under the age of 50 years were excluded from the analysis to
minimize potential biases induced by including younger individuals who were susceptible to
genetic and/or atypical virulent diseases not related to habitual physical activity. NHANES data
have been linked to death records from the National Death Index through December 31, 2011. If
a participant is deceased, the duration of time in months between the NHANES examination and
death is provided. Accidental deaths were excluded from the analysis. Demographic and
comorbidity information that are available included, age, gender, race, education level, smoking
status (never, former, current), drinking status (former or current drinker, heavy drinker,
moderate drinker, non-drinker), body mass index (BMI; kg/m²), mobility difficulty (yes, no),
diagnosis of diabetes, coronary heart disease, congestive heart failure, stroke, and cancer.
Subjects with missing covariates were also excluded. In total, 3400 participants aged 50 to 84
years who fulfilled the inclusion criteria remained in the analysis, where 1773 were from cohorts
2003-2004, and 1627 were from cohorts 2005-2006. There were 542 reported deaths in the
sample over an average of 6.4 follow-up years (7.2 for cohort 2003-2004, and 5.6 for cohort
2005-2006).

Physical activity was measured with the ActiGraph AM-7164 accelerometer (ActiGraph,
LLC, Fort Walton Beach, Florida). The device was placed on an elasticized fabric belt, custom-
fitted for each subject, and worn on the right hip. Participants were instructed to remove the belt
while sleeping, bathing, and swimming. The monitors were programmed to record activity
counts in successive 1-minute epochs for up to 7 consecutive days. Non-wear time was defined
as any interval of 90 minutes or longer in which all count values were 0 with allowance for up to
two minutes of non-zero counts between 1 and 99 [32,33]. Subjects were included if they had at
least one valid day of accelerometer data, defined as at least 10 hours of wear time and all
NHANES generated quality flags for the data were deemed valid [34].

Following previous work [35,36], minutes with activity counts < 100 are defined as
sedentary, and minutes with activity counts ≥100 are defined as active without further
distinguishing between light physical activity (LIPA) and moderate to vigorous physical activity
(MVPA). A sedentary or active bout is defined as being sedentary or active for at least one minute.

**Notations**

First, necessary notations are introduced. To define fragmentation metrics, the following notations are introduced: the duration of the longest active bout is denoted by $D_A$, the number of bouts of length $t$ is denoted by $n_A(t)$, the number of active bouts of length $\leq t$ is denoted by $n^A_c(t)$. Total active time can be represented as $T_A = \sum_{t=1}^{D_A} n_A(t) \cdot t$, and the total number of active bouts can be represented as $n_A = \sum_{t=1}^{D_A} n_A(t)$. Notations for sedentary bouts can be defined similarly, with all subscripts changed to “S”. When introducing and defining the fragmentation metrics, the subscripts are sometimes dropped since they are defined for both sedentary and active. For simplicity of exposition, most of the conceptual derivations below will assume continuous $t$.

Necessary adjustments are required to account for the fact that observed durations are integers. In this context, the continuity assumption is reasonable as a minute-level epoch assumption may be relaxed and a much finer resolution can be considered for high-frequency accelerometry data.

**Nonparametric metrics**

We now review nonparametric fragmentation metrics.

1. **Average duration**

The simplest and the most intuitive fragmentation metric is the average bout duration. Some previous studies have demonstrated usage of average bout durations. Paraschiv-Ionescu used the area above the cumulative distribution function curve (AAC) of bout durations (which statistically is equivalent to average duration) to study the relationship between chronic pain and PA [23]. To study how “breaks” of sedentary was associated with metabolic risk, Healy et al.
considered the mean duration of the breaks [25]. Similarly, Lynch et al calculated the average length of active and sedentary bouts and studied their association with breast cancer biomarkers [37]. Average bout duration is denoted by \( \mu \) and estimated as

\[
\mu = \frac{T}{n} = \frac{\sum_{t=1}^{D} n(t) \cdot t}{\sum_{t=1}^{D} n(t)} \quad (\text{Equation 1})
\]

respectively.

2. Normalized variability

Second most common method for nonparametrically summarizing a distribution is to estimate its variability. The Gini index was originally developed in econometrics to study the statistical dispersion of the distribution of incomes [38] and was used by Chastin et al [13] as a measure of the accumulation of sedentary time. Here, the Gini index is denoted by \( g \), and defined and estimated as

\[
g = \frac{\sum_{t_1=1}^{D} \sum_{t_2=1}^{D} n(t_1)n(t_2)|t_1-t_2|}{2n^2 \mu} \quad (\text{Equation 2})
\]

respectively. The Gini index can be seen as a measure of (absolute, not squared) variability of bout durations normalized by the average duration. It can be shown that \( g \) is bounded between 0 and 1 [38]. When Gini index is close to 1, it indicates that total time is accumulated via a small number of longer bouts. Conversely, when Gini index is close to 0, it indicates that all bouts contribute equally to total time.

