Published in final edited form as:

Physiol Meas. 2016 November ; 37(11): 1981–1992. doi:10.1088/0967-3334/37/11/1981.

Using Symbolic Aggregate approXimation (SAX) to visualize activity transitions among older adults

Amal A. Wanigatunga¹, Paul V. Nickerson², Todd M. Manini¹, and Parisa Rashidi²

¹ Department of Epidemiology and Aging & Geriatric Research, University of Florida, Gainesville, Florida

² Department of Biomedical Engineering, University of Florida, Gainesville, Florida

Abstract

The objective of this analysis was to apply Symbolic Aggregate approXimation (SAX) time-series analysis to accelerometer data for activity pattern visualization stratified by self-reported mobility difficulty. A total of 2,393 (71.6 ± 7.9 years old) participants wore an accelerometer on the hip (4+ days; 10+ hours) during the National Health and Nutrition Examination Survey (NHANES), a biannual series of health assessments of the US population. One minute epoch data was used to perform SAX, which converted accelerometry time series data into four activity levels. Intelligent icons of normalized activity transition prevalence, a visual representation of time-series data, were examined among those who self-reported mobility difficulty. Mobility difficulty questions assessed various levels difficulty performing activities such as walking a quarter mile, walking up ten steps, stooping/crouching/kneeling, and walking between rooms on the same floor. Daily activity counts were estimated across difficulty level using weighted-linear regression after adjusted for demographics, lifestyle factors, medical conditions, and accelerometer wear time. Those reporting higher mobility difficulty tended to be older, female, less educated, not married, and smokers. Additionally, those with higher mobility difficulty self-reported lower education, lower income, lower moderate-to-vigorous physical activity, and higher history of adverse medical conditions. Using SAX-derived intelligent icons, those with no difficulty showed high variations of transitions across all activity levels. With higher difficulty, the variations in transitions were lower and constricted around low activity level transitions. Among those who reported unable/don't do mobility-related activities, there were only transitions in the lower tiered activity levels with high prevalence of prolonged low level activity. Adjusted estimates of daily activity were lower as higher difficulty reported occurred but only significant for those reporting unable/don't do mobility-related activity ($p < 0.01$). In summary, this analysis showed apparent differences in stratified activity patterns even when traditional regression analyses on volumetric accelerometer data yielded null results.

Keywords

discrete representation; accelerometer; activity transition; older adults; mobility difficulty

1. INTRODUCTION

Large-scale research efforts to monitor physical activity (PA) are turning to the use of accelerometers because these devices provide an objective and non-invasive way to continuously collect PA data (1). Additionally, accelerometry yields PA in dimensions of volume, intensity, accumulation, and frequency. However, accelerometer data is large, complex, and difficult to interpret unless reduced to the traditional volumetric measures of time spent in PA. The growing body of accelerometer literature in older adults typically report means and standard deviations of PA averaged over an interval of time (e.g. per day or per week) (2). However, emerging studies have derived ways to elucidate patterns of accelerometer data in terms of bouts or intervals of consecutive minutes spent in PA categories (3-6). Though these bouts are still characterizations of PA volume, breaks or interruptions in these bouts give information of how and when volumes of PA are accrued, giving one of the first methods to assess transitions within daily PA routines.

Activity transitions are an important complement to measures of volume towards accurately profiling daily activity patterns. A large contributor to the way activity is accumulated is through walking and ambulation-like movements, an important predictor of health-related outcomes in adults 60 years and older (7). In sedentary behavior literature, transitions out of prolonged sedentary behaviors (otherwise known as breaks in sedentary time) are particularly important due to associated health benefits on cardio-metabolic biomarkers (8-10). Additionally, within-level PA activity transitions (e.g. engaging in various levels of PA) are just as important, particularly among older adults (11, 12), a population least likely to participate in moderate-to-vigorous intensity levels of PA (MVPA) (13). Yet, there is a gap in the present literature for a simple approach to identify total daily activity transition patterns when combining both activity and inactivity.

