



**Research  
with Heart.**

## ARIC Manuscript Proposal Form

### ARIC Publication Admin Use Only

Publication Committee Review Date: **07/09/24**

ARIC Manuscript Proposal Number: **#4495**

#### 1.a. Full Title:

Deploying and comparing multiple open-source step counting algorithms in the ARIC study

**b. Abbreviated Title (Length 26 characters):** Step algorithms in ARIC

#### 2. Writing Group [please provide a middle initial if available; EX: Adam L Williams]:

Writing group members:

Sunan Gao, BE,  
Xinkai Zhou, PhD,  
Lily J. Koffman, MS,  
Amal A. Wanigatunga, PhD, MPH,  
Jennifer A. Schrack, PhD, MS,  
Ciprian Crainiceanu, PhD,  
John Muschelli, PhD

I, the first author, confirm that all the coauthors have given their approval for this manuscript proposal. SG [please confirm with your initials electronically or in writing]

**First author [please provide a middle initial; EX: Adam L Williams]:**

**Sunan Gao**  
First name Middle initial Last name

Address: 2024 E. Monument Street, Suite 2-700, Baltimore, MD 21287

Phone: 667-335-6100

E-mail: sgao57@jh.edu

**ARIC author** to be contacted if there are questions about the manuscript and the first author does not respond or cannot be located (The ARIC author should be involved enough in ARIC to be able to point the lead author to appropriate ancillary study PIs and to be able to search ARIC manuscript proposals if the lead author doesn't have the access needed to do such a search).

Name: John Muschelli

Address: 615 N. Wolfe Street, Room E3136

Baltimore, MD 21205

Phone: 410-955-3067

E-mail: muschellij2@gmail.com

### **3. Timeline:**

Statistical Analysis: 3 months

Manuscript Preparation: 4 months

### **4. Rationale:**

Physical activity (PA) has been shown to reduce morbidity and mortality associated with multiple chronic conditions, including cardiovascular disease, type 2 diabetes, and several cancers, and is linked to improved quality of life (Paluch et al., 2022). The step count, a simple yet popular approach for providing PA targets to the general public, is easy to understand and remember, making it an ideal metric for PA and disease prevention guidelines (del Pozo Cruz et al., 2022a). With the increasing popularity of wearable monitors and mobile devices, monitoring daily steps has become more feasible than ever for the public. However, estimating steps using commercial devices rarely allows access to raw data or provides open-source algorithms for processing data into clinically meaningful endpoints, limiting reproducibility, generalizability, and scalability (Strackiewicz et al., 2023a, 2023b). In contrast, employing open-source algorithms to estimate step counts using high-density tri-axial data from wearable PA monitors (e.g., accelerometers) is a potentially promising alternative that could provide harmonized and interpretable data across and within studies (Karas et al., 2019).

Open-source algorithms using raw (sub-second) accelerometry measurements have been proposed to estimate step counts (Ducharme et al., 2021; Karas et al., 2021a; Maylor et al., 2022; Small et al., 2023; Strackiewicz et al., 2023a). However, obtaining true step counts in free living settings is difficult due to privacy, cost, and logistics. As such, most open-source algorithms are only validated in controlled or semi-controlled settings, and accuracy in free-living varies widely (Maylor et al., 2022). As physical activity has been shown to be the strongest predictor of mortality (Sheng et al., 2021), public health recommendations are increasingly based on the number of steps, which in turn may depend on the algorithm used for estimation.

Published papers note problems with the number of steps estimated using currently available algorithms, which, on average are around 9,000 to 10,000 steps per day even in older populations (del Pozo Cruz et al., 2022b, Sheng et al., 2021; Maylor et al., 2022; Paluch et al., 2022). This seems to be an overestimate that requires further investigation. Indeed, there is a general lack of transparency on exactly what algorithms provide and what, if any, open-source step counting algorithms are comparable. This can have profound effects on recommendations and intervention studies, where a difference of a few thousand steps may make a huge difference.

We propose to investigate: (1) whether it is feasible to deploy open source step counting algorithms in the ARIC study; (2) compare results of the algorithms across the population and their predictive performance; (3) compare the absolute number of steps and discuss implications for health recommendations. The paper also generates new variables that provide the estimated number of steps in the ARIC population, one of the premier studies of cardiovascular and neurocognitive health in a diverse population. As high-resolution data are collected and stored and each algorithm is known and fixed, the process is perfectly reproducible and can be applied to the same study and/or other studies to produce harmonized measures of number of steps.

To our knowledge, no prior study has focused on comparing the scale and distribution differences in generated step counts and explored whether the association results between estimated steps and health outcomes are sensitive to the choice of open-source step algorithms within the same dataset.

We will deploy five open-source algorithms in more than 1,400 study participants in the ARIC study who had high-resolution wrist accelerometry data collected at V9. We hypothesize that the estimated steps will exhibit high correlations and similar distributions, demonstrate consistent directionality and significance in association results, but differ in scale. Thus, different algorithms may produce similar predictive performance for health outcomes but vary in their interpretation of steps. This analysis will provide a critical perspective on step count algorithms and offer insights into the challenges involved in translating research findings into public health recommendations. Also, the existence of open-source reproducible software will make harmonization of data with other studies feasible. For example, one of the algorithms was designed for and applied to UK Biobank, thus harmonizing ARIC and UK Biobank in terms of the number of steps estimated by one of the algorithms.