3. Average hazard

Lim et al. [28] proposed two metrics \( k_{ra} \) and \( k_{ar} \) to quantify the frequency of switching from active to resting (sedentary) and from resting (sedentary) to active states, respectively, and
studied their association with cognitive impairment in older adults [39]. Similarly, they adopted this concept to study sleep fragmentation and its effects on Alzheimer’s Disease (AD) [40,41]. The estimation procedure focused on estimation of “transition probability” as a function of bout duration, applied smoothing, and identified a range of durations when the “transition probability” function flattened out. The constant values of the “transition probability” from those “constancy” ranges have defined $k_{ra}$ and $k_{ar}$.

From a statistical point of view, the “transition probability” function constructed by Lim et al. is exactly the hazard function widely used and studied in survival analysis. Modeling the hazard function is a principal approach for analyzing time-to-event data. The hazard function can be seen as the instantaneous probability of failure at time $t$ given that the subject has survived until time $t$ [30,42,43]. In this sense, the hazard is a measure of risk: the greater the hazard, the greater the risk of failure. Thus, in terms of dichotomous sedentary-active states, the hazard function can be used to study the probability of transitioning from sedentary to active or from active to sedentary state.

There are a few statistical reservations and modeling limitations to directly use the proposal of Lim et al. [28]. First and the most important, their proposal does not provide a well-defined estimand of interest. Second, the estimation of $k_{ra}$ and $k_{ar}$ tries to identify the “constancy range” of hazard function - a restrictive and hard-to-verify assumption. Third, the proposal employs LOWESS smoothing to identify the “constancy” range, thus, requires “ad hoc” decisions to make. To avoid these limitations, but stay close to the original proposal, we suggest the average hazard (AH) as the primary estimand to non-parametrically summarize the hazard function as a function of bout duration. Next, we outline our proposed procedure to estimate the AH.
For observed durations $t_1, \ldots, t_n$, it is assumed that there are $m$ unique values, which are denoted in increasing order by $\{t_{n_1}, t_{n_2}, \ldots, t_{n_m}\}$. Then, hazard rates can be estimated at these distinct bout durations nonparametrically using the Nelson-Aalen approach [43] while treating all bout durations as non-censored (i.e., all active bouts will transition to sedentary bouts, and vice versa)

$$h(t_{n_i}) = \frac{n(t_{n_i})}{n - n^c(t_{n_{i-1}})} \quad \text{(Equation 3)}$$

where $i = 1, \ldots, m$. Note that Nelson-Aalen approach does not estimate the hazard function at time points that are not observed. The AH is then estimated as

$$\bar{h} = \frac{1}{m} \sum_{t \in D} h(t), \quad \text{(Equation 4)}$$

where $D = \{t_{n_1}, t_{n_2}, \ldots, t_{n_m}\}$. Larger $\bar{h}_S$ indicates a higher frequency of transitioning from sedentary to active state; and larger $\bar{h}_A$ indicates a higher frequency of transitioning from active to sedentary state.

4. Transition probability

Several studies have considered to use the reciprocal of average bout duration. To investigate the efficacy of a multicomponent intervention to reduce office workers’ sitting time, Healy et al. considered “sit to stand transition” which is defined as “number of sit-to stand transitions per hour of sitting” [24]. Chastin [27] studied the accumulation of sedentary time in older adults with obesity and low muscle strength by using the ratio of the number of sedentary bouts divided by the total sedentary time.

Here the reciprocal of average bout duration is denoted by $\lambda$. It can be shown (see proof in S1 Text) that $\lambda$ is equal to between-states (i.e., sedentary-to-active or active-to-sedentary) transition
probability, namely

$$\lambda_S = \frac{n_S}{T_S} = \Pr(y_{t+1} = 1 | y_t = 0) \quad (\text{Equation 5})$$

$$\lambda_A = \frac{n_A}{T_A} = \Pr(y_{t+1} = 0 | y_t = 1), \quad (\text{Equation 6})$$

where $$y_t$$ is the indicator of activity type (sedentary or active) at time epoch $$t$$, and 0 and 1 denote sedentary and active states, respectively. Larger/smaller values of $$\lambda$$ correspond to more/less frequent switching between the states and as a result, may indicate more/less fragmented activity pattern. It is important to note that larger/smaller values of $$\lambda$$ also correspond to shorter/longer average bout duration.

From a parametric point of view, if the bout durations follow an exponential distribution, frequently used to model time-to-event data [30], $$\lambda$$ is exactly the parameter that fully defines the exponential distribution with the following cumulative distribution function (CDF):

$$F(t) = 1 - e^{-\lambda t}, \quad t \geq 0 \quad (\text{Equation 7}).$$

Note that any parametric assumptions should be validated through an appropriate goodness-of-fit tests.

