



Published in final edited form as:

*J Adolesc Health*. 2018 May ; 62(5): 583–590. doi:10.1016/j.jadohealth.2017.11.298.

## Clustering of Health Behaviors and Cardiorespiratory Fitness Among U.S. Adolescents

Jacob Hartz, MD, MPH<sup>1</sup>, Leah Yingling<sup>2</sup>, Colby Ayers, MS<sup>3</sup>, Joel Adu-Brimpong<sup>2</sup>, Joshua Rivers, MS<sup>2</sup>, Chaarushi Ahuja<sup>2</sup>, and Tiffany Powell-Wiley, MD, MPH<sup>4</sup>

<sup>1</sup>Children's National Medical Center

<sup>2</sup>Office of Intramural Training and Education of the National Institutes of Health

<sup>3</sup>Department of Clinical Science, University of Texas – Southwestern

<sup>4</sup>National Heart, Lung, and Blood Institute

### Abstract

**Purpose**—Decreased cardiorespiratory fitness (CRF) is associated with an increased risk of cardiovascular disease. However, little is known how the interaction of diet, physical activity (PA), and sedentary time (ST) affects CRF among adolescents. By using a nationally-representative sample of U.S. adolescents, we used cluster analysis to investigate the interactions of these behaviors on CRF. We hypothesized that distinct clustering patterns exist and that less healthy clusters are associated with lower CRF.

**Methods**—We used 2003–04 National Health and Nutrition Examination Survey data for persons aged 12–19 years (N=1225). PA and ST were measured objectively by accelerometer, and the American Heart Association Healthy Diet Score quantified diet quality.  $\dot{V}O_2$  max was measured by sub-maximal treadmill exercise test. We performed cluster analysis to identify sex-specific clustering of diet, PA, and ST. Adjusting for accelerometer wear-time, age, body mass index (BMI), race/ethnicity, and the poverty-to-income ratio, we performed sex-stratified linear regression analysis to evaluate the association of cluster with  $\dot{V}O_2$  max.

**Results**—Three clusters were identified for girls and boys. For girls, there was no difference across clusters for age ( $p=0.1$ ), weight ( $p=0.3$ ), and BMI ( $p=0.5$ ) and no relationship between clusters and  $\dot{V}O_2$  max. For boys, the youngest cluster ( $p<0.01$ ) had three healthy behaviors, weighed less, and was associated with a higher  $\dot{V}O_2$  max compared with the two older clusters.

---

**Address correspondence to:** Tiffany Powell-Wiley, MD, MPH, Cardiovascular and Pulmonary Branch, Division of Intramural Research, National Heart, Lung, and Blood Institute, National Institutes of Health, Building 10, 5-3330, 10 Center Drive, Bethesda, MD, 20892, United States, Phone: 1-301-594-3735, Fax: 1 301 402 0888, tiffany.powell-wiley@nih.gov.

**Publisher's Disclaimer:** This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

**Financial Disclosure:** The authors have no financial relationships to disclose.

**Conflict of Interest:** The authors have no conflicts of interest to disclose.

**Conclusions**—We observed clustering of diet, PA, and ST in U.S. adolescents. Specific patterns were associated with lower  $\dot{V}O_2$  max for boys, suggesting that our clusters may help identify adolescent boys most in need of interventions.

### Keywords

Cluster analysis; nutritional status; cardiorespiratory fitness; sedentary lifestyles

---

## INTRODUCTION

Approximately one in five adolescents in the United States is obese (body mass index 95th percentile, for age and sex).(1) Obesity often occurs simultaneously with low cardiorespiratory fitness (CRF). A low CRF in adolescence is associated an abnormal lipid profile, increased risk of metabolic syndrome, and increased arterial stiffness.(2, 3) While a number of non-modifiable factors influence CRF, it can be improved with a healthful diet, more physical activity (PA), and less sedentary time (ST).(4) Unfortunately, poor diet quality,(5) inadequate PA,(6) and excessive ST(7) are well-documented among U.S. adolescents.

Although each of these health behaviors is an independent risk factor for low CRF(3, 4) and cardiovascular disease,(8) co-occurrence of more than one of these risk factors may increase cardiovascular risk (as measured by CRF) than any of the factors alone.(9–12) Further, the effects of one health behavior may be influenced by another.(10, 11) One method to better understand the interplay of diet, PA, and ST is cluster analysis. Cluster analysis does not predefine groups, but rather identifies groups based on shared characteristics from the available data.(13) Understanding how these health behaviors cluster can help identify populations at greatest risk for cardiovascular disease and who would benefit most from targeted interventions.

Previous studies evaluating the clustering of different health behaviors in adolescents have been limited by a small sample size,(9, 12) reliance on self-report,(11, 14) or focused on populations outside the U.S.(11, 14) Objective measures of PA and ST are important because they more predictive of cardiovascular disease risk than self-reported data.(15, 16)

In this study, we use the 2003–2004 National Health and Nutrition Examination Survey (NHANES) as it includes a diverse population of 12–19 year olds that are representative of adolescents in the U.S. Further, it provides objective measurements of PA and ST using an accelerometer and an objective measure of cardiorespiratory fitness using a submaximal stress test, which is a validated predictor of maximal oxygen consumption ( $\dot{V}O_2$  max)<sup>1</sup> in adolescents.(17, 18) We hypothesized that cluster analysis would allow U.S. adolescents to be grouped in a meaningful and reliable manner that would help identify sub-populations with elevated cardiovascular risk. We also hypothesized that clusters based on diet quality, PA, and ST would be associated with CRF.

---

<sup>1</sup>  $\dot{V}O_2$  max (mL/kg/min) is a continuous variable that represents the maximal oxygen uptake during intense exercise and was estimated from a submaximal exercise stress test.

## METHODS

We used data from the 2003–2004 NHANES cohort.(19) NHANES is an ongoing series of studies conducted by the National Center for Health Statistics that provides a representative sample of the U.S. civilian, non-institutionalized population. The selection of participants is through a complex, multistage probability design. The specifics of NHANES have been described.(20) The study was conducted according to the guidelines in the Declaration of Helsinki and all procedures involving human subjects were approved by the National Center for Health Statistics Institutional Review Board. Written informed consent was obtained from all participants 18 years of age and older. The NHANES protocol was developed and reviewed to be in compliance with the Health and Human Services Policy for Protection of Human Research Subjects (45 CFR part 46).