## **5. Main Hypothesis/Study Aims:**

**Aim 1:** To generate step counts utilizing five distinct open-source step count algorithms (i.e., ADEPT, Oak, Step Detection Threshold, Verisense, and Stepcount) on sub-second accelerometry data collected from wrist-worn Actigraphs in ARIC Study Visit 9. Subsequently, we will compare and visualize the distribution of the step counts estimated by each open-source algorithm.

**Subaim 1a:** To identify the preferred step algorithm(s) by evaluating the distribution and plausibility of estimated step counts, considering factors such as expected daily step ranges for the ARIC population, consistency across demographic subgroups, and alignment with previous literature on step counts in older adults.

Hypothesis 1: We hypothesize that the step counts derived from these open-source algorithms will vary substantially in terms of number of steps within the population, yet they will be highly correlated. We further hypothesize that one algorithm will emerge as preferred based on producing more plausible step count distributions for ARIC population.

**Aim 2:** To evaluate the cross-sectional linear and non-linear associations between steps estimated by the five algorithms and health outcomes, with a particular emphasis on metabolic diseases, cardiovascular diseases, and cognitive function.

Hypothesis 2: We postulate that step counts will be associated with health outcomes. For a given outcome, we also postulate that the direction of the effect will be consistent between step count algorithms, but the effect sizes will vary. And after the scaling of estimated step counts, the effect size will be similar and comparable.

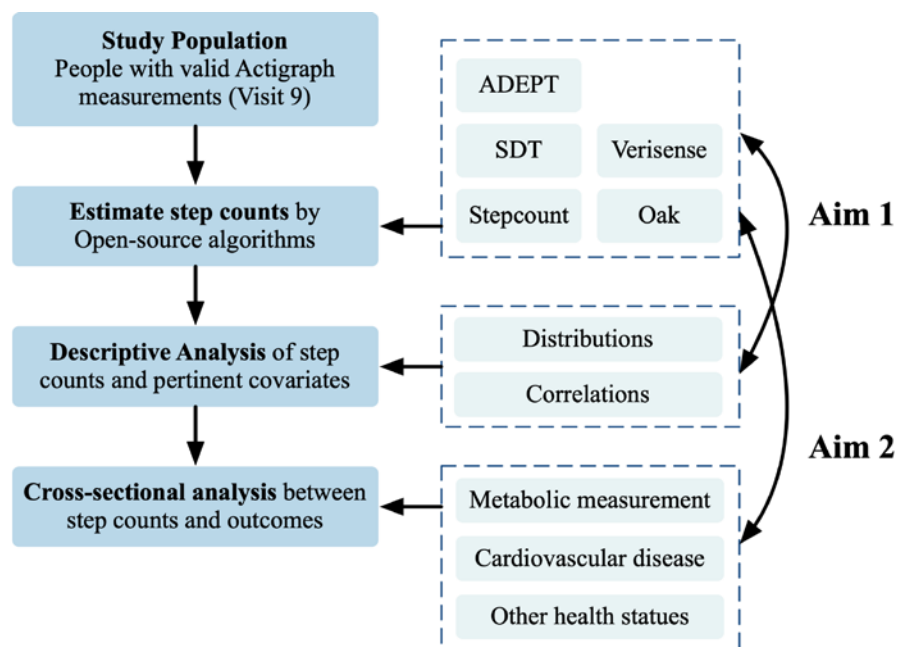
**6. Design and analysis - please address the following aspects:**

- a) inclusion/exclusion
- b) study design
- c) outcome and other variables of interest with specific reference to the time of their collection
- d) summary of data analysis
- e) Any anticipated methodologic limitations or challenges if present
- f) Will the author need Limited data to complete the proposed manuscript? ☐ Yes, Limited data is needed (Provide a brief (2-3 sentences) justification for requesting PHI data) . ☒ No, De-identified data will be sufficient.

\*Please note, Limited dataset access is strict and rarely provided. Limited data includes identifiable information such as dates (birthdays, visit dates, etc.). CMS, Genomic, Geocoded, Proteomics/Somalogic, and other -omic data all fall under the limited data category. De-identified data does not include dates. All dates are date adjusted to "Days since Visit 1".

**Study design:**

Figure 1 represents the proposed analytic workflow. The study focuses on the comparative visualizations of step count generated from five open-source algorithms and the associations between step counts and health outcomes at ARIC Visit 9.



**Figure 1:** The analysis workflow of proposal.

**Study Population:**

We will include all participants who attended Visits 9 in ARIC. We will exclude those whose race was other than Black or White and those who had missing covariates. We will exclude participants who have insufficient Actigraph data, defined as those with more than 10% missing data during the day and fewer than three valid measurement days.