To fill this gap, this analysis used Symbolic Aggregate approXimation, a time-series analysis method commonly used in the computer science field (14), to visualize PA patterns derived from accelerometer data collected through the National Health and Nutrition Examination Survey (NHANES). The SAX method converts time series into discrete representations to identify patterns otherwise known as motifs. The primary goal of this analysis was to visualize PA transition motifs in older adults who self-reported mobility difficulty. We hypothesized that patterns of activity would shift from high activity level patterns (sustained bouts and transitions) to low levels of activity patterns as higher mobility difficulty was reported. A second goal of this analysis was to use a traditional statistical approach to support the SAX visualization. We hypothesized that the average of total daily activity would be negatively associated with self-reported mobility difficulty.

2. METHODS

2.1. Study population

The NHANES study is an ongoing, cross-sectional, observational study which implements a stratified, complex, multistage probability design to collect data from a national, representative sample of noninstitutionalized US citizens. While data is gathered from those 6 years old and older, specific populations such as non-Hispanic black Americans are

oversampled. The Centers for Disease Control and Preventions (CDC) oversees this research program and has gained approved through the National Center for Health Studies Ethics Review Board while obtaining informed consent from all participants. Participant data was collected by questionnaires via home interviews and health assessments using mobility examination centers (MECs). During the 2003-2004 and 2005-2006 cohorts, accelerometer data was collected from eligible participants (e.g. ambulatory, able to wear the device). Participants were asked to wear a custom-fit belt with the accelerometer positioned on the right hip for seven sequential days after assessments were conducted in the MECs. Participants were asked to take off the device during sleep and water-related activities. Further details on the accelerometer protocol have been reported elsewhere (15). In 2003-2004 and 2005-2006 cohorts, data was collected in a total of 20,470 participants. For this analysis, participants who had complete data on interview-based questions on mobility-related questions (n=4,718) and only those aged 60 years and older were included (n=3,471). Our final sample was 2,393 participants who collected at least 4 days of 10+ daily hours of valid accelerometry. Due to the low variability between weekday and weekend activity among older adults, the inclusion of weekend days within valid wear time was not required for this analysis (16, 17).

2.2. Participant characteristics

Sociodemographics included self-reported age collected at the time of screening; race/ethnicity as a combined measure of race and Hispanic origin, income as an estimate of total household income; education as the highest grade or level of education; married as either married or living with a partner; and whether covered by health insurance. Lifestyle factors included self-reported general health, body mass index calculated as kg/m^2 (weight and height were measured at the MECs), self-reported current smoking status, alcohol use defined as ever having 5 or more drinks every day, daily MVPA derived from accelerometry (minutes spent in 2,020+ counts/min (13)), wear time reported as mean wear hours per day, and NHANES cohort assignment. Medical conditions were self-reported as ever been told by a health practitioner of having: congestive heart failure, coronary heart disease, angina, myocardial infarction or heart attack, stroke, pulmonary disease including emphysema or chronic bronchitis, diabetes or sugar diabetes, cancer or malignancy, arthritis, or any fracture of the hip, wrist, and/or spine.

2.3. Accelerometer-derived physical activity

Accelerometers, devices which collect movement continuously, were used to measure PA. ActiGraph AM-7614 monitors (Ft. Walton, Florida, formerly CSA/MTI AM-7164) were utilized to collect movement on the vertical axis from the hip, which has been shown to be most associated with locomotion-type movement (18). This uniaxial monitor collected data as a magnitude of acceleration (intensity) and duration of acceleration (volume). Raw acceleration data was converted to activity counts which are comparable units of band-passed filtered (to eliminate non-human movement) accelerometer data. Then, activity counts were summarized over a 1 minute epoch. We applied a non-wear algorithm provided by the NCI (National Cancer Institute) (19) to identify non-wear periods of at least 90 minutes of consecutive zero counts with a 2 minute limit of artifactual movement. Minutes not labeled as non-wear were considered valid wear periods (20).

