

Original Paper

Clustering-Based Accelerometer Measures of Physical Activity Patterns in Children With Overweight or Obesity: Baseline Cross-Sectional Methodological Analysis

Hyatt Moore IV^{1,2}, PhD; Thomas N Robinson^{3,4}, MPH, MD; Alexandria Jensen², PhD; Fatma Gunturkun², PhD; K Farish Haydel³, BA; Kristopher I Kapphahn², MS; Manisha Desai², PhD

¹Naval Postgraduate School, Monterey, CA, United States

²Quantitative Science Unit, Stanford University, Stanford, CA, United States

³Department of Pediatrics, Stanford Solutions Science Lab and Division of General Pediatrics, Stanford University, Palo Alto, CA, United States

⁴Department of Medicine, Stanford Prevention Research Center, Stanford University, Palo Alto, CA, United States

Corresponding Author:

Hyatt Moore IV, PhD
Naval Postgraduate School
1 University Circle
Monterey, CA 93943
United States
Email: hyatt.moore@nps.edu

Abstract

Background: Accelerometers produce high-resolution physical activity data, but commonly used summary measures often reduce these data to total volume, intensity, or variability and may not retain interpretable temporal structure across the day. Cluster-based summaries may provide a way to characterize daily physical activity profiles while preserving information about when activity occurs.

Objective: This study evaluated whether cluster-derived accelerometer summary measures could represent daily physical activity patterns and explain variation in pediatric cardiometabolic outcomes comparably to traditional accelerometer summary metrics.

Methods: This baseline cross-sectional methodological analysis used data from 268 Latino children with overweight or obesity from low-income families participating in the Stanford GOALS trial. Participants were aged 7 to 11 years old and wore accelerometers for a minimum of 1 week according to study protocol. Valid daily activity profiles were summarized within a 7:00 AM–11:00 PM analytic window using consecutive, nonoverlapping 10-minute intervals. Daily profiles were clustered using unsupervised learning, and participant-level cluster-derived measures were created from the distribution of valid days assigned to each cluster. We compared these measures with traditional accelerometer summaries, including time spent in activity intensity states, Time Active Mean, Time Active Variability, Activity Intensity Mean, and Activity Intensity Variability. Linear regression models were used to evaluate associations with waist circumference, fasting insulin, and fasting triglycerides, adjusting for age and sex. Model performance was compared using R^2 and the Akaike information criterion.

Results: Cluster-derived measures explained a comparable proportion of variation in the 3 cardiometabolic outcomes to traditional accelerometer summary metrics. For example, the highest R^2 values among the cluster-derived measures were 25%, 11%, and 6% for waist circumference, fasting insulin, and fasting triglycerides, respectively, compared with 25%, 10%, and 6% for Time Active Mean. No single summary-measure approach consistently yielded the highest R^2 across all 3 outcomes.

Conclusions: Cluster-derived accelerometer measures provide a regression-ready approach for summarizing daily physical activity patterns while preserving interpretable clock-time structure. In this baseline analysis, these measures performed comparably to traditional accelerometer summaries while capturing temporal features not represented by conventional volume- or intensity-based metrics. Future work should evaluate external validation across diverse populations and settings and assess the use of these measures in longitudinal and intervention analyses.

Trial Registration: ClinicalTrials.gov NCT01642836; <https://clinicaltrials.gov/study/NCT01642836>

JMIR Form Res 2026;10:e87532; doi: [10.2196/87532](https://doi.org/10.2196/87532)

Keywords: accelerometry; wearable sensors; physical activity; unsupervised machine learning; cluster analysis; obesity; pediatric health; waist circumference

Introduction

Measures of Physical Activity

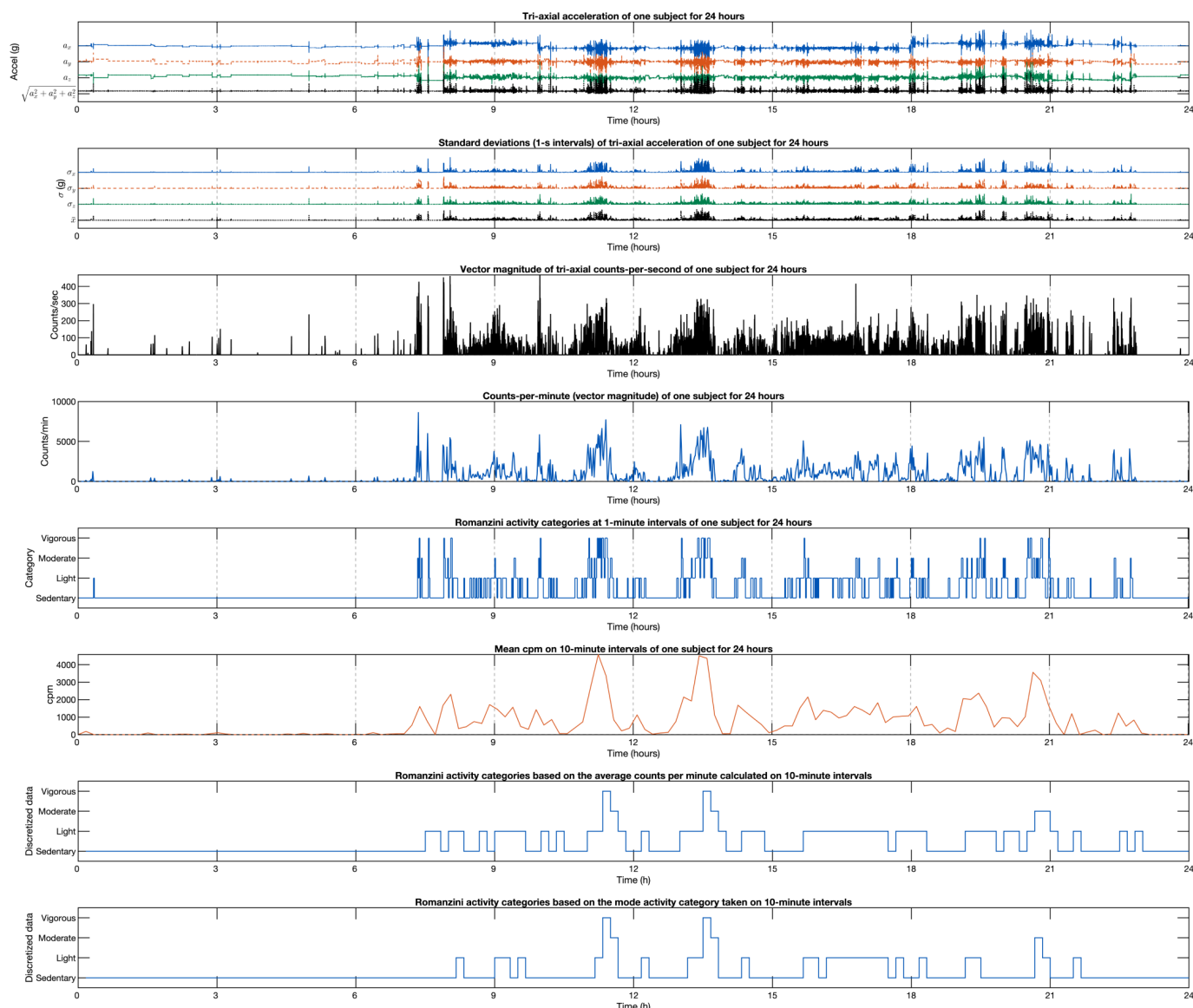
Accelerometers have been increasingly used to measure physical activity objectively [1]. The richness of accelerometer data presents both challenges and opportunities. Modern accelerometers measure acceleration in 3 dimensions at a high frequency (eg, 40 Hz or more) and, often, for long periods of time (eg, 10 d). How to process and summarize high-resolution accelerometer data is an open topic. Accelerometer-based measures of physical activity may convert the accelerometer data to describe the time spent in activity states such as “percent time spent in a sedentary state.” Others summarize the data as mean *counts* per minute (CPM), where count—a unitless measure developed by ActiGraph [2]—reflects the level of intensity of activity at a given epoch or window of time (eg, a typical epoch may cover a 15-s interval) [3-11]. Several studies have demonstrated CPM to correlate with mechanical loading and activity [8-11]. Any measure that summarizes activity based on the raw accelerometer data requires some processing. For example, the count is obtained by filtering and then summing the raw acceleration over consecutive, nonoverlapping epochs (ActiGraph LLC). Other common approaches, termed *cut-point* methods, classify activity intensity using predetermined thresholds based on accelerometer signals [8,12-16]. These approaches have been systematically reviewed in youth by Kim et al [16]. Processing involves defining epochs, transforming raw data, and specifying cut-points to classify intensity. Figure 1 illustrates raw tri-axial accelerations, in units of gravity, over a 24-hour period from a child participating in the Stanford GOALS trial [17,18] along with its corresponding count values, CPM, and activity categories using Romanzini cut-point approach [8].