**Parametric metrics**

In this section, Power law, one of the most popular and widely used distribution to model bout duration, is discussed.

The CDF of a Power law or Type I Pareto distribution is defined as

$$F(t) = 1 - \left(\frac{t}{t_{\text{min}}}\right)^{-\alpha+1}, \quad t \geq t_{\text{min}}, \quad (\text{Equation 8})$$
where \( t_{\text{min}} \) is the lowest bout duration, the scaling parameter \( \alpha \) summarizes information about the pattern of accumulation of total sedentary or total active time. Larger values of \( \alpha \) indicates that subject tends to accumulate sedentary (active) time with a larger proportion of shorter sedentary (active) bouts.

The parameter \( \alpha \) is typically estimated through maximum likelihood [44,45]. For the analysis of bout durations, the distribution is assumed to be discrete and the approximation to MLE estimator is used as follows

\[
\hat{\alpha} = 1 + n \left[ \sum_{i=1}^{n} \ln \frac{t_i}{\hat{t}_{\text{min}} - 0.5} \right]^{-1} \text{ (Equation 9)}
\]

[44]. Here \( \hat{t}_{\text{min}} \) is the lower bound and can be estimated by the minimum bout length. Most studies [13,46] assess the goodness-of-fit of a power law through a visual assessment of the histogram of bout durations on doubly logarithmic plot. While Paraschiv-Ionescu employed a formal statistical goodness-of-fit test based on Anderson-Darling test [23].

Table 1 summarizes all five metrics described above along with their estimation method and interpretation.

<table>
<thead>
<tr>
<th>Metrics</th>
<th>Interpretation</th>
<th>Estimation</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \mu )</td>
<td>Average duration</td>
<td>( T ) ( \frac{n}{n} )</td>
</tr>
<tr>
<td>( g )</td>
<td>Normalized variability</td>
<td>( \sum_{t_1=1}^{D} \sum_{t_2=1}^{D} n(t_1)n(t_2)</td>
</tr>
<tr>
<td>( \bar{h} )</td>
<td>Average hazard</td>
<td>( \bar{h} = \frac{1}{m} \sum_{t \in D} h(t) )</td>
</tr>
</tbody>
</table>
### Statistical analysis

Descriptive statistics were grouped by the survival status at the end of follow-up. For each subject, distributions of sedentary and active bouts were calculated by aggregating data from all bouts across valid days. Marginal densities for each metric were plotted by the survival status (deceased/alive). Survey-weighted Cox proportional hazard models [47–49] were fitted to model mortality. All models were adjusted for the covariates and comorbidities described in the Measures subsection. Two groups of models were fitted. Models in Group A studied the individual effect of each fragmentation metric on the relative risk of death by including one metric at a time. Models in Group B studied the individual effect of each metric on the relative risk of death independently of total sedentary/active times by including total sedentary time in the models with fragmentation metrics summarizing sedentary bouts and including total active time in the models with fragmentation metrics summarizing active bouts. For additional interpretability of the results, each fragmentation metric included in the models was standardized by subtracting population-level mean and dividing by population-level standard deviation, resulting in hazard ratios that correspond to one standard deviation change.

## Results

### Baseline characteristics
Descriptive statistics stratified by survival status at the end of follow-up period are presented in Table 2. All descriptive statistics are survey weighted to be representative of the U.S population. The average age of the participants was 64 years. Slightly more than half of participants were female (54%). Participants recorded as deceased tended to have greater daily sedentary time and lower daily active time. In addition, deceased participants tended to be older and have higher prevalence of comorbidities, mobility problems, and tobacco use.

Table 2: Baseline characteristics for all 3400 subjects, and by the end of follow-up status