2.4. Symbolic Aggregate approXimation (SAX) activity transitions

First, activity counts within total valid wear time were normalized per participant. Each normalized value was determined by taking the area under the curve made by all activity counts (analogous to the empirical cumulative distribution function [ECDF] in non-parametric statistics). Next, data was down-sampled using a method named Piecewise Aggregate Approximation (PAA) which divides the time series data into equi-sized frames, and the average value of each frame replaced all the previous values within that frame. For this analysis, we chose to use a 30-minute scanning window for PAA in our primary analysis, depicted in Figure 1. This chosen length is somewhat arbitrary and can be easily changed based on the population being studied. While a shorter window increases the resolution, it also increases the frequency of transitions that might be less meaningful in a sedentary population. A longer window will decrease the resolution, reducing the frequency of transitions making it more difficult to detect meaningful transitions in a highly active population. A pre-determined number of activity levels, denoted as β where $\beta=10$ for this analysis, was used to equally bin the activity levels according to the underlying distribution which generated accelerometer cut-points relative to the population activity. Two-step motifs were generated which represented activity transitions within a 60-minute period. For example, if the lowest activity was labeled “A” and the second lowest activity level was labeled “B”, an “AB” motif represents an activity transition from the lowest activity level to the second-lowest activity level within a 60-minute period. The number of occurrences for each motif type was then divided by the total number of all motif patterns found in the analytic sample to output the prevalence of the specific motif type. Because wear times varied between participants, we normalized the prevalence value. To do this, the observed prevalence value motif type distributions were compared to the ECDF through bootstrapping by assigning random participants into artificial groups. Then, observed prevalence values were assessed against the ECDF of resampled values. The subsequent derived value, which lies between 0 and 1, represents the normalized proportion of observed motif type prevalence greater than the total potential motif types.

2.5. SAX intelligent icons

Intelligent icons, otherwise known as time-series bitmaps (21), were constructed to visualize normalized prevalence of level transitions by motif type via heat maps. Figure 2 illustrates a map of an intelligent icon of activity transitions. Activity that was sustained without transitions run diagonally in the middle of the icon and increase intensity from left to right. Below the line of sustained activity are “high to low” transitions and above are “low to high” transitions. Since normalized activity transition motifs were bound between 0 and 1, a color was assigned to that value and transitions were arranged according to Figure 2.

2.6. Self-reported difficulty in mobility

A 20-item physical functioning questionnaire collected self-reported difficulty information on a variety of physical activities. Participants were asked for each activity: “By {yourself/himself/herself} and without using any special equipment, how much difficulty {do you/does survey participant} have...?” Four items assessed mobility difficulty which included questions about walking a quarter mile, walking up ten steps, stooping/crouching/kneeling,

and walking between rooms on the same floor. A composite mobility score was created according to categories of reporting “no difficulty”, “some difficulty”, “much difficulty”, and “unable/don't do”. First, we dichotomized those who reported any difficulty versus no difficulty to create the “no difficulty” group. Among those who reported any difficulty, participants who reported much difficulty for at least one of the questions were separated into the much difficulty group while the rest were included in the group who reported some difficulty. Among the entire sample, those who reported “unable to do” and “do not do this activity” were collapsed into the “unable/don't do” group.

2.7. Statistical approach

Unweighted characteristics of the sample were described as either means/standard deviations or frequencies/percentages across sociodemographics, lifestyle factors, medical history, and accelerometer wear measures. Differences among groups were tested between categorical variables using chi-square tests and with analysis of variance tests for continuous variables. Weighted multiple linear regression was used to estimate average PA intensity (activity counts/day) through the day across categories of self-reported mobility difficulty and the alpha level was set at 0.05. Covariates to adjust estimates of activity counts/day included sociodemographics, lifestyle factors, and medical conditions. The 2 year population weights for MEC were divided by 2 to rescale the overall MEC weight to match the survey population at the midpoint of the study periods. Python version 2.7 (Python Software Foundation. Python Language Reference, at <http://www.python.org>), SAS version 9.4 (SAS Institute, Cary, NC, 2004), and STATA version 13 (Stata Corporation; College Station, TX) were used to perform accelerometer pre-processing, SAX application, and post-processing analyses.