### Dietary Intake and Quality Data

Dietary quality was obtained using NHANES 24-hour dietary recall data. NHANES participants underwent two 24-hour dietary recalls; the first consisted of an in-person interview on Day 1 at the mobile examination center (MEC), while the second occurred via a telephone interview conducted 3–10 days following Day 1. We only include those who completed the Day 1 24-hour recall as a majority of participants completed Day 1 testing at the MEC examination and the setting was consistent for all participants.

Each reported food or beverage is recorded based on a specific food code from the US Department of Agriculture Food and Nutrient Database for Dietary Studies (FNDDS). For the 2003–2004 survey, FNDDS 2.0 was used.(21)

A healthy eating score was then calculated to assess diet quality, as previously described by Lloyd-Jones et al. and consistent with the current Dietary Guidelines for Americans and American Heart Association recommendations.(22, 23) The healthy diet score sets ideal criteria for five constituents of dietary intake: fruits and vegetables ( 4.5 cups per day), fish ( two 3.5-oz servings per week), fiber-rich whole grains ( three 1-oz equivalent servings per day), sodium (<1500 milligrams per day), and sugar-sweetened beverages ( 450 kcal (36 oz.) per week). The highest possible diet score is 5 (meeting all five components of the healthy diet criteria) and lower scores represent lower diet quality. This score was then used in the subsequent cluster analysis. The healthy eating score was then converted to a standardized z-score.

### Measurement of Accelerometer-based Activity Level And Wear-Time

NHANES adolescents wore an Actigraph model 7164 accelerometer (Actigraph, LLC; Ft. Walton Beach, FL) over the right hip on an elastic belt for 7 days while awake. The accelerometer was to be removed only for swimming or bathing. The uniaxial Actigraph accelerometer measures and records vertical acceleration as “counts,” which reflect the intensity of PA associated with locomotion. Data were recorded in 1-min epochs for up to 7 days. Details of the accelerometer protocol are available.(24)

Consistent with previous research, only those with at least four valid days of wear time were included in our analysis.(20) A valid day was defined as 10 hours of “wear time.” Wear

time was determined by subtracting non-wear time from 24 hours. Non-wear time was defined as any interval of zero intensity counts that lasted for at least 60 consecutive minutes with allowance for 1–3 consecutive minutes of counts between 0–100. Wear time was categorized and three of the intensity levels were used for analysis: 1) ST, 2) moderate-intensity PA, and 3) vigorous-intensity PA. A composite variable, moderate-to-vigorous physical activity (MVPA), was defined as activity greater than or equal to the threshold for moderate activity. Intensity thresholds have previously been determined for those 18 years, with moderate intensity PA having 2020 counts/minute and vigorous activity having 5999. For adolescents aged 12–17 years, different thresholds were used for each age group to adjust for differences in resting metabolic rates and are provided in Supplementary Table 1.(25) A bout of exercise was defined as having >8–10 minutes above the pre-specified count thresholds and then quantified in mean minutes of activity per day. ST was defined as <100 counts per minute of wear time and quantified in mean minutes per day.(26) The SAS syntax for this procedure was provided by the National Cancer Institute.(27) The amount of time spent in MVPA and ST was then converted to a standardized z-score.

### Cardiorespiratory Fitness

We used CRF as the outcome variable. CRF was estimated by a submaximal treadmill exercise test and quantified using  $\dot{V}O_2$  max (mL/kg/min), which is a participant's maximum rate of oxygen consumption during a stepwise exercise test. A participant's predicted  $\dot{V}O_2$  max determined which of the eight possible treadmill tests was used to measure  $\dot{V}O_2$  max. The predicted  $\dot{V}O_2$  max was based on the formula developed by Jackson, which incorporates the individual's sex, age, body mass index (BMI), and self-reported level of PA. (24, 28) Each protocol included a 2-minute warm-up, two 3-minute exercise stages, and a 2-minute cool-down period. The goal of each protocol was to elicit a heart rate that was 75% of the age-predicted maximum ( $220 - \text{age in years}$ ) by the end of the second stage. The heart rate was monitored continuously via four electrodes connected to the trunk and abdomen and was recorded at the end of warm-up, the end of each exercise stage, and during each minute of recovery. Blood pressure was measured using an STBP-780 automated sphygmomanometer (Colin Medical Instruments Corporation, San Antonio, TX).  $\dot{V}O_2$  max was estimated by extrapolation to an expected, age-specific maximal heart rate by using measured heart rate responses to the two 3-minute exercise stages. This assumes a linear relationship between heart rate and oxygen consumption during exercise. The procedures used in developing the standards for  $\dot{V}O_2$  max have been described in detail.(29)

### Other Covariates

Age, sex, and race/ethnicity were defined using self-reported demographic data from 2003–2004 NHANES. We included non-Hispanic Whites, non-Hispanic Blacks, and Mexican Americans who reported a single racial identity because the group sample sizes were adequate for separate evaluation and estimates. The poverty-to-income ratio (PIR) compares family income to the Health and Human Services federal poverty threshold.(30) A PIR <1 indicates that the family income is <130% of the poverty threshold, a PIR of 1–3 corresponds to 131%–185% of the poverty threshold, and a PIR >3 corresponds to income

>185% of the poverty threshold. Weight and height were measured during the NHANES physical examination and used to calculate BMI (weight in kilograms divided by the square of the height in meters). A BMI between 85<sup>th</sup>–<95<sup>th</sup> percentile was considered overweight and a BMI ≥95<sup>th</sup> percentile for age and sex was considered obese. Computer time and television time were obtained via self-report during the interview session. A physician obtained blood pressures using a manual sphygmomanometer with an appropriately sized arm cuff after participants rested 5 minutes in a seated position. Hypertension was classified as a systolic blood pressure ≥95<sup>th</sup> percentile for sex, age, and height.(31, 32) Triglycerides and total cholesterol were measured from venous samples using enzymatic reactions. High density lipoprotein cholesterol (HDL-cholesterol) and low-density lipoprotein cholesterol (LDL-cholesterol) were measured from venous samples using a direct immunoassay method. The glucose concentration was determined by spectrophotometrically measuring the nicotinamide adenine dinucleotide concentration, which is directly proportional to the glucose concentration. Insulin was measured using the two-site immunoenzymometric assay. Using insulin and fasting plasma glucose, the homeostasis model assessment of insulin resistance (HOMA-IR) was used to estimate insulin resistance with the formula: fasting serum insulin (μU/ml) × fasting plasma glucose (mg/dL)/405.(33) Detailed specimen collection and processing instructions are described in the NHANES Laboratory Procedures Manual.(34)

### Study Population

The 2003–2004 cohort was chosen because those aged 12–19 years were eligible to undergo a submaximal stress test and to wear an accelerometer with a non-missing sample weight. Of 2303 eligible participants, 1600 completed a valid stress test. Of those, 141 were excluded from using an accelerometer because of impairments in walking or wearing the device as determined by NHANES. Of the remaining 1459 participants, 83 were excluded because the device was not in calibration and 119 did not have four days of valid wear time. An additional four participants had invalid intensity counts. Of those with valid accelerometer data and a valid stress test, 20 participants were excluded because of a missing or invalid diet survey. Of the original 12–19-year-olds involved in the 2003–2004 NHANES cohort, 1233 met our inclusion criteria.