<table>
<thead>
<tr>
<th>Variables</th>
<th>Status at end of follow up</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Alive</td>
<td>Deceased</td>
</tr>
<tr>
<td>n</td>
<td>2858</td>
<td>542</td>
</tr>
<tr>
<td>Age (mean (sd))</td>
<td>62.82 (9.62)</td>
<td>72.35 (10.16)</td>
</tr>
<tr>
<td>Male (%)</td>
<td>44.9</td>
<td>53.3</td>
</tr>
<tr>
<td>Education (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Less Than High School</td>
<td>30.8</td>
<td>48.7</td>
</tr>
<tr>
<td>High School</td>
<td>40.3</td>
<td>35.7</td>
</tr>
<tr>
<td>More Than High School</td>
<td>28.9</td>
<td>15.3</td>
</tr>
<tr>
<td>Missing Education</td>
<td>0</td>
<td>0.3</td>
</tr>
<tr>
<td>Race (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>79.7</td>
<td>82.1</td>
</tr>
<tr>
<td>Black</td>
<td>9.5</td>
<td>10.1</td>
</tr>
<tr>
<td>Mexican</td>
<td>4</td>
<td>3.6</td>
</tr>
<tr>
<td>Other Hispanic</td>
<td>2.5</td>
<td>1</td>
</tr>
<tr>
<td>Other Race</td>
<td>4.3</td>
<td>3.2</td>
</tr>
<tr>
<td>Alcohol History (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Never Drinker</td>
<td>13.3</td>
<td>14.1</td>
</tr>
<tr>
<td>Former Drinker</td>
<td>22.1</td>
<td>40.8</td>
</tr>
<tr>
<td>Current Drinker</td>
<td>60</td>
<td>38.7</td>
</tr>
<tr>
<td>Missing Information</td>
<td>4.6</td>
<td>6.4</td>
</tr>
<tr>
<td>Alcohol Usage (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Moderate Drinker</td>
<td>53.2</td>
<td>31.9</td>
</tr>
<tr>
<td>Heavy Drinker</td>
<td>6.7</td>
<td>6.7</td>
</tr>
<tr>
<td>Missing Information</td>
<td>40</td>
<td>61.3</td>
</tr>
<tr>
<td>Smoking (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Never Smoker</td>
<td>47.9</td>
<td>30.4</td>
</tr>
<tr>
<td>Former Smoker</td>
<td>35.7</td>
<td>48</td>
</tr>
<tr>
<td>Current Smoker</td>
<td>16.4</td>
<td>21.6</td>
</tr>
</tbody>
</table>
### Differences by survival status

Density plots of all metrics and total sedentary and active times categorized by the survival status are shown in Fig 1. These plots along with Table 2 estimates marginal associations between the fragmentation metrics and mortality status. Participants in the deceased group (red) tended to have longer sedentary time ($T_S$) and shorter active time ($T_A$), longer average sedentary bout ($\mu_S$) and shorter average active bout ($\mu_A$) durations than participants in the alive group (blue). Deceased participants were also more likely to have larger $g_S$ and smaller $g_A$, indicating that their (normalized by the mean) sedentary bout durations are more variable while their (normalized by the mean) active bout durations are less variable. Considering the average hazard metrics, alive participants were more likely to transition from sedentary to active behavior ($\alpha_S$), and deceased participants were more likely to transition from active to sedentary ($\alpha_A$).
behavior (higher $\bar{h}_A$). Deceased participants also had smaller $\lambda_S$ and higher $\lambda_A$ that corresponded to smaller chances of switching from sedentary to active behavior and higher chances of switching from active to sedentary behavior. Similarly, deceased participants tended to accumulate total active time with shorter active bouts (larger $\alpha_A$), and accumulate total sedentary time with longer sedentary bouts (smaller $\alpha_S$). Another interesting observation is that while the distributions of $\mu_S$ and $\mu_A$ exhibit considerable skewness, $\lambda_S$ and $\lambda_A$ had near-symmetric population-level distributions, a highly desirable statistical property. Note also that because of the skewness in distributions, the population averages of $\mu_S$ and $\mu_A$ are quite different from the reciprocals of the population averages of $\lambda_S$ and $\lambda_A$.

Fig 1. The estimated probability distribution functions of total sedentary time and sedentary fragmentation metrics (TOP) and total active time and active fragmentation metrics (BOTTOM) for deceased (red) and alive (blue) participants.

Pairwise scatterplots and correlations of fragmentation metrics

Fig 2 shows the pairwise scatterplots (bottom triangle) and correlations (upper triangle) between all metrics. As expected, there was a clear parabolic association shape between $\mu$ and $\lambda$ due to their definitions and estimation procedures. A parabolic relationship between $\mu$ and $\alpha$ was observed for both sedentary and active bouts. Meanwhile, $\lambda$ and $\bar{h}$ were highly positively correlated with a linear trend ($\rho = 0.84$ for sedentary bouts, and $\rho = 0.93$ for active bouts); and $\lambda$ and $g$ were highly negatively correlated with a linear trend ($\rho = -0.77$ for sedentary bouts, and $\rho = -0.92$ for active bouts). Moreover, $\lambda$ and $\alpha$ had an almost linear relationship ($\rho = 0.94$ for sedentary bouts, and $\rho = 0.97$ for active bouts). The correlation between the total sedentary time
and the five sedentary bout fragmentation metrics were equal to 0.67, 0.57, -0.74, -0.74, -0.71.

The correlations between the total active time and the five active bout fragmentation metrics were equal to 0.76, 0.74, -0.80, -0.81, -0.79. The shape and variability in the pair-wise scatterplots seem to indicate that fragmentation metrics may provide information about mortality beyond that of already provided by the total sedentary and total active times. In addition to revealing pairwise dependences between fragmentation metrics, Fig 2 allows to visually explore pairs of fragmentation metrics and their potential to separate deceased (red) from alive (blue) participants.