3. RESULTS

3.1. Participant characteristics

Participant characteristics stratified by self-reported mobility difficulty were reported in Table 1. Participants who reported either much difficulty or unable/don't do tended to be older, female, less educated, have lower income, and not married while there were no observed differences in non-Hispanic whites and having health insurance across groups. Participants who self-reporting higher difficulty tended to report poorer health, had high BMI, engaged in less daily MVPA but there were no differences in smoking status. History of adverse medical conditions was higher as self-reported difficulty was higher except for pulmonary disease ($p=0.07$) and cancer ($p=0.16$).

3.2. Activity transitions

Intelligent icons of activity transitions stratified by self-reported mobility difficulty were illustrated in Figure 3. Those who reported no difficulty had high transition prevalence in all activity levels (low-to-high and high-to-low) with high prolonged engagement in the highest activity level. Time spent in lower activity levels and subsequent transitions within those levels were far underrepresented. We observed a similar pattern among those who reported some mobility difficulty but the amount of lower activity levels were higher, with less low-to-high transitions. There was no sustained activity engagement in the highest activity levels,

but prolonged engagement in the lowest levels of activity became prevalent. Among those who reported much difficulty, there was a clear representation of almost fully sustained low level activity and transitional activity at the lowest levels. The clearest difference in activity transitions was observed among those reporting unable/don't do, where sustained and transitional activity was constricted to the lower intensities of activity. There were still some transitions into and out of higher levels of activity but these were less prevalent. However, no gaps were observed at the lowest activity intensities for both sustained activity and transitions between activity levels.

3.3. Activity count by self-reported mobility difficulty

Associations between daily PA intensity and self-reported mobility difficulty were estimated with the traditional approach of linear regression to understanding activity stratified by population group (Table 2). In Models 1 and 2, those reporting no difficulty had the highest average activity counts/day. Those reporting some, much, and unable/don't do had statistically lower estimates compared to the no difficulty group in both models. After further covariate adjustments in Models 3 and 4, estimated counts/day was lower as mobility difficulty reported became higher. However, only those reporting unable/don't do mobility-related activities had statistically lower daily activity compared to the no mobility difficulty group ($p < 0.01$ for both models).

4. DISCUSSION

To our knowledge, this is the first study to apply a discretization technique to accelerometer time series data to visualize activity transitions at a population level. Among a US sample of adults 60 years and older who self-reported status on mobility difficulty, we were able to visualize both prolonged and transitional activity based on daily cumulative activity throughout daily waking periods. We found that as difficulty became more severe, transitions and sustained activity shifted from high to low intensity levels. We also observed that sustained high level activity engaged by those reporting no mobility difficulty was not presented in groups with any reported mobility difficulty. Additionally, sustained low level activity was not present in the group who reported no difficulty, but engaged by groups with any mobility difficulty. Furthermore, activity transitions starting from low and ending in higher activity intensity levels were the first to be lost before any other type of transition pattern. To confirm the SAX visualizations, we estimated average daily activity counts by mobility difficulty status through a traditional linear regression approach. We found those reporting any mobility difficulty or those who don't participate in mobility-related activities engaged in less overall daily activity than those reporting no mobility difficulty when only adjusting for age and wear time. After adjusting for sociodemographics, lifestyle factors, and medical conditions, only those who were unable/did not do mobility-related activity accumulated significantly lower daily activity. This reveals that these covariates play a greater role in determining daily activity among groups who reported some or much mobility difficulty when compared to those without difficulty.