### Statistical Analysis

To account for the complex survey design of NHANES, we used sample weighting with SURVEY procedures in the statistical software package SAS version 9.4 (SAS Institute Inc., Cary, NC). Because of the differences in units, diet quality, ST, and MVPA were transformed to standardized z-scores. Scores above the mean were labeled “healthful” for diet and PA, whereas scores below the mean were labeled as “healthful” for ST. Patterns among these behaviors were then determined using latent class analysis. The patient clustering was an agglomerative hierarchical clustering approach where each patient formed his/her own singleton cluster and clusters were subsequently merged together. Utilizing Ward’s minimum variance method, the distance between two clusters was defined via Euclidean distances, and this distance was calculated between all possible combinations of two clusters. Clusters with the smallest distance were merged with new scores calculated. At each stage, the semi-partial  $R^2$  statistic was calculated to measure the loss of homogeneity

within clusters caused by the merging process. Clusters were derived for the overall population and sex-specific clusters were identified, with only data from sex-specific clusters shown. To ensure we created meaningful clusters, we tested the reliability of the clustered solution in a random subsample (50% of the sample) of the data and compared the degree of agreement between the clusters and the subsample using weighted kappa statistic. The 3-cluster solution gave a very good degree of agreement ( $\kappa=0.91$ ), while the 4-cluster solution gave a slightly less degree of agreement ( $\kappa=0.80$ ). The sample-weighted  $\chi^2$  test was used to compare baseline characteristics of the NHANES population across the different clusters for boys and girls. A sample-weighted F test was used to compare energy intake across the different clusters. Finally, we used sample-weighted, sex-stratified multivariate linear regression models to determine the association between the clusters and  $\dot{V}O_2$  max, adjusting for accelerometer wear-time, BMI, age, race, and PIR as a categorical variable. Two-sided p values  $<0.05$  were considered statistically significant.

## RESULTS

The study population (N=1233) was composed of 48.5% girls and 12.6% Mexican American, 16.3% non-Hispanic blacks, and 71.2% non-Hispanic whites. The mean age of the population was 15.4 years ( $\pm 0.7$  years).

We identified three clusters for girls and three clusters for boys (Figure 1). For girls, Cluster 1A had high diet scores, high levels of MVPA, and high levels of ST. Cluster 2A had high diet scores, low MVPA, and low ST. Cluster 3A had low diet scores, low MVPA, and a high ST.

For boys, Cluster 1B had high diet scores, high MVPA, and low ST. Cluster 2B had low ST, but poor diet quality and low MVPA. Cluster 3B had high diet scores, low MVPA, and high ST.

Most U.S. girls were in Clusters 2A and 3A (with a nearly equal distribution between the two), and significantly fewer girls were in Cluster 1A (Table 1). Cluster 1A had the highest proportion of Mexican Americans; whereas, Cluster 2A had the highest proportion of non-Hispanic whites, and Cluster 3A had the highest proportion of non-Hispanic blacks ( $p<0.01$ ). There was no difference in age distribution among clusters ( $p=0.13$ ).

For boys, cluster 3B was older than those in the remaining clusters ( $p<0.01$ ) (Table 1). Cluster 1B had the highest percentage of Non-Hispanic black adolescents ( $p<0.01$ ), whereas Cluster 2B had the highest percentage of Mexican American ( $p<0.01$ ) and non-Hispanic white ( $p<0.01$ ) adolescents.

Among girls, there was no statistically significant difference in weight or BMI ( $p>0.05$  across clusters (Table 2). There was a statistically significant difference across the three clusters for diet quality, MVPA, and ST ( $p<0.01$ ). Interestingly, although the objective measure of sedentary time was statistically significant, the self-reported sedentary behaviors, television time ( $p=0.88$ ) and computer time ( $p=0.75$ ), were not statistically significant across the clusters. Both systolic blood pressure ( $p=0.04$ ) and diastolic blood pressure ( $p<0.01$ )



were statistically different across the clusters. There was no significant difference in estimated  $\dot{V}O_2$  max across the clusters.

For boys, those in Cluster 3B had the highest weight ( $p<0.01$ ). BMI increased across clusters, but was not statistically significant ( $p=0.08$ ) (Table 3). There was a statistically significant difference across the clusters for diet quality, MVPA, and ST ( $p<0.01$ ). While computer time was statistically different across the clusters ( $p=0.03$ ), television time was not ( $p=0.79$ ). Cluster 1B also had lower diastolic blood pressure ( $p=0.02$ ) and higher HDL-cholesterol ( $p<0.01$ ). Systolic blood pressure increased across clusters, but was not statistically significant ( $p=0.08$ ). Boys in Cluster 1B had the highest  $\dot{V}O_2$  max ( $p<0.01$ ).

In multivariable modeling, there was no association among clusters and  $\dot{V}O_2$  max among girls (Table 4). Among boys, there was a statistically significant relationship among the clusters and  $\dot{V}O_2$  max. For boys, Cluster 3B, which had the highest amount of ST, had the lowest  $\dot{V}O_2$  max, even though it had highest diet scores and similar MVPA to Cluster 2B.