The initially proprietary nature of “count” data previously garnered some criticism by those who viewed the

lack of standardization and transparency as a limitation. In their comprehensive investigation, Bai et al [19] addressed such limitations of many existing approaches that rely on counts. They proposed accelerometer-based summary measures derived from raw tri-axial acceleration signals rather than proprietary counts. These measures summarize 2 conceptual dimensions of activity: duration and intensity. Time Active Mean (TAM) and Time Active Variability (TAV) describe the mean and variability of time spent active, whereas Activity Intensity Mean (AIM) and Activity Intensity Variability (AIV) describe the mean and variability of activity intensity during active periods. These 4 measures provide a more transparent approach to the analysis of raw accelerometer data, although they still require a cut-point to distinguish activity from rest for the sample and accelerometer equipment under investigation, as described in the *Methods* section.

While these more recent methods offer greater transparency, there is still an opportunity to leverage the richness of the data to address certain types of research questions in new and innovative ways. For example, our team has been interested in questions that include the timing and patterns of physical activity that traditionally used measures may not reflect, such as whether individuals with the same amount of cumulative daily exercise who work out in the evening have more favorable cardiovascular health than those who work out in the morning. The summary measures described above do not address the pattern of physical activity signals across the day. In this work, we formally propose an alternative summary measure that leverages daily physical activity patterns preserving key temporal aspects of the accelerometer data. We evaluate the performance of the measure relative to traditional measures in a regression framework by examining the proportion of variation explained in 3 prespecified clinical outcomes.

Figure 1. Overview of accelerometer-based activity visualization of a Stanford GOALS trial participant for a 24-hour period: (1) tri-axial acceleration signals and their vector magnitude (VM), (2) SDs of acceleration signals and their mean, (3) VM of count signal at 1-second intervals, (4) VM of count signal at 1-minute intervals (*counts* per minute [CPM]), (5) Romanzini activity categories at 1-minute intervals, (6) mean CPM of VM count signal in 10-minute frames, (7) Romanzini activity categories at 10-minute intervals based on mean CPM, and (8) Romanzini activity categories at 10-minute intervals based on the mode of 1-minute intervals.



Clustering-Based Physical Activity Patterns

We are not the first to apply unsupervised machine learning to accelerometer data to gain insights into physical activity patterns. In a systematic review by Jones et al [20], 13 papers were identified that applied unsupervised machine learning to accelerometer data, and researchers have continued in this direction. For example, Nawrin et al [21] describe the need for measures that capture the temporal nature of intensively measured physical activity and demonstrate the diversity of physical activity patterns observed in a small dataset of 42 healthy individuals. Several studies have further associated patterns—discovered by machine learning techniques—with health outcomes such as obesity, cardiovascular disease (CVD), and mental health conditions within a regression framework [22–27]. These clustering-based methods provide a more granular perspective of physical activity patterns by

capturing variations in the timing and intensity of activity throughout the day, rather than solely summarizing total physical activity volume. For instance, Niemelä et al [24] identified 4 temporal physical activity clusters in midlife adults and found that these patterns were significantly related to cardiovascular disease risk. Similarly, Nawrin et al [23] and Smagula et al [27] highlighted the role of physical activity timing in metabolic health and depression risk, reinforcing the idea that physical activity patterns—not just duration or intensity—may influence key health outcomes. These studies have demonstrated the promise of unsupervised machine learning approaches for gaining insight into physical activity patterns and their role in health.

There are methodological challenges, however, in using a summary measure derived from unsupervised machine learning within a regression context. As Jones et al [20] noted, the lack of consensus on analytic approach and feature

selection limits comparability across studies and underscores the need for more transparent and standardized methods. Our study addresses this issue by formally evaluating cluster membership, derived from aligned daily activity profiles, as a physical activity summary measure in comparison with established summary measures from Romanzini et al [8] and Bai et al [19]. We further examine the sensitivity of cluster-based findings to data-processing decisions and assess performance within a regression framework focused on the proportion of variance explained in prespecified clinical outcomes. In this way, the study is designed to clarify when pattern-based measures may add value, particularly for research questions in which the timing of activity across the day is of substantive interest.

The aim of this study was to develop and evaluate clustering-based accelerometer measures derived from aligned daily activity profiles and to examine their relationships with key outcomes at baseline within the Stanford GOALS trial [17,18]. The methodological focus was to preserve interpretable clock-time structure across the day rather than reduce physical activity to aggregate volume measures alone. Accordingly, this paper is intended as a baseline cross-sectional methodological analysis, not as a longitudinal or intervention-effects study.

Methods

Study Design

This study was a secondary baseline cross-sectional methodological analysis using data from the Stanford GOALS trial, a community-based randomized controlled trial. The present

analysis focused on deriving clustering-based accelerometer measures from daily activity profiles and evaluating their relationships with key outcomes at baseline. It was designed to characterize daily activity structure at a single study time point rather than to assess longitudinal change or intervention response.

Stanford GOALS

The Stanford GOALS trial was a National Institutes of Health-funded, 3-year, community-based randomized controlled trial comparing strategies for weight control among 268 children aged 7 to 11 years with overweight or obesity from low-income, Latino (98%) families. Children had to be at or above the 85th percentile of the Centers for Disease Control and Prevention’s growth charts for their age and sex to participate [18]. The baseline sample used in the current analysis consisted of 121 male participants and 147 female participants with an average age of 9.53 (SD 1.46) years and BMI of 25.01 kg/m² (Table 1).

Participants were provided with ActiGraph GTX3+ ambulatory monitors and instructed to wear them on their hip, continually, for a period of at least 7 days, except when bathing or swimming. The devices measured acceleration (gravity) on 3 perpendicular axes at a frequency of 40 Hz. Upon the completion of the baseline study period, the devices were returned, and the data were exported to comma-separated value (.csv) files using ActiGraph’s companion software, ActiLife, which derives additional *count* measures from the raw data. These include counts for each axis of acceleration (*x*, *y*, and *z*) as well as their vector magnitude (VM, defined as $\sqrt{x^2 + y^2 + z^2}$).

Table 1. Measures from Stanford GOALS trial participants taken at baseline^a.

Measure	Male (n=121), mean (SD)	Female (n=147), mean (SD)	All (n=268), mean (SD)
Age (y)	9.49 (1.50)	9.57 (1.43)	9.53 (1.46)
Height (cm)	139.36 (9.83)	139.08 (9.67)	139.20 (9.73)
Weight (kg)	49.64 (12.53)	48.81 (12.99)	49.18 (12.77)
Waist (cm)	85.31 (10.79)	85.21 (11.03)	85.25 (10.90)
BMI (kg/m ²)	25.20 (3.79)	24.85 (4.08)	25.01 (3.95)
Insulin (μIU/mL)	13.79 (8.86)	17.45 (12.54)	15.80 (11.16)
Triglycerides (mg/dL)	90.00 (53.97)	105.1 (56.46)	98.28 (55.75)

^aWaist circumference, fasting insulin levels, and fasting triglyceride levels were selected for our analysis measures.

Ethical Considerations

The parent Stanford GOALS trial was approved by the Stanford University Administrative Panel on Human Subjects in Medical Research (protocol #19311). Parent or guardian written informed consent and Health Insurance Portability and Accountability Act authorization were obtained, and child assent was obtained as part of the parent trial. The parent trial was registered at ClinicalTrials.gov (NCT01642836). The present manuscript reports a baseline cross-sectional secondary methodological analysis of deidentified accelerometer and clinical data collected during the Stanford GOALS trial and does not evaluate longitudinal or intervention effects.

Physical Activity Data

This baseline analysis used accelerometer data collected from 268 boys and girls participating in the Stanford GOALS trial. Participants were instructed to wear ActiGraph GT3X+ accelerometers on the hip continuously for at least 7 days, except when bathing or swimming. Count data for the *x*-, *y*-, and *z*-axes and their vector magnitude were exported in consecutive, nonoverlapping 1-minute epochs. Raw acceleration signals were summarized by calculating the SDs of each axis in consecutive, nonoverlapping 1-sec intervals. Before screening for nonwear or device malfunction, the accelerometer files yielded 2322 candidate 24-hour participant-days for processing.