Fig 2: **Pairwise scatterplots (lower triangular) and correlations (upper triangular) for the fragmentation metrics for deceased (red) and alive (blue). Deceased group was plotted over the alive group.**

**Cox PH models**

Table in S1 Table shows the results of the baseline models adjusted for all covariates and comorbidities described in Measures and adjusted for the total sedentary time (Model 1) or the total active time (Model 2). The results were consistent with those reported in previous studies [35,50,51] and demonstrate that one minute increase of total active time is associated with lower mortality risk (HR = 0.99, 95% CI = 0.99-1.00), and one minute increase of total sedentary time is associated with higher mortality risk (HR = 1.002, 95% CI = 1.001-1.002).

The left panel of Table 3 shows the odds ratio based on 1 SD increase for models in Group A (unadjusted for total sedentary and total active time). All ten fragmentation metrics were significantly associated with the relative odds of mortality. Five fragmentation metrics had a negative significant association with the relative odds of mortality: \( \mu_A(\text{HR} = 0.50, 95\% \text{ CI} = \)
0.42-0.60), $g_A$ (HR = 0.61, 95% CI = 0.55-0.69), $\bar{h}_S$ (HR = 0.80, 95% CI = 0.70-0.92), $\lambda_S$ (HR = 0.67, 95% CI = 0.57-0.79), and $\alpha_S$ (HR = 0.69, 95% CI = 0.60-0.78). Conversely, the other five fragmentation metrics had a positive association with the relative odds of mortality: $\mu_S$ (HR = 1.11, 95% CI = 1.05-1.16), $g_S$ (HR = 1.18, 95% CI = 1.04-1.34), $\bar{h}_A$ (HR = 1.52, 95% CI = 1.39-1.66), $\lambda_A$ (per SD HR = 1.59, 95% CI = 1.45-1.75), and $\alpha_A$ (HR = 1.56, 95% CI = 1.42-1.72).

Table 3: The results of fully adjusted Cox proportional hazard models (HR corresponding to 1SD increase)

<table>
<thead>
<tr>
<th>Metrics</th>
<th>Group A</th>
<th>pvalues</th>
<th>Group B</th>
<th>p values</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu_S$</td>
<td>1.11$^*$ (1.05, 1.16)</td>
<td>4.88e-05</td>
<td>1.07$^*$ (1.01, 1.13)</td>
<td>2.82e-02</td>
</tr>
<tr>
<td>$\mu_A$</td>
<td>0.50$^*$ (0.42, 0.60)</td>
<td>4.26e-13</td>
<td>0.72$^*$ (0.59, 0.88)</td>
<td>1.44e-03</td>
</tr>
<tr>
<td>$g_S$</td>
<td>1.18 (1.04, 1.34)</td>
<td>9.33e-03</td>
<td>1.00 (0.85, 1.18)</td>
<td>0.97</td>
</tr>
<tr>
<td>$g_A$</td>
<td>0.61$^*$ (0.55, 0.69)</td>
<td>6.66e-16</td>
<td>0.75$^*$ (0.64, 0.86)</td>
<td>9.70e-05</td>
</tr>
<tr>
<td>$\lambda_S$</td>
<td>0.67$^*$ (0.57, 0.79)</td>
<td>8.56e-07</td>
<td>0.77$^*$ (0.61, 0.96)</td>
<td>2.07e-02</td>
</tr>
<tr>
<td>$\lambda_A$</td>
<td>1.59 (1.45, 1.75)</td>
<td>&lt;2e-16</td>
<td>1.40$^*$ (1.23, 1.58)</td>
<td>1.86e-07</td>
</tr>
<tr>
<td>$\bar{h}_A$</td>
<td>1.52$^*$ (1.39, 1.66)</td>
<td>&lt;2e-16</td>
<td>1.32$^*$ (1.18, 1.48)</td>
<td>9.04e-07</td>
</tr>
<tr>
<td>$\alpha_S$</td>
<td>0.69$^*$ (0.60, 0.78)</td>
<td>4.05e-08</td>
<td>0.75$^*$ (0.63, 0.90)</td>
<td>1.95e-03</td>
</tr>
<tr>
<td>$\alpha_A$</td>
<td>1.56 (1.42, 1.72)</td>
<td>&lt;2e-16</td>
<td>1.33 (1.16, 1.52)</td>
<td>2.85e-05</td>
</tr>
</tbody>
</table>

* Indicates significant association at 5% significance level.