By discretizing activity and transitions between activity levels, we visualized activity patterns stratified by the health-related factor of mobility difficulty. While complex

algorithms are now being applied to accelerometer data due to the increasing impetus to interpret raw sensor-collected daily PA, issues arise computationally with high dimensionality and increased difficulty with interpretability when applying intelligent system analyses (22, 23). The SAX method is a simple time-series technique to reduce dimensionality of accelerometer data while also maintaining lower bounding to the original time series data. Moreover, this method is flexible to user-defined activity periods and levels which may be further appropriate to condition-specific populations. A potential use of this technique is to explore longer activity motifs and determine unique patterns against directly observed activity (such as walking at usual pace) conducted in a laboratory setting. Then, using free-living accelerometer data, the stratified differences of engaging in the investigated pattern can be examined across population-based factors such as age, functional status, and disease condition. Beyond visualization, intelligent icons can be assigned similarity scores by deriving the cosine of the angle of ECDF vector components. Cosine similarity can be assessed where the higher the value between icons, the greater the similarity between patterns which provides a way identify meaningful patterns (24).

A recent study using the same NHANES cohort as this analysis reported patterns of sedentary time and activity in terms of bouts (consecutive minutes in specified activity/inactivity level) and breaks (interruptions in bouts) (25). This study was assessed among individuals with and without mobility disability, as similar designation to this analysis's use of mobility difficulty. Their findings parallel our own, where people with mobility disability were found to be less active and engage in more prolonged sedentary behavior than those without mobility disability. Even though we used relative activity cut-points rather than absolute thresholds based on energy expenditure, did not distinguish between activity and inactivity, and defined mobility difficulty differently from the same questionnaire, we showed that those with no difficulty engaging in higher activity levels and transitioning more in these higher levels compared to those reporting any mobility difficulty. We also described how breaks in these levels occurred between functional groups, and characterized different types of breaks according to direction (high to low) and intensity (low to lowest).

Limitations of this analysis require acknowledgement. One limitation is the method by which mobility difficulty was collected. Self-report data may be impacted by social desirability and recall bias and may lead to underestimated effects. Moreover, we acknowledge that those who report “do not do this activity” are not typically categorized as “unable to do”. However, after running sensitivity analyses with both groups separated, there results did not change. Another limitation is the way accelerometer data is categorized by PA intensity. Traditionally, absolute cut-points calibrated from energy expenditure studies are used in objectively-collected PA literature. However, absolute cut-points are subject to misclassification, particularly by age. For example, cut-points for MVPA generated from Freedson et al. were calibrated using younger adults and traditionally used in a vast amount of literature (26). However, these thresholds are not appropriate for older adults as this age segment generally does not reach the higher levels of PA. Furthermore, cut-points generated for older adults such as in Copeland et al. study may be age appropriate but not calibrated for functional status (27). For this analysis, we chose to use relative activity cut-points to avoid this type of misclassification. However, SAX is flexible to where the user can set thresholds. Therefore, absolute cut-points such as those calibrated with energy expenditure

can be implemented if deemed necessary. Though relative activity cut-points increase comparability across population factors such as age and functional status, using relative activity introduces interpretability issues towards understanding activity transitions related to energy expenditure. Nonetheless, relative activity representations give a descriptive map of target goals for intervention development to either reduce prolonged inactivity by transition type, or increase activity tailored to functional status. Other approaches may be taken, for example, if “high to low” activity transitions are lost before any other type of transitions, a behavioral component to increase confidence may be prioritized over increasing activity when tailoring physical activity interventions. Further research should apply statistical techniques to SAX visualizations that can be used at the individual level to apply to subgroups.

5. CONCLUSION

To our knowledge, this is the first study to discretize activity patterns relative to an older adult population by self-reported mobility difficulty status. The activity maps showed those who reported no difficulty are active, engage in high-level activity transitions, and spent little to no prolonged time spent in lower activity levels compared to those reporting any mobility difficulty. As self-reported mobility difficulty becomes higher, low-to-high transitions became less prevalent than all other transitions and eventually prolonged inactivity was observed with no occurrences of high intensity level activity transitions. These visualizations provide useful information towards targeting specific activity patterning rather than volume of activity engaged.