Data Quality

Daily accelerometer profiles were constructed within a prespecified analytic window of 07:00 to 23:00. This window was selected as part of the accelerometer-processing workflow to support aligned daily-profile analyses while retaining adequate participant-level profile availability; candidate windows and resulting profile-retention summaries are presented in Table S1 ([Multimedia Appendix 1](#)). Nonwear was identified from the count data at 1-minute resolution using Choi method [28]. Periods of device malfunction were identified in the raw gravity signal along the *x*-, *y*-, and *z*-axes using custom code by clustering the raw signal and examining outlying groups for shared signal patterns, which were subsequently confirmed by ActiGraph's support team as device malfunction. Profiles were flagged and excluded from the analyses if nonwear or device malfunction was detected within the selected analytic window. Within the 7:00 AM to 11:00 PM window, 1009 of 2322 (43.5%) candidate daily profiles were flagged during screening, leaving 1313 (56.5%) valid daily profiles; 256 (95.5%) participants retained at least 1 valid daily profile, whereas 12 (4.5%) participants retained none. Baseline clinical outcome data were complete for all 268 participants; missingness in the present analysis was therefore limited to accelerometer-derived daily profiles affected by nonwear or device malfunction within the analytic window.

Accelerometer Summary Measurements

Physical Activity Categories: Cut-Point-Based Summary Measures

CPMs were calculated from the vector-magnitude of the GT3X tri-axial count data and categorized as sedentary behavior (SB), light physical activity (LPA), and moderate-to-vigorous physical activity (MVPA), a combination of moderate physical activity (MPA) and vigorous physical activity (VPA). The cut-points for these categories were taken from Romanzini et al [8]. Every minute from 7:00 AM to 11:00 PM was categorized in this way, and the total amount of time spent in each category was tabulated per subject-day. The average amount of time spent in each physical category was derived from the tabulated activity data of each subject.

Daily Accelerometer Summary Measures

Accelerometer summary measures—TAM, TAV, AIM, and AIV—are calculated from variation measured from the accelerometers' raw signal data as outlined in Bai et al [19] for each day. Achieving this requires the use of a binary activity label, $L(t)$, which categorizes accelerometer readings at time t as either active (1) or inactive (0). This label is obtained on a 1-second interval by comparing the SD of the accelerometer signal to a cut-point, C , which is determined from the data. Bai and others did this in their data by examining density and cumulative distributions of the SD and then selecting the point at which there was a clear, visible flattening of the density. We followed the same methodology and selected $C=0.6$ based on our results (Figure S1 in [Multimedia Appendix 1](#)). TAM and TAV are, respectively,

the mean and SD of the time spent active ($L(t)=1$) for each person for each day examined. This gets at the duration of time spent active, but not the intensity of the activity. Activity intensity is defined as the SD of the original signal relative to the average standard deviation of the signal when inactive, which is calculated per subject-day. The AIM and AIV metrics are the average and SDs, respectively, of the activity intensity signal during active periods. Tables S2 and S3 ([Multimedia Appendix 1](#)) present these metrics for 3 Stanford GOALS participants to help orient readers to the dataset.

Daily Physical Activity Patterns

In this study, we propose a metric based on patterns of activity. We used *k*-means clustering to categorize each participant's daily activity profile within a 7:00 AM to 11:00 PM analytic window using average counts from consecutive, nonoverlapping 10-minute intervals. This window was selected from candidate windows as part of the prespecified processing workflow to retain broad daytime and evening coverage while maintaining participant-level profile availability (Table S1 in [Multimedia Appendix 1](#)). The 16-hour duration also allowed the sorting variants to partition the analytic day evenly into 1-, 2-, 4-, 8-, and 16-hour segments. Representing each valid day in 10-minute intervals yielded a 96-element profile vector for each day, with the number of retained profiles varying across participants. This interval length was used to preserve interpretable clock-time structure while reducing sensitivity to minute-level irregularity.

Figure 2 illustrates how these profiles can be clustered to identify common patterns of daily activity within the group. In the unsorted profiles shown in Figure 2, cluster 1 was characterized by comparatively low activity throughout the day, Cluster 2 by higher activity from the morning through late afternoon, and Cluster 3 by a pronounced late-afternoon and evening increase, peaking near 7:00 PM. Although overall activity magnitude contributed to cluster differentiation, timing also contributed: the Cluster 2 and Cluster 3 centroids crossed in the late afternoon and exhibited their highest activity during different portions of the day. We therefore interpreted the clusters as recurring activity-profile patterns rather than as definitive latent classes. The sorting variants (Sort_01, Sort_02, Sort_04, Sort_08, and Sort_16) were evaluated as a structured sensitivity analysis, spanning broader to more localized within-day sorting, to examine how partial relaxation of exact clock-time ordering affected clustering results while preserving differing degrees of temporal structure.

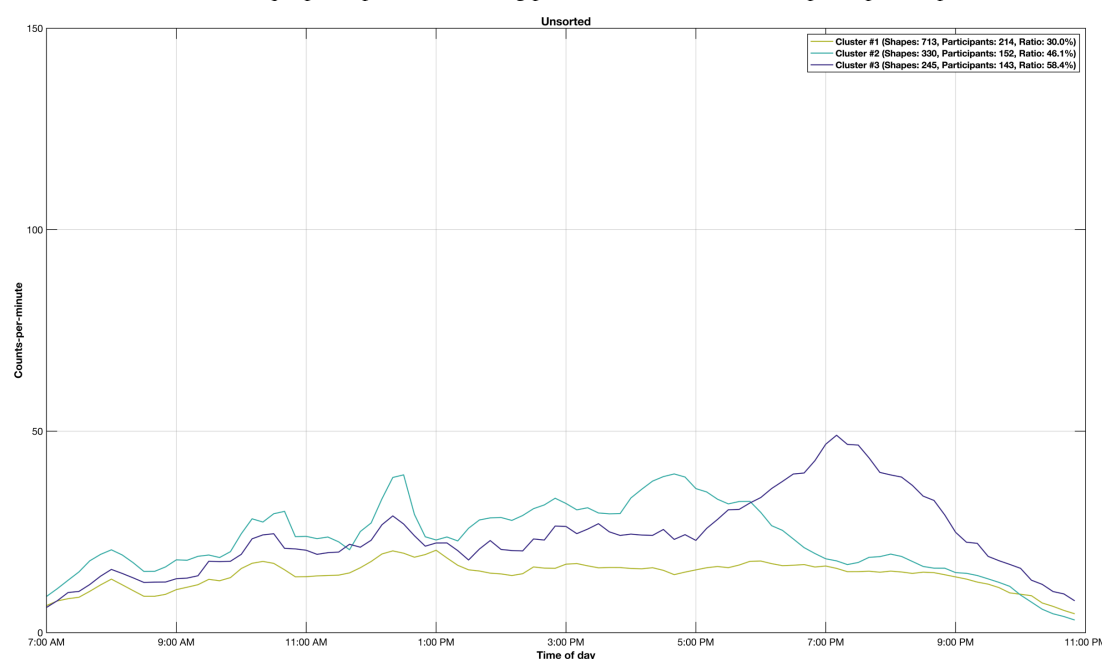
Clustering was performed using Padaco, a publicly available MATLAB-based accelerometer visualization and analysis program with clustering functionality. Daily profile vectors were clustered using *k*-means with squared Euclidean distance. This approach compares activity at corresponding positions in the represented daily profiles and therefore preserves the intended clock-time interpretation. An elastic alignment measure such as dynamic time warping was not used because the scientific objective was not to treat profiles

as equivalent after temporal warping across the day. Instead, limited tolerance for differences in the precise timing of activity was introduced explicitly through the within-segment sorting variants described below. To support the reproducibility of the clustering solution, we used Padaco's reproducibility setting, which initializes the MATLAB random number generator with `rng('default')` before clustering. The Padaco source code is available at GitHub [29].

To estimate the number of clusters, k , we used the prediction strength metric developed by Tibshirani and Walther [30], together with the recommendations of Yu et al [31]. We calculated the prediction strength statistic for up

to $k=50$ and selected the highest value of k above 0.75 or, when applicable, a local maximum or elbow in the prediction-strength curve. In this study, we used counts for comparison, but the same approach can be applied to other time-series units, including raw acceleration, monitor independent movement summaries, and related measures. To visually assess the clustering solution, we projected the 96-bin daily activity profiles into 2 dimensions using principal component analysis (PCA) and overlaid the assigned cluster labels. This PCA projection was used only as a post hoc visual diagnostic of cluster structure and was not used in the clustering algorithm or subsequent regression analyses.