HR: hazard ratio associated with one standard deviation increase
CI: confidence interval

The right panel of Table 3 shows the odds ratio based on 1 SD increase for models in Group B. Models in Group B additionally included total sedentary time in the models with sedentary bouts fragmentation metrics and included total active time in the models with active bouts fragmentation metrics. After the adjustment for the total sedentary/active times, four fragmentation metrics had a negative significant association with the relative odds of mortality: $\mu_A$ (HR=0.72, 95% CI = 0.59-0.88), $g_A$ (HR = 0.75, 95% CI = 0.64-0.86), $\alpha_S$ (HR = 0.75, 95% CI = 0.63-0.90), $\lambda_S$ (per SD HR = 0.77, 95% CI = 0.62-0.96), and four fragmentation metrics had a...
positive significant association with the relative odds of mortality: $\mu_S$ (HR = 1.07, 95% CI = 1.01-1.13), $\bar{h}_A$ (HR = 1.32, 95% CI = 1.18-1.48), $\alpha_A$ (HR = 1.33, 95% CI = 1.16-1.52), and $\lambda_A$ (HR = 1.40, 95% CI = 1.23-1.58).

**Sensitivity Analysis**

Two sensitivity analyses were conducted. The first explored possible effects of reversed causality by excluding all deaths within the first year of follow-up. The second excluded days with wear time longer than 20 hours to eliminate any bias from days when subjects wore accelerometers during sleep.

The first sensitivity analysis excluded deaths within the first year of follow-up. After the exclusion, 3334 (with 476 deaths) subjects remained and the same Cox PH models were re-estimated. The results are shown in Table 4. The exclusion of the 1-st year follow-up deaths slightly attenuated results, but has not changed neither the significance or direction of most of the associations, with the exception of $\lambda_S$ that became insignificant in the model of Group B.

**Table 4: The results of fully adjusted Cox proportional hazard models after excluding deaths within the first year of follow-up (HR corresponding to 1SD increase)**

<table>
<thead>
<tr>
<th>Metrics</th>
<th>Group A</th>
<th>Group B</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>HR (95% CI)</td>
<td>p values</td>
</tr>
<tr>
<td>$\mu_S$</td>
<td>1.11 (1.07, 1.16)</td>
<td>7.18e-07</td>
</tr>
<tr>
<td>$\mu_A$</td>
<td>0.54 (0.45, 0.66)</td>
<td>1.76e-09</td>
</tr>
<tr>
<td>$g_S$</td>
<td>1.15 (0.99, 1.34)</td>
<td>0.056</td>
</tr>
<tr>
<td>$g_A$</td>
<td>0.64 (0.57, .73)</td>
<td>4.77e-13</td>
</tr>
<tr>
<td>$\bar{h}_S$</td>
<td>0.83 (0.73, 0.95)</td>
<td>8.67e-03</td>
</tr>
<tr>
<td>$\bar{h}_A$</td>
<td>1.45 (1.32, 1.59)</td>
<td>1.48e-14</td>
</tr>
<tr>
<td>$\lambda_S$</td>
<td>0.73* (0.62, 0.85)</td>
<td>4.92e-05</td>
</tr>
<tr>
<td>$\lambda_A$</td>
<td>1.51 (1.37, 1.67)</td>
<td>&lt;2e-16</td>
</tr>
<tr>
<td>$\alpha_S$</td>
<td>0.73* (0.64, 0.84)</td>
<td>3.60e-06</td>
</tr>
<tr>
<td>$\alpha_A$</td>
<td>1.48 (1.34, 1.63)</td>
<td>7.66e-15</td>
</tr>
</tbody>
</table>
The second sensitivity analysis excluded days with wear time longer than 20 hours. “Valid” days in the original analysis is defined based on the wearing time longer than 10 hours. However, there are numbers of subject-days that had more than 20 hours of wear time, with up to 24 hours of wear per day. Including days with wear time longer than 20 hours might lead to counting sleep time as sedentary that in turn could result in biased estimated summaries for sedentary and active time. Therefore, for the second sensitivity analysis, valid day was defined as a day with wear time between 10 and 20 hours. The threshold of 20 hours has been previously used to identify “bed-time” periods in NHANES [52]. After exclusion of invalid subject-days from the original samples, 3362 (with 534 deaths) subjects remained and same Cox PH models where re-estimated. The results are shown in Table 5. Although, the direction of association and significance for majority of the metrics remained the same, both $\mu_S$ and $\lambda_S$ became insignificant.