## DISCUSSION

The present study uses objective measures to demonstrate that modifiable risk factors for cardiovascular disease cluster in clinically meaningful ways in U.S. adolescents. Consistent with previous research, this study shows that both healthful and unhealthful behaviors cluster together and that health behaviors cluster differently in boys and girls.(9–11) However, our findings are the first to demonstrate a correlation between clusters of health behaviors in boys with an objective measure of CRF, a strong predictor of cardiovascular and all-cause mortality.(2) Further, our study was the first to use objective measures of PA, ST, and CRF in an ethnically diverse, nationally representative sample of U.S. adolescents. By using objective measures, we overcome the limitations of self-reported data that were used in the prior studies.(9, 11, 12, 14) The clustering patterns seen in our study emphasize the importance of holistic interventions that target multiple health behaviors.(16)

Our findings suggest that there is a need for lifestyle interventions that target younger adolescent boys in order to help them maintain healthful behaviors. For boys, Cluster 1B was the youngest cluster and had healthful behaviors across all three domains. Unfortunately, it was also the smallest cluster, comprising only 10.3% of boys. As in previous studies, we show significant changes in health behaviors occur during adolescence, particularly in regards to MVPA and ST.(4, 6, 9, 20, 35, 36) Unfortunately, this is also a time of reduced contact with health care providers, which makes intervening more challenging. (37)

Additionally, our study demonstrates the need to target girls who do not have any healthful behaviors. Although, more than half of U.S. adolescent girls had at least two healthful behaviors, 44.7% of girls had no healthful behaviors. For girls, screening for dietary quality, PA, and ST in the clinical setting may identify those in greatest need of intervention. Moreover, non-Hispanic black girls were disproportionately in the cluster with no healthful behaviors, highlighting an at-risk population in need of effective, targeted interventions.(38)

Our work also suggests the synergistic effects of MVPA and ST in predicting CRF. Among both girls and boys, there were clusters with disparate levels of MVPA and ST. This is consistent with previous findings that increased MVPA does not necessarily displace ST.(10) Only when both high MVPA and low ST were present (Cluster 1B) were the clusters predictive of CRF. This was true even if a high MVPA was associated with a healthful diet. Although previous research has found that there is increased caloric intake with increased levels of ST, our study suggests for U.S. adolescents, a healthful diet can be associated with both high and low levels of ST.(10, 14, 36)

Finally, the various combinations of health behaviors in our study emphasize the importance of holistic interventions addressing all aspects of a healthful lifestyle including diet, MVPA, and ST.(39) Including tools and strategies to reduce ST in multifaceted interventions will likely increase the effectiveness of interventions in adolescents and not be overly resource intensive.(39)

### Limitations

The study had certain limitations. First, as a cross-sectional study we cannot comment on causation between CRF and the independent variables. Second, a submaximal treadmill test is inferior to a symptom-limited exercise test because it relies on more assumptions. However, the submaximal stress test has a fairly high correlation between estimated and measured  $\dot{V}O_2$  max. (17, 18) Third, we used cluster analysis rather than latent class analysis, despite that cluster analysis is more subjective. However, the 3-cluster solution had the highest level of agreement and is likely easier to use than having more clusters. Lastly, there are limitations of our dataset. The accelerometer data available in this dataset was in 1-minute epochs, rather than 30 second epochs, which are typically used to measure the shorter periods of activity in adolescents. Similarly, although a more comprehensive diet survey would provide more accurate information on a participant's diet, the 24-hour diet recall provides high quality data.(40)

### CONCLUSIONS

In conclusion, we find that modifiable risk factors for cardiovascular disease cluster in a meaningful manner and that healthful characteristics do not necessarily cluster together. For boys, our clusters are associated with CRF, but the clusters are not associated with CRF for girls. In boys, interventions that start before adolescence are needed to prevent the development of unhealthful behaviors. For girls, resources should be focused on the cluster of girls with no healthful behaviors, as this group is at particular risk. Further, holistic interventions that improve the diet, increase PA, and reduce ST are likely needed in order to have appreciable effects on cardiovascular health.

### Supplementary Material

Refer to Web version on PubMed Central for supplementary material.



## Acknowledgments

**Sources of Funding:** Dr. Powell-Wiley, Ms. Yingling, Mr. Adu-Brimpong, Mr. Rivers, and Ms. Ahuja are funded by the Division of Intramural Research of the National Heart, Lung, and Blood Institute.

**Funding Source:** The views expressed in this manuscript are those of the authors and do not necessarily represent the views of the National Heart, Lung, and Blood Institute; the National Institutes of Health; or the U.S. Department of Health and Human Services. Funding for Tiffany Leah Yingling, Joel Adu-Brimpong, Joshua Rivers, Chaarushi Ahuja, and Powell-Wiley is provided through the Division of Intramural Research of the National Heart, Lung, and Blood Institute (NHLBI) of the National Institutes of Health (NIH).

## Biographies

**Jacob Hartz:** Dr. Hartz made substantial contributions to conception and design, acquisition of data, analysis and interpretation of data, drafting the article and revising it critically for important intellectual content. He gave final approval of the version to be published, and agrees to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

**Leah Yingling:** Ms. Yingling made substantial contributions to conception and design, acquisition of data, analysis and interpretation of data, drafting the article and revising it critically for important intellectual content. She gave final approval of the version to be published, and agrees to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

**Colby Ayers:** Mr. Ayers made substantial contributions to conception and design, acquisition of data, analysis and interpretation of data, drafting the article and revising it critically for important intellectual content. He gave final approval of the version to be published, and agrees to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

**Joel Adu-Brimpong:** Mr. Adu-Brimpong made substantial contributions to conception and design, acquisition of data, analysis and interpretation of data, drafting the article and revising it critically for important intellectual content. He gave final approval of the version to be published, and agrees to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

**Joshua Rivers:** Mr. Rivers made substantial contributions to conception and design, acquisition of data, analysis and interpretation of data, drafting the article and revising it critically for important intellectual content. He gave final approval of the version to be published, and agrees to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

**Chaarushi Ahuja:** Ms. Ahuja made substantial contributions to conception and design, acquisition of data, analysis and interpretation of data, drafting the article and revising it critically for important intellectual content. He gave final approval of the version to be

published, and agrees to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Tiffany Powell-Wiley: Dr. Powell-Wiley made substantial contributions to conception and design, acquisition of data, analysis and interpretation of data, drafting the article and revising it critically for important intellectual content. She gave final approval of the version to be published, and agrees to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

## Abbreviations

<b>BMI</b>	Body mass index
<b>CRF</b>	cardiorespiratory fitness
<b>MVPA</b>	moderate-to-vigorous physical activity
<b>NHANES</b>	National Health and Nutrition Examination Survey
<b>PA</b>	physical activity
<b>ST</b>	sedentary time
$\dot{V}O_2$ max	maximal oxygen consumption