Figure 2. Cluster-derived daily physical activity profiles from accelerometer count data summarized in consecutive, nonoverlapping 10-minute intervals from 7:00 AM to 11:00 PM. The profiles illustrate 3 recurring patterns identified from the unsorted daily activity data: low activity across the day, higher activity from morning through late afternoon, and relatively higher evening activity. The legend reports the number of daily profiles assigned to each cluster, the number of unique participants contributing profiles to the cluster, and the participant-to-profile ratio.

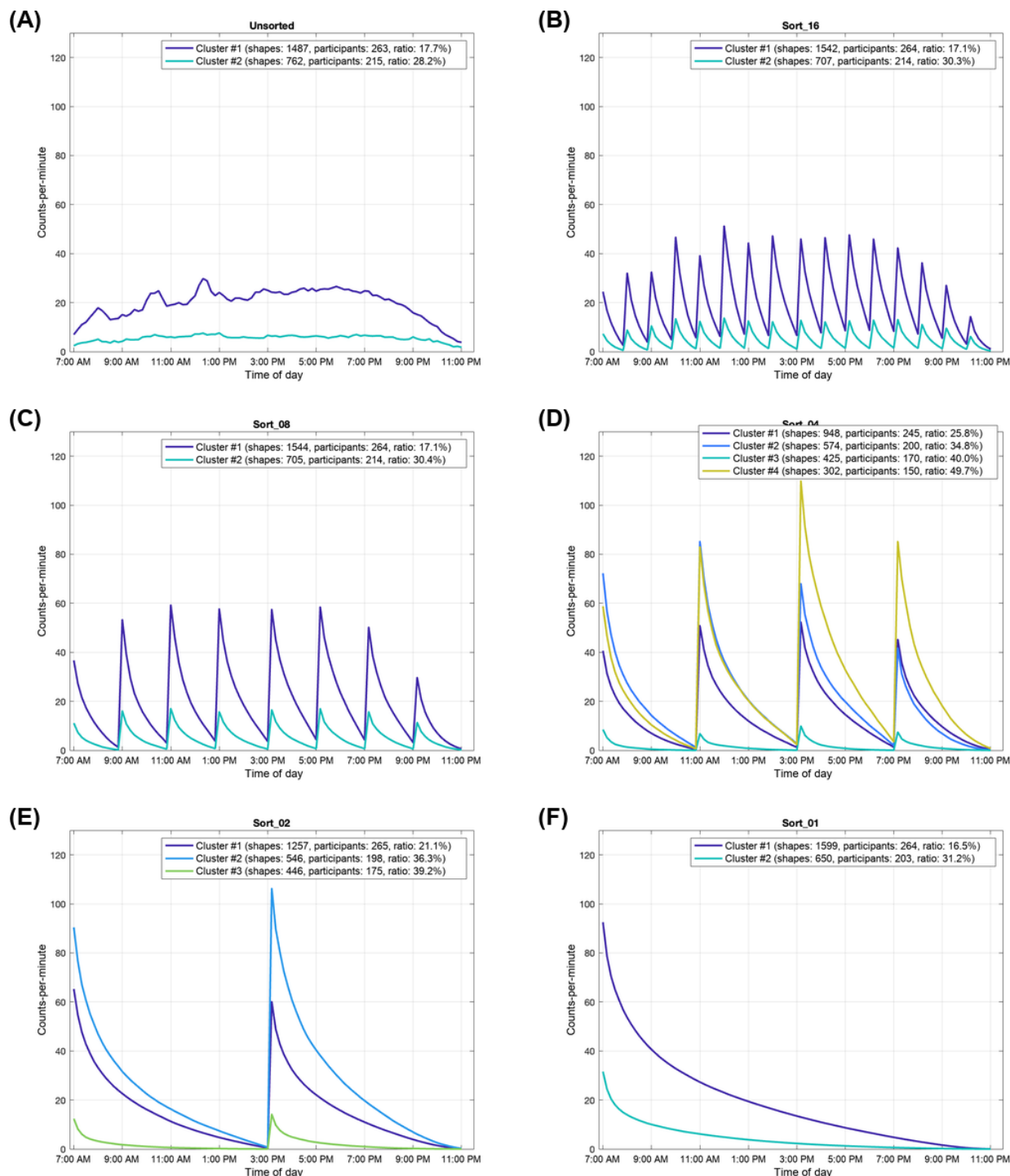


Sorting Activity Levels Prior to Clustering

Minor shifts in the precise timing of activity could cause otherwise similar days to appear dissimilar when profiles are compared at corresponding 10-minute intervals. At the same time, fully aligning temporal features across the day could remove substantively meaningful distinctions between activity occurring in different portions of the day. We therefore sorted activity values by magnitude within prespecified consecutive time segments while retaining the order of those segments before applying k-means clustering. This preprocessing allowed profiles with minor within-segment timing differences—for example, a brief morning activity peak at 8:00 AM versus 8:30 AM—to remain similar, while preserving distinctions between activity occurring in broader portions of the day. For example, it continued to distinguish a profile characterized by high morning activity and low afternoon and evening activity from one

characterized by high afternoon activity and lower morning and evening activity. A limiting version of this approach is to sort the vectors from the highest to the lowest activity values *over the whole day*, which removes clock-time ordering while retaining information about the level and distribution of activity across the analytic window. We included this whole-day sorting approach (Sort_01) as an extreme comparison condition in the analysis. To preserve progressively more temporal structure, we also sorted activity values within prespecified time segments: two 8-hour segments (Sort_02), four 4-hour segments (Sort_04), eight 2-hour segments (Sort_08), and sixteen 1-hour segments (Sort_16; Figure 3). Figure 3 shows the cluster centroids for the unsorted and sorted representations used in the analysis, illustrating the recurring activity-profile categories from which participant-level membership variables were derived.

Figure 3. Daily activity profiles found by sorting accelerometer data, within consecutive segments of the day, prior to clustering. Accelerometer count data on 10-minute intervals from 7:00 AM to 11:00 PM were sorted for each participant prior to clustering. (A) Unsorted. (B) Sort_16: daily accelerometer data are first split into 16 consecutive 1-hour segments that are individually sorted prior to clustering. (C) Sort_08: daily accelerometer data are first split into 8 consecutive 2-hour segments that are each sorted. (D) Sort_04: daily accelerometer data are split into 4 consecutive 4-hour segments, which are locally sorted prior to clustering. (E) Sort_02: daily accelerometer data are split into 2 consecutive 8-hour segments, which are each sorted prior to clustering. (F) Sort_01: accelerometer data are sorted across the entire day prior to clustering. Temporal information is discarded in this scenario.



Regression Models and Metrics for Comparing Summary Measures

Model Comparison Framework

To evaluate the proposed cluster-derived summary measure and its sorting-based variations, we compared the variation in

a given outcome explained by each of the candidate accelerometer summary measures using the coefficient of determination, R^2 . We selected 3 clinical outcomes for the primary regression analysis: waist circumference, fasting triglyceride levels, and fasting insulin levels. These outcomes were selected because they represent clinically relevant markers of central adiposity, lipid metabolism, and insulin-related

metabolic risk in pediatric obesity, consistent with pediatric cardiovascular risk frameworks [32]. Each outcome was modeled as the dependent variable in a linear regression model that included the accelerometer-derived physical activity measure of interest and adjusted for age and sex. Model performance was summarized using R^2 and Akaike information criterion (AIC). R^2 was used to describe the proportion of variation explained, whereas AIC was used to compare relative model fit among candidate models with the same outcome.

Modeling Outcome as a Function of Cut-Point-Based Summary Measures

To compare the proportion of outcome variation explained by traditional accelerometer summary measures, each clinical outcome was modeled as a function of the participant-level physical activity measure while adjusting for age and sex. The physical activity measures included the duration spent in each cut-point-based physical activity category defined by Romanzini et al [8] and the daily accelerometer summary measures proposed by Bai et al [19]. Each candidate measure was summarized across the valid observed days contributed by the participant and evaluated in a separate regression model. For example, 1 model regressed the clinical outcome on the participant's average sedentary behavior across valid observed days, while adjusting for age and sex:

$$Y_i = \beta_0 + \beta_1 \cdot A_i + \beta_2 \cdot S_i + \beta_3 \cdot \text{AverageSB}_i + \varepsilon_i$$

where Y_i is the clinical outcome for participant i , A_i is age, S_i is sex, AverageSB_i is the participant-level average sedentary behavior measure, and ε_i is the error term.