<table>
<thead>
<tr>
<th>Metrics</th>
<th>Group A</th>
<th>Group B</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu_S$</td>
<td>1.11 (1.05, 1.16)</td>
<td>1.05 (0.99, 1.11)</td>
</tr>
<tr>
<td>$\mu_A$</td>
<td>0.52 (0.43, 0.62)</td>
<td>0.78 (0.65, 0.95)</td>
</tr>
<tr>
<td>$\lambda_S$</td>
<td>0.67 (0.58, 0.79)</td>
<td>0.79 (0.62, 1.02)</td>
</tr>
<tr>
<td>$\lambda_A$</td>
<td>1.54 (1.40, 1.70)</td>
<td>1.30 (1.14, 1.48)</td>
</tr>
<tr>
<td>$g_S$</td>
<td>1.21 (1.06, 1.38)</td>
<td>1.05 (0.90, 1.22)</td>
</tr>
<tr>
<td>$g_A$</td>
<td>0.63 (0.56, 0.71)</td>
<td>0.78 (0.68, 0.90)</td>
</tr>
<tr>
<td>$\bar{h}_S$</td>
<td>0.84 (0.74, 0.96)</td>
<td>1.03 (0.91, 1.16)</td>
</tr>
<tr>
<td>$\bar{h}_A$</td>
<td>1.49 (1.36, 1.63)</td>
<td>1.27 (1.15, 1.41)</td>
</tr>
<tr>
<td>$\alpha_S$</td>
<td>0.70 (0.61, 0.80)</td>
<td>0.79 (0.63, 1.00)</td>
</tr>
<tr>
<td>$\alpha_A$</td>
<td>1.51 (1.37, 1.66)</td>
<td>1.23 (1.08, 1.42)</td>
</tr>
</tbody>
</table>

Nonlinear effect of average active bout and average sedentary bout durations
Among all considered fragmentation metrics, average bout duration, $\mu$, is probably the most straightforward to calculate and communicate. Below, we show that, when modeled via $\lambda = 1/\mu$, the effect of $\mu$ on proportional HR is highly nonlinear and the largest risk increase is observed in subjects with average active bout duration less than 3 minutes.

To illustrate our argument, we refer to Table in S2 Table that reports estimated effect for non-standardized fragmentation metrics with respective to 1 unit increase for $\mu_A$ and $\lambda_A$, which equal to -0.26 and 3.88, respectively. The top panel in Fig 3 shows the hazard ratios with respect to a 1 minute increase of average activity bout calculated based on $\mu_A$ (dashed line) and $\lambda_A$ (solid line). When estimated from the Cox-PH model with $\mu_A$, the effect on the hazard ratio is constant and equal to $e^{-0.27} = 0.77$ (dashed line in the top panel). This implies that increasing average active bout by 1 minute always reduces hazard by approximately 23%, regardless of “original” average bout duration. On contrary, when considering the effect of increasing $\mu_A$ by 1 minute estimated in the Cox-PH model with $\lambda_A$, the nonlinear change is given by

$$\exp \left( 3.88 \left( \frac{1}{\mu_A+1} - \frac{1}{\mu_A} \right) \right) = \exp \left( -\frac{3.88}{\mu_A(\mu_A+1)} \right). \text{ (Equation 10)}$$

The solid black line on top panel of Fig 3 demonstrates this nonlinear effect on the hazard ratio. This shows that effect is drastically large and nonlinear when $\mu_A$ is small, and as $\mu_A$ gets larger, the effect flattens out. For example, if $\mu_A = 1$, then the hazard ratio corresponding to one-minute-increase is 0.14, i.e. increasing from a 1-minute to 2-minute average active bout makes hazard almost 86% (or 7 times) smaller. However, if $\mu_A = 4$, a one-minute increase in the average active bout reduces hazard ratio to 0.82, which translates to roughly 18% decrease. The observed non-linear effect is consistent with the density plots of average active bout durations.
shown at the bottom panel of Fig 3. It also shows that participants with average active bout shorter than 3 minutes are at highest risk of non-surviving.

**Fig 3.** Effect of the change in $\mu_A$ on hazard ratio. Top: Change in hazard ratios as a function of 1-minute change in $\mu_A$, based on $\mu_A$ (dashed) and $\lambda_A$ (solid). Bottom: density of $\mu_A$ for deceased (red) and alive (blue).

**Discussion**

To our knowledge, this study is the first that demonstrated that patterns of sedentary and active time accumulation are strongly associated with all-cause mortality independently of total sedentary and total active time, respectively. The major findings suggest that patterns of active time accumulation have much stronger significant associations with mortality than patterns of sedentary time accumulation. Longer average duration of active bouts, a lower probability of transitioning from active to sedentary behavior, and a higher normalized variability of active bout durations were strongly negatively associated with all-cause mortality independently of total active time. Although, the main analysis showed that a longer average duration of sedentary bouts and a lower probability of transitioning from sedentary to active behavior were positively associated with all-cause mortality, these associations became non-significant during sensitivity analysis. A larger proportion of longer sedentary bouts remained positively significantly associated with all-cause mortality during both sensitivity analyses.

The results for durations of active bouts suggest nonlinear associations of average active bout duration with mortality. The largest nonlinear risk increase was observed in subjects with average active bout duration less than 3. These results are consistent with previous studies showing positive health effects of accumulating active time in prolonged active bouts [25,37],
but provide more insight in developing further individualized guidelines of physical activity.