ACKNOWLEDGEMENTS

AW is supported by a Diversity Supplement awarded by NIH/NIA P30AG028740 Claude D. Pepper Older Americans Independence Center. PN is supported by a Diversity Supplement awarded by NIH/NIGMS R01GM114290-01. We would like to acknowledge Dr. Patrick J. Tighe (ptighe@anest.ufl.edu), who initially applied the SAX methodology to pain data. We adopted this methodology towards accelerometry data from this effort. We would also like to acknowledge the Data Science Core at the University of Florida's Claude D. Pepper Older Americans Independence Center who provided assistance with managing and processing the raw accelerometry count data.

REFERENCES

1. Troiano RP, McClain JJ, Brychta RJ, Chen KY. Evolution of accelerometer methods for physical activity research. *British journal of sports medicine*. 2014; 48(13):1019–23. [PubMed: 24782483]
2. Bassett DR, Troiano RP, McClain JJ, Wolff DL. Accelerometer-based physical activity: total volume per day and standardized measures. *Medicine and science in sports and exercise*. 2015; 47(4):833–8. [PubMed: 25102292]
3. Glazer NL, Lyass A, Esliger DW, Blease SJ, Freedson PS, Massaro JM, et al. Sustained and shorter bouts of physical activity are related to cardiovascular health. *Medicine and science in sports and exercise*. 2013; 45(1):109. [PubMed: 22895372]
4. Shiroma EJ, Freedson PS, Trost SG, Lee IM. Patterns of accelerometer-assessed sedentary behavior in older women. *JAMA*. 2013; 310(23):2562–3. [PubMed: 24346993]
5. Straker L, Campbell A, Mathiassen SE, Abbott RA, Parry S, Davey P. Capturing the pattern of physical activity and sedentary behavior: exposure variation analysis of accelerometer data. *J Phys Act Health*. 2014; 11(3):614–25. [PubMed: 23416959]
6. Orme M, Wijndaele K, Sharp SJ, Westgate K, Ekelund U, Brage S. Combined influence of epoch length, cut-point and bout duration on accelerometry-derived physical activity. *International Journal of Behavioral Nutrition and Physical Activity*. 2014; 11(1):1. [PubMed: 24405936]