Modeling Outcome as a Function of Cluster Membership Distribution

A linear regression framework was also used to model each clinical outcome as a function of cluster membership distribution while adjusting for age and sex. For each clustering solution, participant-level cluster membership was represented as the proportion of days assigned to each cluster. Because these proportions sum to 1 across clusters for each participant, 1 cluster-membership proportion was omitted as the reference category in each regression model. Specifically, the outcome was modeled as:

$$Y_i = \beta_0 + \beta_1 \cdot A_i + \beta_2 \cdot S_i + \sum_{j=1}^{k-1} \beta_{j+2} \cdot C_{j,i} + \varepsilon_i$$

where $C_{j,i}$ is the proportion of days that participant i had in cluster j relative to the total number of valid days observed for participant i , such that $\sum_{j=1}^k C_{j,i} = 1$. For example, if the total clusters are $k=3$ and a given participant had 7 days of activity, with 4 days belonging to cluster #1, 3 days belonging to cluster #2, and zero days belonging to cluster #3, then their

cluster membership distribution would be $C_1=4/7=0.5714$, $C_2=3/7=0.4286$, and $C_3=0$. Thus, we are not constrained by the total number of clusters or differences in the total number of days available for each participant, which may vary (eg, removal due to data cleaning), because they are normalized according to the participant's total contributions (7 in the example given). This modeling approach was applied to the clusters defined by each of our preprocessing approaches: unsorted and sorted, in day-wise partitions of 1, 2, 4, 8, and 16 hours.

Results

The analytic sample for this baseline cross-sectional analysis was drawn from participants in the Stanford GOALS trial with accelerometer data meeting the prespecified daily-profile screening criteria. Under the selected 7:00 AM to 11:00 PM analytic window, 1313 valid daily profiles remained after screening, and 256 participants contributed at least 1 valid daily profile for the clustering-based analyses. Table 2 presents variance explained (R^2) and goodness of fit (AIC) corresponding to each candidate summary measure when modeling waist circumference, fasting insulin levels, and fasting triglyceride levels; values for a model using only age and sex are included for reference as well.

Across outcomes, the results were broadly comparable across candidate summary measures. No single summary-measure approach consistently yielded the highest R^2 across all 3 outcomes. For example, the proportion of variation in insulin explained ranged from only 0.0636 (LPA) to 0.1064 (VPA) for the physical activity categories. The daily accelerometer summary measures have a similar level of variation explained and range from 0.0618 (AIV) to 0.1024 (TAM). Cluster member distribution explains from 0.0872 for the unsorted clusters to 0.1051 for Sort_02. Further, the ability of all measures to explain variation varied widely depending on the clinical outcome. For example, the unsorted cluster membership distribution explains only 0.0221 of the variation observed in triglycerides, 0.0872 of insulin variation, and a proportion of 0.2300 in waist circumference variation. Across all summary measures, the proportion of variation explained was the highest for waist circumference, followed by fasting insulin and then fasting triglyceride levels.

We additionally evaluated how alternative preprocessing or sorting methods prior to clustering influenced model performance for each outcome. When modeling with cluster membership distribution, variation explained was comparable across sorting approaches. However, sorting the days in eight 2-hour sections (Sort_08) prior to clustering yielded the best results for waist circumference ($R^2=0.2516$, $\text{AIC}=1880.5$), while presorting in two 8-hour sections (Sort_02) had the highest R^2 for insulin ($R^2=0.1051$, $\text{AIC}=1949.4$) and presorting within four 4-hour sections was best for triglycerides ($R^2=0.0585$, $\text{AIC}=2763.6$).

Table 2. Age- and sex-adjusted regression models comparing physical activity summary measures as predictors of waist circumference, fasting insulin, and fasting triglycerides in Stanford GOALS participants^a.

Health outcome and statistic	Age + sex covariate	Physical activity categories			Accelerometer summary measures					Cluster membership						
		SB ^b	LPA ^c	MVPA ^d	MPA ^e	VPA ^f	TAM ^g	TAV ^h	AIM ⁱ	AIV ^j	Unsorted	Sort_01	Sort_02	Sort_04	Sort_08	Sort_16
Waist (cm)																
<i>R</i> ²	0.1911	0.2422	0.2332	0.2183	0.2039	0.2238	0.2539	0.2418	0.2349	0.2149	0.2300	0.2348	0.2281	0.2190	0.2516	0.2266
AIC ^k	1869.2	1881.7	1884.7	1889.6	1894.3	1887.8	1850.8	1854.9	1857.2	1863.7	1887.8	1886.1	1890.4	1895.4	1880.5	1888.9
Insulin																
<i>R</i> ²	0.0522	0.0786	0.0636	0.0953	0.0735	0.1064	0.1024	0.0956	0.0796	0.0618	0.0872	0.0970	0.1051	0.1011	0.0995	0.0937
AIC	1925.5	1952.8	1957.0	1948.1	1954.2	1945.0	1913.8	1915.7	1920.1	1925.0	1952.4	1949.7	1949.4	1952.5	1949.0	1950.6
Triglyceride																
<i>R</i> ²	0.0143	0.0238	0.0223	0.0256	0.0225	0.0288	0.0555	0.0484	0.0449	0.0300	0.0221	0.0227	0.0344	0.0585	0.0230	0.0221
AIC	2742.1	2766.9	2767.3	2766.4	2767.2	2765.5	2733.5	2735.2	2736.2	2740.1	2769.3	2769.2	2768.1	2763.6	2769.1	2769.3
Measures include cut-point-based activity categories, daily accelerometer summary measures, and cluster-derived measures based on 10-minute count profiles from 7:00 AM to 11:00 PM. Sorting approaches are described in the <i>Methods</i> section. Activity categories include sedentary behavior, light physical activity, moderate-to-vigorous physical activity, and vigorous physical activity.																
Daily accelerometer summary measures include Time Active Mean, Time Active Variability, Activity Intensity Mean, and Activity Intensity Variability.																
SB: sedentary behavior.																
LPA: light physical activity.																
MVPA: moderate-to-vigorous physical activity.																
MPA: moderate physical activity.																
VPA: vigorous physical activity.																
TAM: Time Active Mean.																
TAV: Time Active Variability.																
AIM: Activity Intensity Mean.																
AIV: Activity Intensity Variability.																
AIC: Akaike information criterion.																

^aMeasures include cut-point-based activity categories, daily accelerometer summary measures, and cluster-derived measures based on 10-minute count profiles from 7:00 AM to 11:00 PM. Sorting approaches are described in the *Methods* section. Activity categories include sedentary behavior, light physical activity, moderate-to-vigorous physical activity, and vigorous physical activity.

Daily accelerometer summary measures include Time Active Mean, Time Active Variability, Activity Intensity Mean, and Activity Intensity Variability.

^bSB: sedentary behavior.

^cLPA: light physical activity.

^dMVPA: moderate-to-vigorous physical activity.

^eMPA: moderate physical activity.

^fVPA: vigorous physical activity.

^gTAM: Time Active Mean.

^hTAV: Time Active Variability.

ⁱAIM: Activity Intensity Mean.

^jAIV: Activity Intensity Variability.

^kAIC: Akaike information criterion.

PCA overlays showed that the assigned clusters generally occupied different regions along continuous distributions of daily activity profiles across the preprocessing specifications considered (Supplementary Figure S2 in [Multimedia Appendix 1](#)). The distribution of variance across PC1 and PC2 changed with the preprocessing step, indicating that the sorting specification affected which aspects of the transformed daily activity profiles were emphasized in the reduced 2D space. Sort_02 showed stronger organization along PC1, whereas Sort_04 showed additional organization along PC2. These projections were used as visual diagnostics of the cluster assignments and illustrate their organization in 2 dimensions but were not used to define the clustering solutions or interpreted as evidence of discrete latent groups.

When modeling with physical activity categories, the duration of SB was best for waist circumference ($R^2=0.2422$, AIC=1881.7), while the duration of VPA was best for insulin ($R^2=0.1064$, AIC=1945.0) and triglycerides ($R^2=0.0288$, AIC=2765.5). When using Bai's physical activity summary measures, TAM performed best for modeling each outcome: waist circumference ($R^2=0.2539$, AIC=1850.8), insulin ($R^2=0.1024$ and AIC=1913.8), and triglycerides ($R^2=0.0555$, AIC=2733.5).

Discussion

This baseline cross-sectional methodological analysis evaluated a clustering-based approach for summarizing accelerometer-derived daily physical activity profiles in the Stanford GOALS trial. The primary contribution is the development and evaluation of cluster-derived participant-level measures that preserve interpretable clock-time structure across the day while remaining usable in conventional regression models. Across the cardiometabolic outcomes considered, these cluster-based measures explained a comparable proportion of variation to traditional accelerometer summary metrics. Although several physical activity summary measures showed statistically detectable associations with cardiometabolic outcomes, the absolute variance explained by these models was modest, especially for triglycerides. This finding is expected, as triglyceride levels and related cardiometabolic markers are influenced by many factors beyond accelerometer-derived activity patterns, including diet, adiposity, genetics, pubertal status, medication use, sleep, and other behavioral or clinical characteristics. Therefore, small differences in R^2 or AIC across models should not be overinterpreted. These comparisons are most useful for showing that cluster-derived measures can be incorporated into conventional regression models and can perform comparably to traditional accelerometer summaries while preserving interpretable temporal structure, rather than for establishing the stand-alone clinical predictive utility of any single measure.