The results regarding sedentary bout duration may imply that subjects that accumulate their sedentary time through prolong sedentary bouts may benefit from breaking those long bouts. This is consistent with the emerging evidence that breaking up prolong sedentary time can have multiples positive effects including improvement of cardiovascular and cardiometabolic health [3,5,6,24,26,53,54].

Methodologically, this work compared various approaches to quantify patterns of sedentary time accumulation through fragmentation of accelerometry measured physical activity. Active-to-Sedentary and Sedentary-to-Active transition probabilities, co-expressed as reciprocals of average active bout and average sedentary bout durations, are easily interpretable modifiable fragmentation metrics with insightful connection to other metrics of fragmentation, and stronger associations to mortality. Importantly, the use of $\lambda$ results in a non-linear effect on hazard ratio that could potentially help with providing more subject-specific guidelines. The parameter of power law distribution, $\alpha$, has been a popular metric in a few recent studies on accumulation patterns of sedentary time [13,46,53]. Fig 2 demonstrates a close-to-linear relationship between $\alpha$ and $\lambda$ with correlation coefficients of 0.94 for sedentary bouts and 0.97 for active bouts. Nevertheless, $\alpha_S$ was the only sedentary time fragmentation metric that remained significant through the sensitivity analyses. This supports recent findings where among several sedentary fragmentation metrics only $\alpha_S$ was significantly associated with biomarkers of glucose metabolism [53].

There are a few limitations in the present study. First, the minimum bout length was defined based on the Actigraph epoch duration of 1 minute. The results may change, if the sedentary and active bouts are defined in a different way. Second, the threshold of 100 AC used
to define sedentary state is a widely accepted for the NHANES data. However, there is an on-
going debate on whether this threshold can be uniformly applied across all age groups and
genders and whether such uniform thresholds induce bias in classifying sedentary behaviors.

Therefore, it is imperative to fragmentation metrics as a function of different threshold values in
future work. Third, the non-wear criteria may overestimate the total amount of sedentary time by
including those have not complied with the protocol and wear the device during sleep time.

Finally, this study is cross-sectional, therefore, all findings are associative, not casual.

Fragmentation metrics can provide unique translatable insights into accumulation
patterns for sedentary and active time and lead to the better understanding of associations
between those patterns and health outcomes.

Acknowledgements

Competing interests: The authors have declared that no competing interest exist.

References


52. Urbanek JK, Spira A, Di J, Leroux A, Crainiceanu C, Zipunnikov V. Epidemiology of Objectively Measured Bedtime and Chronotype in the US adolescents and adults:


Supporting Information

S1 Table. Results of the baseline models adjusted for all covariates and comorbidities described in Measures and adjusted for the total sedentary time (Model 1) or the total active time (Model 2).

S2 Table. Results of Cox PH models in original scale without standardization.

S1 Text. Heuristic Proof that $\lambda$ is equivalent to the between states transition probability.

S2 Text. Expressing nonparametric metrics using the empirical CDF.
<table>
<thead>
<tr>
<th></th>
<th>$T_S$</th>
<th>0.67</th>
<th>0.57</th>
<th>-0.74</th>
<th>-0.74</th>
<th>-0.71</th>
<th>-0.49</th>
<th>-0.58</th>
<th>-0.50</th>
<th>0.49</th>
<th>0.61</th>
<th>0.63</th>
</tr>
</thead>
<tbody>
<tr>
<td>$h_S$</td>
<td>0.61</td>
<td>-0.63</td>
<td>-0.83</td>
<td>-0.84</td>
<td>-0.70</td>
<td>-0.42</td>
<td>-0.46</td>
<td>0.58</td>
<td>0.57</td>
<td>0.58</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$g_S$</td>
<td>-0.74</td>
<td>-0.77</td>
<td>-0.53</td>
<td>-0.41</td>
<td>-0.16</td>
<td>-0.13</td>
<td>0.19</td>
<td>0.20</td>
<td>0.23</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\lambda_S$</td>
<td>0.84</td>
<td>0.74</td>
<td>0.52</td>
<td>0.40</td>
<td>0.31</td>
<td>-0.27</td>
<td>-0.40</td>
<td>-0.43</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\alpha_S$</td>
<td>0.94</td>
<td>0.76</td>
<td>0.47</td>
<td>0.41</td>
<td>-0.50</td>
<td>-0.54</td>
<td>-0.57</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$T_A$</td>
<td>0.76</td>
<td>0.74</td>
<td>-0.80</td>
<td>-0.81</td>
<td>-0.79</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\mu_A$</td>
<td>0.85</td>
<td>-0.77</td>
<td>-0.87</td>
<td>-0.84</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$g_A$</td>
<td>-0.90</td>
<td>-0.92</td>
<td>-0.80</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$h_A$</td>
<td>0.93</td>
<td>0.86</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\lambda_A$</td>
<td>0.97</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\alpha_A$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- **_alive**
- **deceased**