7. Hardy SE, Kang Y, Studenski SA, Degenholtz HB. Ability to walk 1/4 mile predicts subsequent disability, mortality, and health care costs. *Journal of general internal medicine*. 2011; 26(2):130–5. [PubMed: 20972641]
8. Bankoski A, Harris TB, McClain JJ, Brychta RJ, Caserotti P, Chen KY, et al. Sedentary activity associated with metabolic syndrome independent of physical activity. *Diabetes Care*. 2011; 34(2): 497–503. [PubMed: 21270206]
9. Dunstan DW, Howard B, Healy GN, Owen N. Too much sitting—a health hazard. *Diabetes research and clinical practice*. 2012; 97(3):368–76. [PubMed: 22682948]
10. Healy GN, Dunstan DW, Salmon J, Cerin E, Shaw JE, Zimmet PZ, et al. Breaks in sedentary time beneficial associations with metabolic risk. *Diabetes care*. 2008; 31(4):661–6. [PubMed: 18252901]
11. Committee PAGA. Physical activity guidelines advisory committee report, 2008. Vol. 2008. US Department of Health and Human Services; Washington, DC: 2008. p. A1-H14.
12. Nelson ME, Rejeski WJ, Blair SN, Duncan PW, Judge JO, King AC, et al. Physical activity and public health in older adults: recommendation from the American College of Sports Medicine and the American Heart Association. *Circulation*. 2007; 116(9):1094. [PubMed: 17671236]
13. Troiano RP, Berrigan D, Dodd KW, Masse LC, Tilert T, McDowell M. Physical activity in the United States measured by accelerometer. *Medicine and science in sports and exercise*. 2008; 40(1):181–8. [PubMed: 18091006]
14. Lin, J.; Keogh, E.; Lonardi, S.; Chiu, B., editors. Proceedings of the 8th ACM SIGMOD workshop on Research issues in data mining and knowledge discovery. ACM; 2003. A symbolic representation of time series, with implications for streaming algorithms..
15. NHANES. US Department of Health and Human Services, Centers for Disease Control and Prevention; Hyattsville (MD): 2005-2006. Documentation, codebook and frequencies: physical activity monitor.
16. Marshall S, Kerr J, Carlson J, Cadmus-Bertram L, Patterson R, Wasilenko K, et al. Patterns of Weekday and Weekend Sedentary Behavior Among Older Adults. *Journal of aging and physical activity*. 2015; 23(4):534–41. [PubMed: 25415205]
17. Nicolai S, Benzinger P, Skelton DA, Aminian K, Becker C, Lindemann U. Day-to-day variability of physical activity of older adults living in the community. *J Aging Phys Act*. 2010; 18(1):75–86. [PubMed: 20181995]
18. Bouten CV, Koekkoek KT, Verduin M, Kodde R, Janssen JD. A triaxial accelerometer and portable data processing unit for the assessment of daily physical activity. *IEEE Transactions on Biomedical Engineering*. 1997; 44(3):136–47. [PubMed: 9216127]
19. SAS Programs for Analyzing NHANES 2003-2004 Accelerometer Data: National Cancer Institute. 2014
20. Choi L, Liu Z, Matthews CE, Buchowski MS. Validation of accelerometer wear and nonwear time classification algorithm. *Medicine and science in sports and exercise*. 2011; 43(2):357–64. [PubMed: 20581716]
21. Kumar, N.; Lolla, VN.; Keogh, EJ.; Lonardi, S.; Ratanamahatana, C., editors. SDM. SIAM; 2005. Time-series Bitmaps: a Practical Visualization Tool for Working with Large Time Series Databases..
22. Mannini A, Sabatini AM. Machine learning methods for classifying human physical activity from on-body accelerometers. *Sensors*. 2010; 10(2):1154–75. [PubMed: 22205862]
23. Ravi N, Dandekar N, Mysore P, Littman ML. Activity recognition from accelerometer data. *AAAI*. 2005
24. Tighe PJ, Nickerson P, Fillingim RB, Rashidi P. Characterizations of Temporal Postoperative Pain Signatures with Symbolic Aggregate Approximations. *The Clinical journal of pain*. 2016
25. Manns P, Ezeugwu V, Armijo-Olivo S, Vallance J, Healy GN. Accelerometer-Derived Pattern of Sedentary and Physical Activity Time in Persons with Mobility Disability: National Health and Nutrition Examination Survey 2003 to 2006. *J Am Geriatr Soc*. 2015; 63(7):1314–23. [PubMed: 26173621]

26. Freedson PS, Melanson E, Sirard J. Calibration of the Computer Science and Applications, Inc. accelerometer. *Medicine and science in sports and exercise*. 1998; 30(5):777–81. [PubMed: 9588623]
27. Copeland JL, Esliger DW. Accelerometer assessment of physical activity in active, healthy older adults. 2009

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Figure 1.
Activity levels derived from Symbolic Aggregate Approximation

Figure 2.
Intelligent icon map of activity transitions

Figure 3. SAX-derived intelligent icons stratified by self-reported mobility difficulty. Note that prevalence of transitions is marked by light colors. Higher prevalence is marked with increasing darker colors.

Table 1

Unweighted participant characteristics with valid accelerometry who self-reported mobility difficulty