Summarizing accelerometer-derived physical activity patterns in a meaningful way remains a challenge in the field. Traditional summary measures, including cut-point-based classifications and intensity-derived metrics, capture important dimensions of total activity volume, intensity, and

variability, but they generally do not retain the sequence and timing of movement across the day [19]. Recent studies using clustering and related temporal-pattern approaches have shown that the timing, regularity, and daily distribution of physical activity may provide information beyond total activity volume alone [21,22,24,27]. Other work has also linked later or more irregular activity patterns with poorer behavioral or metabolic profiles [23,27]. The present study builds on this growing literature by evaluating whether clock-time daily activity patterns can be converted into regression-ready participant-level measures and compared directly with traditional accelerometer summaries.

The methodological motivation for this approach also draws from profile-clustering applications in other time-series domains. In civil engineering and energy systems, load-shape clustering has been used to summarize recurring temporal profiles from smart meter data and to support questions about energy-use patterns and optimization [33]. Here, we adapted that profile-based logic to accelerometer-derived daily activity patterns. Each valid day was represented as an aligned activity profile, clustered into common daily patterns, and then summarized at the participant level as the distribution of days assigned to each cluster. This construction is important because it allows temporal pattern information to be retained while producing predictors that can be incorporated into familiar regression frameworks.

The comparison with traditional accelerometer summaries suggests that cluster-derived measures can complement, rather than replace, existing metrics. Traditional measures are well suited for questions about total activity, time above intensity thresholds, and overall movement volume. Cluster-derived measures are better aligned with questions about when activity occurs, whether daily profiles follow recognizable temporal shapes, and whether particular patterns of activity across the day are associated with health outcomes. In this study, the cluster-based measures did not produce uniformly larger model fit improvements, but they performed comparably while preserving an interpretable representation of daily temporal structure. This comparable performance is reassuring for an alternative measure intended to address questions about within-day activity patterns, but it does not establish added clinical utility.

The sorting analyses further illustrate how preprocessing choices can be used to align the clustering approach with different scientific questions. Without sorting, clusters represent activity patterns tied to specific clock times, and the corresponding regression coefficients describe associations between outcomes and the proportion of days spent in those clock-time activity patterns. Sorting the entire day changes the interpretation by emphasizing sustained activity levels regardless of when they occurred, making the resulting measures closer to distributional or intensity-based summaries. Sorting within time segments provides an intermediate option: it relaxes exact ordering within a period while retaining some information about whether activity occurred in the morning, afternoon, or evening. Thus, sorting is not merely a technical sensitivity analysis; it defines the degree of temporal invariance permitted by the representation.

Other pairwise similarity frameworks, including dynamic time warping or Gaussian-process-based approaches, could be combined with hierarchical clustering. Dynamic time warping defines similarity by allowing temporal features to shift to improve alignment and is valuable when overall profile shape is of primary interest and precise timing is treated as nuisance variation. In the present study, however, broad clock-time placement was part of the construct of interest, while only limited within-period timing variation was intentionally relaxed. The selected fixed-profile and within-segment sorting approaches therefore reflect the scientific objective rather than an arbitrary computational partition.

Because cluster-derived measures depend on analytic choices made before and during clustering, transparent reporting is important for interpretation and comparison across studies. Jones et al [20] emphasized that clustering applications require clear reporting of methodological choices, including the underlying machine learning method and tuning decisions. For accelerometer-derived clustering, useful reporting elements include the analytic window, preprocessing pipeline, nonwear and device-malfunction rules, exclusion criteria, missing-data handling, sorting or transformation strategy, clustering algorithm, distance or similarity measure, method for selecting the number of clusters, random seeds or initialization settings, regression model specifications, and code availability. These details do not guarantee that another sample will yield the same clusters, but they allow the analysis to be evaluated, repeated within the same dataset, and compared across studies. This distinction matters: reproducibility within the same dataset concerns whether the same pipeline can be rerun with the same data, whereas external replicability concerns whether similar patterns and associations are observed across populations, devices, monitoring durations, and study settings.

Several limitations provide context for these findings. First, the present analysis was conducted at baseline within the Stanford GOALS trial and was not designed to evaluate longitudinal changes, intervention effects, or causal relationships. Although GOALS is an intervention trial, the objective of this manuscript is methodological: to derive and evaluate cluster-based accelerometer summaries in the baseline data. Second, the analysis was based on 1 week of accelerometer monitoring per participant. Longer monitoring periods may reveal additional within-person variability and could affect the stability of cluster-derived summaries. Third, the analytic window was selected to retain broad daytime and evening coverage while reducing loss of usable profiles due to nonwear or device malfunction. This standardized representation was necessary for aligned daily-profile clustering, but it also required excluding profiles with nonwear or device malfunction within the analytic window. If missingness was related to activity behavior or cardiometabolic outcomes, this complete-case approach could affect estimated associations. These considerations support prespecification of analytic windows, exclusion rules, preprocessing decisions, and missing-data handling in future applications.

The clustering results may also be sample dependent. Patterns discovered in one dataset can reflect characteristics

of the sample, device, monitoring protocol, preprocessing pipeline, and selected clustering method. The number of clusters is one especially consequential choice. We used prediction strength to guide selection of k , but other approaches, such as the Calinski-Harabasz index or silhouette index, could lead to different solutions [34,35]. In addition, the uncertainty associated with deriving clusters was not fully propagated into the downstream regression models. Resampling or bootstrap-based approaches may help evaluate the stability of discovered patterns and the uncertainty of their associations with outcomes. Generalizability is also bounded by the study population and design: participants were Latino children with overweight or obesity from low-income families, monitored over a 1-week period in a baseline cross-sectional setting. External validation is needed to determine how well these activity-pattern summaries transfer to other populations, devices, monitoring durations, and settings.

Future work can extend this framework in several directions. Within GOALS and similar studies, longitudinal and intervention analyses could evaluate whether baseline cluster-derived summaries predict change in cardiometabolic outcomes or whether changes in activity-pattern membership correspond to intervention response. External validation studies can assess whether similar temporal patterns emerge in other cohorts and with other accelerometer-processing pipelines. A systematic methodological comparison of fixed-alignment clustering with dynamic time warping or Gaussian-process-based similarity measures followed by hierarchical clustering could further characterize the tradeoff between preserving clock-time interpretability and allowing elastic shape-based alignment. Individual-level clustering may also be useful when the research question concerns deviations from a person's usual activity pattern rather than differences across participants. Finally, the same general profile-based logic could be evaluated using other accelerometer scales, such as monitor-independent movement summary measures or raw acceleration signals, and possibly other intensively sampled behavioral or sensor-derived time series.

In conclusion, cluster-derived daily activity measures provide a regression-ready way to summarize accelerometer profiles while retaining information about both activity magnitude and within-day timing. In this baseline methodological analysis, these measures performed comparably to traditional accelerometer summaries while capturing a different aspect of physical activity behavior: the shape and timing of daily movement patterns. The approach is therefore best viewed as a complement to existing accelerometer metrics rather than as a demonstrably superior predictor of the cardiometabolic outcomes examined here. With transparent reporting, external validation across diverse populations and settings, and further evaluation in longitudinal and intervention analyses, cluster-derived profile measures may help expand how accelerometer data are used to study relationships between daily activity patterns and health outcomes.

Acknowledgments

We thank the children and families who participated in the Stanford GOALS trial.

During revision, the authors used ChatGPT to assist with interpreting reviewer comments, organizing revision tasks, and drafting or refining language for author review. All revised manuscript text was reviewed, edited, and verified by the authors, who take full responsibility for the final content. Generative AI was not used to conduct analyses, generate results, or create or alter study data.