## References

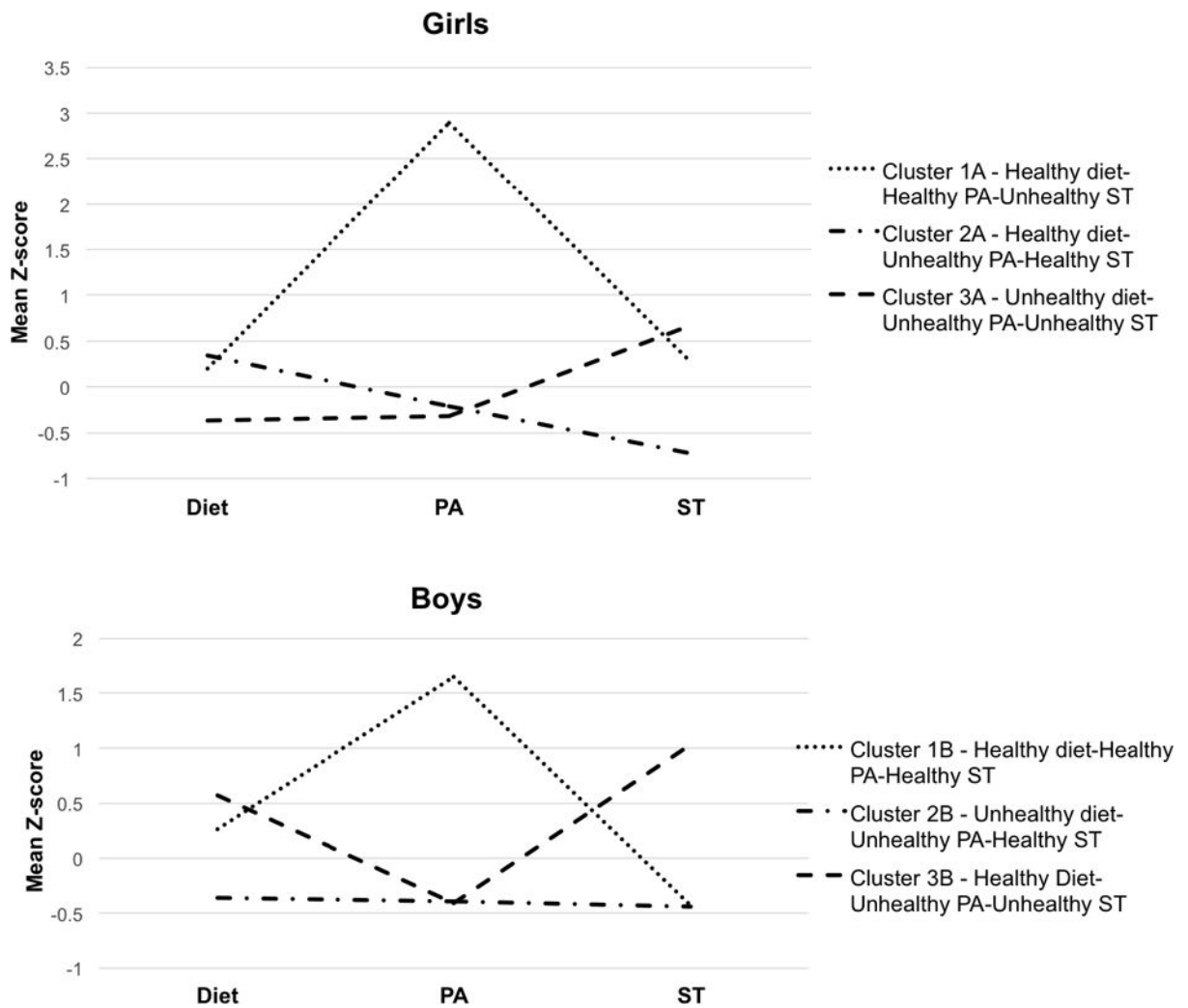
1. Ogden CL, Carroll MD, Curtin LR, et al. Prevalence of high body mass index in US children and adolescents, 2007–2008. *JAMA*. 2010; 303:242–249. [PubMed: 20071470]
2. Kodama S, Saito K, Tanaka S, et al. Cardiorespiratory fitness as a quantitative predictor of all-cause mortality and cardiovascular events in healthy men and women: a meta-analysis. *JAMA*. 2009; 301:2024–2035. [PubMed: 19454641]
3. Ross R, Blair SN, Arena R, et al. Importance of Assessing Cardiorespiratory Fitness in Clinical Practice: A Case for Fitness as a Clinical Vital Sign: A Scientific Statement From the American Heart Association. *Circulation*. 2016; 134:e653–e699. [PubMed: 27881567]
4. Janz KF, Dawson JD, Mahoney LT. Tracking Physical Fitness and Physical Activity from Childhood to Adolescence: The Muscatine Study. *Med Sci Sports Exerc*. 2000; 32:1250–1257. [PubMed: 10912890]
5. Banfield EC, Liu Y, Davis JS, et al. Poor Adherence to US Dietary Guidelines for Children and Adolescents in the National Health and Nutrition Examination Survey Population. *J Acad Nutr Diet*. 2016; 116:21–27. [PubMed: 26391469]
6. Fakhouri TH, Hughes JP, Burt VL, et al. Physical activity in U.S. youth aged 12–15 years, 2012. *NCHS Data Brief*. 2014:1–8.
7. Pate RR, Mitchell JA, Byun W, et al. Sedentary behaviour in youth. *Br J Sports Med*. 2011; 45:906–913. [PubMed: 21836174]
8. Poitras VJ, Gray CE, Borghese MM, et al. Systematic review of the relationships between objectively measured physical activity and health indicators in school-aged children and youth. *Appl Physiol Nutr Metab*. 2016; 41:S197–239. [PubMed: 27306431]