	None n=1,054	Some n=679	Much n=321	Unable/don't do n=339	p-value
Sociodemographics					
Age, mean(SD)	70.1 (7.4)	72.0 (7.8)	72.6 (8.4)	74.7 (8.3)	<0.001
80+, n(%)	161 (15.3)	152 (22.4)	94 (29.3)	131 (38.6)	<0.001
Female, n(%)	456 (43.3)	335 (49.3)	170 (53.0)	225 (66.4)	<0.001
Non-Hispanic White, n(%)	616 (58.4)	425 (62.6)	193 (60.1)	215 (63.4)	0.227
< High school education, n(%)	334 (31.7)	265 (39.0)	130 (40.5)	158 (46.6)	<0.001
< \$20,000, n(%)	220 (20.9)	211 (31.1)	132 (41.1)	135 (39.8)	<0.001
Married, n(%)	716 (67.9)	410 (60.4)	176 (54.8)	152 (44.8)	<0.001
Have health insurance, n(%)	975 (92.5)	635 (93.5)	303 (94.4)	321 (94.7)	0.693
Lifestyle factors					
2003-2004 NHANES cohort	548 (52.0)	359 (52.9)	198 (61.7)	205 (60.5)	0.002
Fair to poor health, n(%)	172 (16.3)	191 (28.1)	137 (42.7)	155 (45.7)	<0.001
Body mass index (kg/m ²), mean(SD)	26.9 (4.5)	28.6 (5.2)	30.1 (6.7)	30.3 (7.0)	<0.001
Current smoker, mean(SD)	131 (12.4)	85 (12.5)	41 (12.8)	36 (10.6)	0.754
5 or more drinks/day ever, n(%)	130 (12.3)	92 (13.6)	47 (14.6)	39 (11.5)	0.012
Daily MVPA min, mean(SD)	14.5 (16.7)	8.1 (12.2)	5.7 (10.4)	3.9 (8.5)	<0.001
Wear time hrs, mean(SD)	14.7 (1.9)	14.5 (2.1)	14.3 (2.1)	14.3 (2.3)	0.009
Medical conditions					
Congestive heart failure, n(%)	39 (3.7)	49 (7.2)	36 (11.2)	56 (16.5)	<0.001
Coronary heart disease, n(%)	80 (7.6)	68 (10.0)	43 (13.4)	54 (15.9)	<0.001
Angina, n(%)	45 (4.3)	53 (7.8)	51 (15.9)	46 (13.6)	<0.001
Myocardial infarction, n(%)	79 (7.5)	63 (9.3)	62 (19.3)	51 (15.0)	<0.001
Stroke, n(%)	41 (3.9)	62 (9.1)	39 (12.2)	43 (12.7)	<0.001
Pulmonary disease, n(%)	13 (1.2)	10 (1.5)	8 (2.5)	10 (3.0)	0.116
Diabetes, n(%)	171 (16.2)	150 (22.1)	93 (29.0)	102 (30.1)	<0.001
Cancer, n(%)	193 (18.3)	132 (19.4)	63 (19.6)	83 (24.5)	0.141
Arthritis, n(%)	345 (32.7)	386 (56.9)	232 (72.3)	246 (72.6)	<0.001
Fracture, n(%)	117 (11.1)	92 (13.6)	60 (18.7)	74 (21.8)	<0.001

Note: MVPA – moderate to vigorous physical activity defined as accelerometry minutes registering over 2,020 counts/min

Table 2

Adjusted daily physical activity intensity (counts/min) by self-reported mobility difficulty

	Model 1			Model 2			Model 3			Model 4		
	Mean	SE	p-value	Mean	SE	p-value	Mean	SE	p-value	Mean	SE	p-value
No difficulty	244.2	3.0	Reference	237.6	4.1	Reference	215.3	3.0	Reference	214.8	3.0	Reference
Some	188.3	3.8	<0.001	190.3	5.1	<0.001	207.7	3.5	0.109	208.0	3.4	0.144
Much	169.6	5.7	<0.001	169.1	7.4	<0.001	205.7	5.3	0.135	206.5	5.3	0.203
Unable/don't do	154.0	5.6	<0.001	157.5	7.5	<0.001	191.2	5.7	<0.001	191.7	5.7	0.001

Note: Model 1 adjusted for age and wear time. Model 2 adjusted for Model 1 + race/ethnicity, sex, education, and income. Model 3 adjusted for Model 2 + cohort year, marital status, health insurance, self-reported health status, body mass index, smoking status, average # of alcoholic drinks/ in the past year, moderate-to-vigorous physical activity. Model 4 adjusted for Model 3 + history of congestive heart failure, coronary heart disease, angina, myocardial infarction, stroke, pulmonary disease, diabetes, cancer, arthritis, fracture.