Funding

Research reported in this publication was supported in part by grants from the Stanford Maternal and Child Health Research Institute, the Department of Pediatrics at Stanford University, and the Li Ka Shing Foundation Stanford–Oxford Big Data for Human Health seed grant program. Data collection for the parent Stanford GOALS trial was supported by the National Heart, Lung, and Blood Institute of the National Institutes of Health (NIH) under award number U01HL103629. This manuscript is also partially supported by the following NIH grants: R01LM013355, Novel machine learning and missing data methods for improving estimates of physical activity, sedentary behavior, and sleep using accelerometer data; UL1TR003142, Stanford's Center for Clinical and Translational Education and Research under the Biostatistics, Epidemiology, and Research Design Program; P30DK116074, the Clinical and Translational Core of the Stanford Diabetes Research Center; and P30CA124435, the Biostatistics Shared Resource of the National Cancer Institute–sponsored Stanford Cancer Institute. The content is solely the responsibility of the authors and does not necessarily represent the official views of Stanford University or the Li Ka Shing Foundation; the National Heart, Lung, and Blood Institute; the NIH; the US Department of Health and Human Services; or the US government.

Data Availability

The individual-level Stanford GOALS data analyzed in this study are not publicly available because of participant confidentiality and data-use restrictions associated with the parent trial. The Padaco source code used for accelerometer visualization and clustering is publicly available at GitHub [29].

Authors' Contributions

Conceptualization: HM, TNR, KFH, KIK, MD

Data curation: HM, KFH, KIK

Funding acquisition: TNR, MD

Methodology: HM, TNR, KFH, KIK, MD

Supervision: TNR, MD

Writing - original draft: HM, TNR, MD

Writing - review and editing: HM, TNR, AJ, FG, KFH, KIK, MD

Conflicts of Interest

None declared.

Multimedia Appendix 1

Supplementary figures and tables describing accelerometer preprocessing, analytic-window selection, activity summary measures, and clustering diagnostics.

[DOCX File (Microsoft Word File), 1385 KB-Multimedia Appendix 1]

References

1. Troiano RP, McClain JJ, Brychta RJ, Chen KY. Evolution of accelerometer methods for physical activity research. *Br J Sports Med*. Jul 2014;48(13):1019-1023. [doi: [10.1136/bjsports-2014-093546](https://doi.org/10.1136/bjsports-2014-093546)] [Medline: [24782483](https://pubmed.ncbi.nlm.nih.gov/24782483/)]
2. Neishabouri A, Nguyen J, Samuelsson J, et al. Quantification of acceleration as activity counts in ActiGraph wearable. *Sci Rep*. Jul 13, 2022;12(1):11958. [doi: [10.1038/s41598-022-16003-x](https://doi.org/10.1038/s41598-022-16003-x)] [Medline: [35831446](https://pubmed.ncbi.nlm.nih.gov/35831446/)]
3. Arigo D, Mogle JA, Brown MM, et al. Differences between accelerometer cut point methods among midlife women with cardiovascular risk markers. *Menopause*. May 2020;27(5):559-567. [doi: [10.1097/GME.0000000000001498](https://doi.org/10.1097/GME.0000000000001498)] [Medline: [32049926](https://pubmed.ncbi.nlm.nih.gov/32049926/)]
4. Banda JA, Haydel KF, Davila T, et al. Effects of varying epoch lengths, wear time algorithms, and activity cut-points on estimates of child sedentary behavior and physical activity from accelerometer data. *PLoS ONE*. 2016;11(3):e0150534. [doi: [10.1371/journal.pone.0150534](https://doi.org/10.1371/journal.pone.0150534)] [Medline: [26938240](https://pubmed.ncbi.nlm.nih.gov/26938240/)]
5. Brailey G, Metcalf B, Lear R, Price L, Cumming S, Stiles V. A comparison of the associations between bone health and three different intensities of accelerometer-derived habitual physical activity in children and adolescents: a systematic review. *Osteoporos Int*. Jun 2022;33(6):1191-1222. [doi: [10.1007/s00198-021-06218-5](https://doi.org/10.1007/s00198-021-06218-5)] [Medline: [35089364](https://pubmed.ncbi.nlm.nih.gov/35089364/)]

6. Bruijns BA, Truelove S, Johnson AM, Gilliland J, Tucker P. Infants' and toddlers' physical activity and sedentary time as measured by accelerometry: a systematic review and meta-analysis. *Int J Behav Nutr Phys Act*. Feb 7, 2020;17(1):14. [doi: [10.1186/s12966-020-0912-4](https://doi.org/10.1186/s12966-020-0912-4)] [Medline: [32028975](https://pubmed.ncbi.nlm.nih.gov/32028975/)]
7. Leroux A, Di J, Smirnova E, et al. Organizing and analyzing the activity data in NHANES. *Stat Biosci*. Jul 2019;11(2):262-287. [doi: [10.1007/s12561-018-09229-9](https://doi.org/10.1007/s12561-018-09229-9)] [Medline: [32047572](https://pubmed.ncbi.nlm.nih.gov/32047572/)]
8. Romanzini M, Petroski EL, Ohara D, Dourado ACR, Reichert FF. Calibration of ActiGraph GT3X, Actical and RT3 accelerometers in adolescents. *Eur J Sport Sci*. 2014;14(1):91-99. [doi: [10.1080/17461391.2012.732614](https://doi.org/10.1080/17461391.2012.732614)] [Medline: [24533499](https://pubmed.ncbi.nlm.nih.gov/24533499/)]
9. Rowlands AV, Rennie K, Kozarski R, et al. Children's physical activity assessed with wrist-and hip-worn accelerometers. *Med Sci Sports Exerc*. Dec 2014;46(12):2308-2316. [doi: [10.1249/MSS.0000000000000365](https://doi.org/10.1249/MSS.0000000000000365)] [Medline: [24781890](https://pubmed.ncbi.nlm.nih.gov/24781890/)]
10. Santos-Lozano A, Marín PJ, Torres-Luque G, Ruiz JR, Lucía A, Garatachea N. Technical variability of the GT3X accelerometer. *Med Eng Phys*. Jul 2012;34(6):787-790. [doi: [10.1016/j.medengphy.2012.02.005](https://doi.org/10.1016/j.medengphy.2012.02.005)] [Medline: [22417978](https://pubmed.ncbi.nlm.nih.gov/22417978/)]
11. Zhang S, Rowlands AV, Murray P, Hurst TL. Physical activity classification using the GENE wrist-worn accelerometer. *Med Sci Sports Exerc*. Apr 2012;44(4):742-748. [doi: [10.1249/MSS.0b013e31823bf95c](https://doi.org/10.1249/MSS.0b013e31823bf95c)] [Medline: [21988935](https://pubmed.ncbi.nlm.nih.gov/21988935/)]
12. Bammann K, Thomson NK, Albrecht BM, Buchan DS, Easton C. Generation and validation of ActiGraph GT3X+ accelerometer cut-points for assessing physical activity intensity in older adults. The OUTDOOR ACTIVE validation study. *PLOS ONE*. 2021;16(6):e0252615. [doi: [10.1371/journal.pone.0252615](https://doi.org/10.1371/journal.pone.0252615)] [Medline: [34081715](https://pubmed.ncbi.nlm.nih.gov/34081715/)]
13. Colley RC, Tremblay MS. Moderate and vigorous physical activity intensity cut-points for the Actical accelerometer. *J Sports Sci*. May 2011;29(8):783-789. [doi: [10.1080/02640414.2011.557744](https://doi.org/10.1080/02640414.2011.557744)] [Medline: [21424979](https://pubmed.ncbi.nlm.nih.gov/21424979/)]
14. Evenson KR, Herring AH, Wen F. Accelerometry-assessed latent class patterns of physical activity and sedentary behavior with mortality. *Am J Prev Med*. Feb 2017;52(2):135-143. [doi: [10.1016/j.amepre.2016.10.033](https://doi.org/10.1016/j.amepre.2016.10.033)] [Medline: [28109457](https://pubmed.ncbi.nlm.nih.gov/28109457/)]
15. Fraysse F, Post D, Eston R, Kasai D, Rowlands AV, Parfitt G. Physical activity intensity cut-points for wrist-worn GENEActiv in older adults. *Front Sports Act Living*. 2021;2:579278. [doi: [10.3389/fspor.2020.579278](https://doi.org/10.3389/fspor.2020.579278)] [Medline: [33521631](https://pubmed.ncbi.nlm.nih.gov/33521631/)]
16. Kim Y, Beets MW, Welk GJ. Everything you wanted to know about selecting the “right” Actigraph accelerometer cut-points for youth, but...: a systematic review. *J Sci Med Sport*. Jul 2012;15(4):311-321. [doi: [10.1016/j.jsams.2011.12.001](https://doi.org/10.1016/j.jsams.2011.12.001)] [Medline: [22306372](https://pubmed.ncbi.nlm.nih.gov/22306372/)]
17. Robinson TN, Matheson D, Desai M, et al. Family, community and clinic collaboration to treat overweight and obese children: Stanford GOALS-A randomized controlled trial of a three-year, multi-component, multi-level, multi-setting intervention. *Contemp Clin Trials*. Nov 2013;36(2):421-435. [doi: [10.1016/j.cct.2013.09.001](https://doi.org/10.1016/j.cct.2013.09.001)] [Medline: [24028942](https://pubmed.ncbi.nlm.nih.gov/24028942/)]
18. Robinson TN, Matheson D, Wilson DM, et al. A community-based, multi-level, multi-setting, multi-component intervention to reduce weight gain among low socioeconomic status Latinx children with overweight or obesity: the Stanford GOALS randomised controlled trial. *Lancet Diabetes Endocrinol*. Jun 2021;9(6):336-349. [doi: [10.1016/S2213-8587\(21\)00084-X](https://doi.org/10.1016/S2213-8587(21)00084-X)] [Medline: [33933181](https://pubmed.ncbi.nlm.nih.gov/33933181/)]
19. Bai J, He B, Shou H, Zipunnikov V, Glass TA, Crainiceanu CM. Normalization and extraction of interpretable metrics from raw accelerometry data. *Biostatistics*. Jan 2014;15(1):102-116. [doi: [10.1093/biostatistics/kxt029](https://doi.org/10.1093/biostatistics/kxt029)] [Medline: [23999141](https://pubmed.ncbi.nlm.nih.gov/23999141/)]
20. Jones PJ, Catt M, Davies MJ, et al. Feature selection for unsupervised machine learning of accelerometer data physical activity clusters—a systematic review. *Gait Posture*. Oct 2021;90:120-128. [doi: [10.1016/j.gaitpost.2021.08.007](https://doi.org/10.1016/j.gaitpost.2021.08.007)] [Medline: [34438293](https://pubmed.ncbi.nlm.nih.gov/34438293/)]
21. Nawrin SS, Inada H, Momma H, Nagatomi R. Examining physical activity clustering using machine learning revealed a diversity of 24-hour step-counting patterns. *J Act Sedentary Sleep Behav*. Aug 12, 2024;3(1):19. [doi: [10.1186/s44167-024-00059-3](https://doi.org/10.1186/s44167-024-00059-3)] [Medline: [40217371](https://pubmed.ncbi.nlm.nih.gov/40217371/)]
22. Aqeel M, Guo J, Lin L, et al. Temporal physical activity patterns are associated with obesity in U.S. adults. *Prev Med*. Jul 2021;148:106538. [doi: [10.1016/j.ypmed.2021.106538](https://doi.org/10.1016/j.ypmed.2021.106538)] [Medline: [33798532](https://pubmed.ncbi.nlm.nih.gov/33798532/)]
23. Nawrin SS, Inada H, Momma H, Nagatomi R. Twenty-four-hour physical activity patterns associated with depressive symptoms: a cross-sectional study using big data-machine learning approach. *BMC Public Health*. May 7, 2024;24(1):1254. [doi: [10.1186/s12889-024-18759-5](https://doi.org/10.1186/s12889-024-18759-5)] [Medline: [38714982](https://pubmed.ncbi.nlm.nih.gov/38714982/)]
24. Niemelä M, Kangas M, Farrahi V, et al. Intensity and temporal patterns of physical activity and cardiovascular disease risk in midlife. *Prev Med*. Jul 2019;124:33-41. [doi: [10.1016/j.ypmed.2019.04.023](https://doi.org/10.1016/j.ypmed.2019.04.023)] [Medline: [31051183](https://pubmed.ncbi.nlm.nih.gov/31051183/)]
25. Smagula SF, Boudreau RM, Stone K, et al. Latent activity rhythm disturbance sub-groups and longitudinal change in depression symptoms among older men. *Chronobiol Int*. 2015;32(10):1427-1437. [doi: [10.3109/07420528.2015.1102925](https://doi.org/10.3109/07420528.2015.1102925)] [Medline: [26594893](https://pubmed.ncbi.nlm.nih.gov/26594893/)]