9. Leech RM, McNaughton SA, Timperio A. The clustering of diet, physical activity, and sedentary behavior in children and adolescents: a review. *International Journal of Behavioral Nutrition and Physical Activity*. 2014; 11:4. [PubMed: 24450617]
10. Cuenca-Garcia M, Huybrechts I, Ruiz JR, et al. Clustering of multiple lifestyle behaviors and health-related fitness in European adolescents. *J Nutr Educ Behav*. 2013; 45:549–557. [PubMed: 23702476]
11. Ottevaere, C., Huybrechts, I., Benser, J. Clustering patterns of physical activity, sedentary and dietary behavior among European adolescents: The HELENA study. *BMC Public Health*; 2011. p. 11
12. Boddy LM, Murphy MH, Cunningham C, et al. Physical activity, cardiorespiratory fitness, and clustered cardiometabolic risk in 10- to 12-year-old school children: the REACH Y6 study. *Am J Hum Biol*. 2014; 26:446–451. [PubMed: 24599609]
13. Everitt, BS., Landau, S., Leese, M., et al. *Cluster Analysis*. 5th. West Sussex: John Wiley & Sons, Ltd; 2011.
14. Rey-Lopez JP, Vicente-Rodriguez G, Repasy J, et al. Food and drink intake during television viewing in adolescents: the Healthy Lifestyle in Europe by Nutrition in Adolescence (HELENA) study. *Public Health Nutr*. 2011; 14:1563–1569. [PubMed: 21338558]
15. Taber DR, Stevens J, Murray DM, et al. The effect of a physical activity intervention on bias in self-reported activity. *Ann Epidemiol*. 2009; 19:316–322. [PubMed: 19230711]
16. Celis-Morales CA, Perez-Bravo F, Ibanez L, et al. Objective vs. self-reported physical activity and sedentary time: Effects of measurement method on relationships with risk biomarkers. *PLoS One*. 2012; 7:e36345. [PubMed: 22590532]
17. Ferrar K, Evans H, Smith A, et al. A systematic review and meta-analysis of submaximal exercise-based equations to predict maximal oxygen uptake in young people. *Pediatr Exerc Sci*. 2014; 26:342–357. [PubMed: 24722792]
18. Unnithan VB, Murray LA, Timmons JA, et al. Reproducibility of cardiorespiratory measurements during submaximal and maximal running in children. *Br J Sports Med*. 1995; 29:66–71. [PubMed: 7788223]
19. Centers for Disease Control and Prevention. Statistics NCfH. National Health and Nutrition Examination Survey Data. Hyattsville, MD: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention; 2003–2004.
20. Troiano RP, Berrigan D, Dodd KW, et al. Physical activity in the United States measured by accelerometer. *Med Sci Sports Exerc*. 2008; 40:181–188. [PubMed: 18091006]
21. USDA Food and Nutrient Database for Dietary Studies. Agricultural Research Service, Food Surveys Research Group. Available at: Website:<http://www.ars.usda.gov/News/docs.htm?docid=12068>.
22. Lloyd-Jones DM, Hong Y, Labarthe D, et al. Defining and setting national goals for cardiovascular health promotion and disease reduction: the American Heart Association’s strategic Impact Goal through 2020 and beyond. *Circulation*. 2010; 121:586–613. [PubMed: 20089546]
23. Johnson RK, Appel LJ, Brands M, et al. Dietary sugars intake and cardiovascular health: a scientific statement from the American Heart Association. *Circulation*. 2009; 120:1011–1020. [PubMed: 19704096]
24. Centers for Disease Control and Prevention; Control CfD. National Health and Nutrition Examination Survey: Cardiovascular Fitness Procedures Manual. 2004
25. Trost SG, Pate RR, Sallis JF, et al. Age and gender difference in objectively measured physical activity in youth. *Med Sci Sports Exerc*. 2002
26. Tudor-Locke C, Camhi SM, Troiano RP. A catalog of rules, variables, and definitions applied to accelerometer data in the National Health and Nutrition Examination Survey, 2003–2006. *Prev Chronic Dis*. 2012; 9:E113. [PubMed: 22698174]
27. National Cancer Institute. SAS programs for analyzing NHANES 2003–2004 accelerometer data. Available at: [http://appliedresearch.cancer.gov/tools/nhanes\\_pam/](http://appliedresearch.cancer.gov/tools/nhanes_pam/) Accessed August 20, 2016
28. Jackson AS, Blair SN, Mahar MT, et al. Prediction of functional aerobic capacity without exercise testing. *Med Sci Sports Exerc*. 1990; 22:863–870. [PubMed: 2287267]

29. Welk GJ, Laurson KR, Eisenmann JC, et al. Development of youth aerobic-capacity standards using receiver operating characteristic curves. *Am J Prev Med.* 2011; 41:S111–116. [PubMed: 21961610]
30. U.S. Census Bureau PD, Fertility & Family Statistics Branch. Current Population Survey: Definitions and explanations. 2004
31. National Health and Nutrition Examination Survey - Physician Examination Procedures Manual. 2003
32. Egan BM, Zhao Y, Axon RN. US trends in prevalence, awareness, treatment, and control of hypertension, 1988–2008. *JAMA.* 2010; 303:2043–2050. [PubMed: 20501926]
33. Matthews DR, Hosker JP, Rudenski AS, et al. Homeostasis model assessment: insulin resistance and  $\beta$ -cell function from fasting plasma glucose and insulin concentrations in man. *Diabetologia.* 1985; 28:412–419. [PubMed: 3899825]
34. National Health and Nutrition Examination Survey - Laboratory Procedures Manual. 2004
35. Tremblay MS, LeBlanc AG, Kho ME, et al. Systematic review of sedentary behaviour and health indicators in school-aged children and youth. *Int J Behav Nutr Phys Act.* 2011; 8:98. [PubMed: 21936895]
36. Pearson N, Haycraft E, Johnston JP, et al. Sedentary behaviour across the primary-secondary school transition: A systematic review. *Prev Med.* 2017; 94:40–47. [PubMed: 27856338]
37. Bloom, B., Freeman, G. Tables of Summary Health Statistics for US Children: 2014 National Health Interview Survey 2015. CDC/NCHS, National Health Interview Survey; 2014. p. 2014
38. Fitzgibbon ML, Tussing-Humphreys LM, Porter JS, et al. Weight loss and African-American women: a systematic review of the behavioural weight loss intervention literature. *Obes Rev.* 2012; 13:193–213. [PubMed: 22074195]
39. Altenburg TM, Kist-van Holthe J, Chinapaw MJ. Effectiveness of intervention strategies exclusively targeting reductions in children’s sedentary time: a systematic review of the literature. *Int J Behav Nutr Phys Act.* 2016; 13:65. [PubMed: 27276873]
40. Prentice RL, Mossavar-Rahmani Y, Huang Y, et al. Evaluation and comparison of food records, recalls, and frequencies for energy and protein assessment by using recovery biomarkers. *Am J Epidemiol.* 2011; 174:591–603. [PubMed: 21765003]

### Implications and Contribution

Diet, physical activity, and sedentary time affect cardiorespiratory fitness independently and in concert. This study uses objective measurements of these variables in cluster analysis of a nationally representative cohort of adolescents. Clusters contained both healthy and unhealthy behaviors and are predictive of cardiorespiratory fitness.