26. Smagula SF, Krafty RT, Thayer JF, Buysse DJ, Hall MH. Rest-activity rhythm profiles associated with manic-hypomanic and depressive symptoms. *J Psychiatr Res*. Jul 2018;102:238-244. [doi: [10.1016/j.jpsychires.2018.04.015](https://doi.org/10.1016/j.jpsychires.2018.04.015)] [Medline: [29705489](https://pubmed.ncbi.nlm.nih.gov/29705489/)]
27. Smagula SF, Zhang G, Gujral S, et al. Association of 24-hour activity pattern phenotypes with depression symptoms and cognitive performance in aging. *JAMA Psychiatry*. Oct 1, 2022;79(10):1023-1031. [doi: [10.1001/jamapsychiatry.2022.2573](https://doi.org/10.1001/jamapsychiatry.2022.2573)] [Medline: [36044201](https://pubmed.ncbi.nlm.nih.gov/36044201/)]
28. Choi L, Liu Z, Matthews CE, Buchowski MS. Validation of accelerometer wear and nonwear time classification algorithm. *Med Sci Sports Exerc*. Feb 2011;43(2):357-364. [doi: [10.1249/MSS.0b013e3181ed61a3](https://doi.org/10.1249/MSS.0b013e3181ed61a3)] [Medline: [20581716](https://pubmed.ncbi.nlm.nih.gov/20581716/)]
29. Padaco. GitHub. URL: <https://github.com/informaton/padaco> [Accessed 2026-09-18]
30. Tibshirani R, Walther G. Cluster validation by prediction strength. *J Comput Graph Stat*. Sep 1, 2005;14(3):511-528. [doi: [10.1198/106186005X59243](https://doi.org/10.1198/106186005X59243)]
31. Yu J, Kapphahn K, Moore H, Haydel F, Robinson T, Desai M. Prediction strength for clustering activity patterns using accelerometer data. *J Meas Phys Behav*. 2023;6(2):134-144. [doi: [10.1123/jmpb.2022-0049](https://doi.org/10.1123/jmpb.2022-0049)]
32. Expert Panel on Integrated Guidelines for Cardiovascular Health and Risk Reduction in Children and Adolescents, National Heart, Lung, and Blood Institute. Expert panel on integrated guidelines for cardiovascular health and risk reduction in children and adolescents: summary report. *Pediatrics*. Dec 1, 2011;128(Supplement_5):S213-S256. [doi: [10.1542/peds.2009-2107C](https://doi.org/10.1542/peds.2009-2107C)]
33. Kwac J, Flora J, Rajagopal R. Household energy consumption segmentation using hourly data. *IEEE Trans Smart Grid*. 2014;5(1):420-430. [doi: [10.1109/TSG.2013.2278477](https://doi.org/10.1109/TSG.2013.2278477)]
34. Caliński T, Harabasz J. A dendrite method for cluster analysis. *Commun Stat*. Jan 1974;3(1):1-27. [doi: [10.1080/03610927408827101](https://doi.org/10.1080/03610927408827101)]
35. Rousseeuw PJ. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. *J Comput Appl Math*. Nov 1987;20:53-65. [doi: [10.1016/0377-0427\(87\)90125-7](https://doi.org/10.1016/0377-0427(87)90125-7)]

Abbreviations

AIC: Akaike information criterion
AIM: Activity Intensity Mean
AIV: Activity Intensity Variability
CPM: counts per minute
LPA: light physical activity
MPA: moderate physical activity
MVPA: moderate-to-vigorous physical activity
PCA: principal component analysis
SB: sedentary behavior
TAM: Time Active Mean
TAV: Time Active Variability
VM: vector magnitude
VPA: vigorous physical activity

Edited by Ivan Steenstra; peer-reviewed by Ahmed Torad, Ruben Buendia Lopez; submitted 10.Nov.2025; final revised version received 28.Aug.2026; accepted 31.Aug.2026; published 08.Oct.2026

Please cite as:

Moore IV H, Robinson TN, Jensen A, Gunturkun F, Haydel KF, Kapphahn KI, Desai M
 Clustering-Based Accelerometer Measures of Physical Activity Patterns in Children With Overweight or Obesity: Baseline Cross-Sectional Methodological Analysis
JMIR Form Res 2026;10:e87532
 URL: <https://formative.jmir.org/2026/1/e87532>
 doi: [10.2196/87532](https://doi.org/10.2196/87532)

© Hyatt Moore IV, Thomas N Robinson, Alexandria Jensen, Fatma Gunturkun, K Farish Haydel, Kristopher I Kapphahn, Manisha Desai. Originally published in *JMIR Formative Research* (<https://formative.jmir.org>), 08.Oct.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original

work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on <https://formative.jmir.org>, as well as this copyright and license information must be included.