**Figure 1.** Standardized z-scores for each of the health behaviors by cluster  
<sup>a</sup>ST, “healthy” is characterized by a negative z-score (z-score below the mean). Sedentary time, moderate activity, vigorous activity, and total moderate-vigorous activity were obtained by accelerometer.



**Table 1**  
Sample-weighted baseline demographics of the 2003–2004 NHANES adolescents, aged 12–19 years

Cluster	Girls <i>n</i> = 576						Boys, <i>n</i> = 653					
	1A	2A	3A	P	1B	2B	3B	P				
<i>n</i>	48	269	263	N/A	99	360	194	N/A				
Weighted % of population	5.7% (1.1)	49.6%	44.7 (2.7)		10.3 (1.7)	62 (3.5)	27.8 (2.6)					
Age (years)	15.3 (0.7)	15 (0.1)	15.6 (0.2)	0.1	14.3 (0.3)	15.3 (0.1)	16.1 (0.2)	<0.01				
Mexican American (%)	28.6 (6.2)	11.6 (2.5)	11.9 (4.2)		20.2 (6.9)	11.3 (3.1)	12.2 (2.8)					
Non-Hispanic, White (%)	58.8 (7.8)	73.8 (4.5)	68.8 (5.5)	<0.01	49.4 (8.8)	76.8 (4.4)	69.0 (5.5)	<0.01				
Non-Hispanic, Black (%)	17.6 (6.5)	14.6 (2.9)	19.3 (3.2)		30.3 (6.3)	11.9 (2.7)	18.8 (3.5)					
Poverty:Income Ratio <sup>a</sup> <1 (%)	23.5 (10.6)	20.1 (2.4)	25.9 (4.6)		23.7 (6.5)	14.4 (2.4)	22.3 (5.4)					
Poverty:Income Ratio 1–3 (%)	38.2 (8.4)	40.5 (4.5)	26.6 (3.6)	0.22	43 (8.2)	36.8 (3)	34.5 (4.7)	0.11				
Poverty:Income Ratio >3 (%)	38.3 (11.6)	39.4 (4.4)	46.5 (5.5)		33 (8.7)	48.8 (2.9)	43.2 (6)					

<sup>a</sup> A Poverty:Income Ratio (PIR) below 1 indicates that the family income is <130% of the poverty threshold, a PIR of 1–3 corresponds to 131–185% of the poverty threshold, and a PIR of >3 corresponds to income >185% of the poverty threshold.

**Table 2**

Cardio-metabolic factors and health behaviors for girls in 2003–2004 NHANES, aged 12–19 years

	Cluster 1A <sup>a</sup>	Cluster 2A	Cluster 3A	p
	N=48	N=318	N=261	
	Mean (SD)	Mean (SD)	Mean (SD)	
Weight (kg)	59.9 (2)	60.1 (1.26)	62.3 (1.86)	0.34
BMI (kg/m <sup>2</sup> )	23 (0.7)	23.1 (0.4)	23.8 (0.7)	0.45
BMI, 85 <sup>th</sup> – <95 <sup>th</sup> percentile <sup>b</sup>	17.8% (8)	18.7% (3.9)	15.7% (2.1)	0.78
BMI, 95 <sup>th</sup> percentile <sup>c</sup>	9% (6.3)	14.7% (3.1)	15.1% (4.3)	0.81
Waist-to-Height Ratio	0.49 (0.01)	0.5 (0.01)	0.5 (0.01)	0.76
Blood pressure, systolic (mm Hg)	104.6 (1.6)	104.5 (0.8)	106.5 (1)	0.04
Blood pressure, diastolic (mm Hg)	58.7 (1.9)	60 (0.7)	62.6 (0.5)	<0.001
Blood pressure, 95 <sup>th</sup> percentile <sup>d</sup>	0% (0)	0.7% (0.3)	2% (1.2)	N/A <sup>1</sup>
Cholesterol, total (mg/dL)	175.6 (4.7)	164.4 (2)	162.9 (2.8)	0.16
Triglycerides (mg/dL)	57.4 (2.8)	56.3 (1.1)	82.2 (7.9)	0.12
HDL (mg/dL)	57.4 (2.8)	56.3 (1.1)	56.2 (1.1)	0.82
LDL (mg/dL)	105.9 (9.2)	86.5 (2.2)	90.5 (3.6)	0.81
Insulin (uU/mL)	12.9 (1.6)	11.6 (0.6)	11.4 (0.8)	0.51
Glucose, fasting (mg/dL)	87 (1.2)	88.6 (1.1)	87.5 (1.1)	0.52
Glycohemoglobin (%)	5.2 (0.1)	5.1 (0.01)	5.1 (0.02)	0.33
HOMA-IR	2.8 (0.4)	2.6 (0.2)	2.5 (0.2)	0.43
Calories <sup>e</sup> (kcal/day)	2161 (106.8)	1974 (40.3)	2047 (44.4)	0.78
Computer time <sup>e</sup> (hr)	1.3 (0.3)	1.1 (0.1)	1.2 (0.1)	0.75
Television time <sup>e</sup> (hr)	2.2 (0.6)	2.2 (0.1)	2.2 (0.2)	0.88
Sedentary time <sup>f</sup> (mins)	472 (40.3)	351 (8.6)	524 (4.6)	<0.001
Diet score <sup>g</sup>	0.2 (0.1)	0.3 (0.1)	−0.4 (0.03)	<0.001
Sedentary time <sup>f</sup>	0.3 (0.3)	−0.7 (0.1)	0.7 (0.04)	<0.001
Vigorous physical activity <sup>f</sup> (min/d)	27.6 (8.9)	1.77 (0.8)	0.75 (0.3)	0.010
Moderate physical activity <sup>f</sup> (min/d)	124 (38.8)	14.3 (1.9)	11.9 (2.3)	<0.008
MVPA <sup>g</sup>	2.9 (0.5)	−0.2 (0.1)	−0.3 (0.05)	<0.001
Estimated $\dot{V}O_2$ max <sup>h</sup> (ml/kg/min)	38.3 (1.2)	36.9 (1)	37.6 (0.7)	0.69

<sup>a</sup>Cluster 1A: high diet scores, high MVPA, high sedentary time. Cluster 2A: high diet scores, low MVPA, and low sedentary time. Cluster 3A low diet scores, low MVPA, and high sedentary time.

<sup>b</sup>Percentage of participants with a BMI 85<sup>th</sup> to <95<sup>th</sup> percentile based on age and gender.

<sup>c</sup>Percentage of participants with a BMI 95<sup>th</sup> percentile based on age and gender.

<sup>d</sup>Percentage of participants with a systolic and/or diastolic blood pressure 95<sup>th</sup> percentile based on age, gender, and height.

<sup>e</sup>Caloric intake, duration of computer time and television time are self-reported.

<sup>f</sup>Sedentary time, moderate physical activity, and vigorous physical activity determined using an accelerometer.

<sup>g</sup>Diet score, Sedentary time, and MVPA are reported as normalized z-scores.

<sup>h</sup>Estimated  $\dot{V}O_2$  max, obtained from a submaximal cardiopulmonary exercise test.

<sup>i</sup>Sample size did not allow statistical comparison across categories.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

**Table 3**

Cardio-metabolic factors and health behaviors for boys in 2003–2004 NHANES, aged 12–19 years

	Cluster 1B <sup>a</sup>	Cluster 2B	Cluster 3B	p
	N=99	N = 360	N=194	N/A
	Mean (SD)	Mean (SD)	Mean (SD)	
Weight (kg)	62.2 (2.7)	70.4 (1.2)	73.7 (2.1)	0.003
BMI (kg/m <sup>2</sup> )	21.9 (0.6)	23.6 (0.3)	24 (0.6)	0.06
BMI, 85 <sup>th</sup> – <95 <sup>th</sup> percentile <sup>b</sup>	16.5% (5.3)	18% (3.8)	21.3% (5.3)	0.8
BMI, 95 <sup>th</sup> percentile <sup>c</sup>	12.3% (4.8)	20.7% (1.7)	15.2% (4.8)	0.33
Waist-to-Height Ratio	0.46 (0.01)	0.48 (0.01)	0.48 (0.01)	0.40
Blood pressure, systolic (mm Hg)	111.1 (1)	112.1 (0.9)	114.1 (1.2)	0.08
Blood pressure, diastolic (mm Hg)	54.4 (1.8)	57.8 (1)	61.1 (1.4)	0.02
Blood pressure, 95 <sup>th</sup> percentile <sup>d</sup>	1.7% (1.1)	2.2% (1.1)	2.3% (1.4)	0.96
Cholesterol, total (mg/dL)	159.9 (3.7)	155 (2.9)	161.2 (4.5)	0.46
Triglycerides (mg/dL)	75.7 (6.6)	97.6 (4.4)	105.5 (12.2)	0.17
HDL (mg/dL)	53.1 (1.7)	48.8 (1.1)	46.9 (1)	0.002
LDL (mg/dL)	88.5 (7.6)	86.3 (3.1)	96.6 (6.9)	0.24
Insulin (uU/mL)	10.4 (1.3)	12.2 (1.2)	13.8 (3.8)	0.53
Glucose, fasting (mg/dL)	88 (1.1)	92.3 (0.8)	92.4 (1.3)	0.13
Glycohemoglobin (%)	5.2 (0.03)	5.2 (0.01)	5.1 (0.03)	0.02
HOMA-IR	2.3 (0.3)	2.8 (0.3)	3.7 (1.4)	0.42
Calories <sup>e</sup> (kcal/day)	2786 (180.1)	2684 (74.7)	2797 (136.7)	0.69
Computer time <sup>e</sup> (hr)	0.9 (0.2)	1 (0.1)	1.5 (0.2)	0.03
Television time <sup>e</sup> (hr)	2.4 (0.2)	2.2 (0.1)	2.3 (0.1)	0.79
Sedentary time <sup>f</sup> (mins)	354 (8.1)	353 (10.5)	558 (16.1)	<0.001
Diet score <sup>g</sup>	0.3 (0.2)	–0.4 (0.1)	0.6 (0.1)	0.001
Sedentary time <sup>f</sup>	–0.4 (0.1)	–0.4 (0.1)	1 (0.1)	<0.001
Vigorous physical activity <sup>f</sup> (min/d)	26 (4.1)	3.7 (1)	4.3 (1.4)	0.001
Moderate physical activity <sup>f</sup> (min/d)	166.1 (11.7)	35.2 (4)	30.3 (5.6)	<0.001
MVPA <sup>g</sup>	1.7 (0.1)	–0.4 (0.04)	–0.4 (0.1)	<0.001
Estimated $\dot{V}O_2$ max <sup>h</sup> (ml/kg/min)	49.7 (1.4)	44.8 (0.7)	43.3 (0.9)	0.005

<sup>a</sup>Cluster 1B: high diet scores, high MVPA, and low sedentary time. Cluster 2B: low sedentary time, poor diet quality and low MVPA. Cluster 3B: high diet scores, low MVPA, and high sedentary time.

<sup>b</sup>Percentage of participants with a BMI 85<sup>th</sup> to <95<sup>th</sup> percentile based on age and gender.

<sup>c</sup>Percentage of participants with a BMI 95<sup>th</sup> percentile based on age and gender.

<sup>d</sup>Percentage of participants with a systolic and/or diastolic blood pressure 95<sup>th</sup> percentile based on age, gender, and height.

<sup>e</sup>Caloric intake, duration of computer time and television time are self-reported.

<sup>f</sup>Sedentary time, moderate physical activity, and vigorous physical activity determined using an accelerometer.

<sup>g</sup>Diet score, Sedentary time, and MVPA are reported as normalized z-scores.

<sup>h</sup>Estimated  $\dot{V}O_2$  max, obtained from a submaximal cardiopulmonary exercise test.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

**Table 4**

Sex-stratified, multivariate linear regression analysis of the clusters and maximal oxygen consumption for girls and boys of the 2003–2004 NHANES adolescents, aged 12–19 years

		Girls, <i>n</i> = 576			Boys, <i>n</i> = 649		
	$\beta$	SE <sup>a</sup>	P		$\beta$	SE	P
Cluster 1A	Ref <sup>b</sup>	Ref	Ref	Cluster 1B	Ref	Ref	Ref
Cluster 2A	-1.2	1.2	0.17	Cluster 2B	-4.6	1.6	0.01
Cluster 3A	-0.01	1	0.66	Cluster 3B	-7.2	1.9	<0.01

<sup>a</sup>SE, standard error

<sup>b</sup>Ref, reference cluster

The model adjusted for accelerometer wear-time, age, race/ethnicity, body mass index, and the poverty-to-income ratio as a categorical variable. The  $\beta$  coefficient represents the unit change in  $\dot{V}O_2$  max for being in the cluster.