## Exposures:

The ARIC-NCS collects wrist accelerometry data using the triaxial Actigraph GT9X (Actigraph Corp, Pensacola, FL). Participants wore the device on their non-dominant wrist 24 hours per day for 7 days. Wrist placement is validated for physical activity measurement and has become popular due to its associated higher compliance, wear time, and ability to measure 24-hour rhythms and both upper and lower body movements (Schrack et al., 2016). Accelerometry data were collected at 80 Hz, and raw sub-second data were downloaded using ActiLife (version 6.13.4). Wear time was evaluated using the validated Choi algorithm, and days with >10% missing data were considered invalid and excluded from analysis (Choi et al., 2011).

Then, step counts will be generated from five different open-source step count algorithms using raw sub-second level accelerometry data.

1. **ADEPT:** ADEPT is a method for automatic segmentation of individual walking strides from high-resolution accelerometry data (Karas et al., 2021a). ADEPT computes the covariance between a scaled and translated empirical stride pattern and the vector magnitude of the triaxial accelerometry signal. Segments with high enough correlation are flagged as potential steps, and then various filters are applied to remove potential false steps. The empirical patterns are derived from manually segmented strides, enabling ADEPT to be applied to data from various body locations. ADEPT outputs the length and precise start and end times of each detected stride. We applied ADEPT, implemented in the adept R package (Karas et al., 2021a), to the accelerometry data using left wrist templates from the adeptdata package (Karas et al., 2021c) and the optimal parameter combination that maximized accuracy in recognizing walking in the IU data (Karas et al., 2021b).
2. **Oak:** Oak leverages the inherent features of walking based data from accelerometers or smart phones (Strackiewicz et al., 2023a). The approach was validated against 20 publicly available datasets covering various walking styles, sensor locations, and measurement settings, demonstrating high sensitivity across body locations and high specificity for common daily activities. Specifically, the method first preprocesses the raw accelerometer data to a standardized vector magnitude format then resamples the data to 10Hz. It then identifies high-amplitude data segments exceeding a predefined threshold. Continuous wavelet transforms (CWT) is applied to these segments to reveal temporal oscillation frequencies. Segments with important frequency components outside the typical step frequency range or of insufficient duration are excluded. Finally, the method classifies the remaining segments as walking. We will fit Oak using the R package walking, which interfaces with the Python ‘forest’ module (Muschelli, 2024).
3. **Step Detection Threshold (SDT):** This algorithm aims to determine the minimum acceleration value needed to be treated as a “step” using raw acceleration data (Ducharme et al., 2021). To account for device orientation, raw vector magnitude acceleration data is demeaned and then band-pass filtered (0.25-2.5 Hz) using Butterworth filter. Peaks were identified using a peak picking algorithm that detected increasing followed by decreasing acceleration. An acceleration threshold was applied to remove errant peaks, with values iteratively adjusted between 0.0 and 0.2g. Step counts for each threshold were compared to hand-tallied steps, and the threshold with the lowest absolute error across participants was selected. We will implement SDT in the walking R package (Muschelli, 2024).
4. **Verisense:** This method is proposed based on peak detection to address the over-counting problem caused by false walking (Gu et al., 2017; Femiano et al., 2022). The algorithm detects

peaks in the acceleration signal and applies constraints on step periodicity (time between peaks falls within a typical range for walking), similarity (peak amplitudes for the same foot are similar), and continuity (motion state is continuous without frequent significant changes) to eliminate false peaks. We will fit Verisense using walking package (Muschelli, 2024), and the R code is available at [https://github.com/ShimmerEngineering/Verisense-Toolbox/tree/master/Verisense\\_step\\_algorithm](https://github.com/ShimmerEngineering/Verisense-Toolbox/tree/master/Verisense_step_algorithm).

5. **Stepcount:** This method utilized a hybrid approach combining self-supervised deep learning for walking classification and peak detection for step counting. The algorithm includes wear/non-wear detection and step counting (Small et al., 2023). The pre-trained model (an 18-layer ResNet-V2 on the UK Biobank) was used to classify periods of walking and non-walking, and then fine-tuned for supervised gait classification using the OxWalk dataset, with 10-second epochs of accelerometer data labeled as walking or non-walking. Step counting was performed using the "find\_peaks" method from the SciPy Python package on the classified walking time windows. The peak finding algorithm is applied after low-pass filtering of the raw signal. We implement stepcount using a R package (<https://github.com/jhuwit/stepcount>), which wrap up the python package (stepcount: <https://github.com/OxWearables/stepcount>).

### **Outcomes:**

One recent review offered a complete overview of the evidence of step counts measured by devices and multiple health outcomes (Ao et al., 2024). A total of 27 health outcomes in the 20 studies included were related to step counts, with metabolic diseases accounting for 44%, cardiovascular diseases for 31%, all-cause mortality for 17%, physical function points for 2%, and fractures, depression score, frailty, global cognitive function, falls, and fear of falling for 1% each (Ao et al., 2024). Considering the observed health associations and cross-sectional study design limitations, we will consider the following health outcomes.

1. Metabolic outcomes: body mass index (BMI), waist circumference, triglyceride (TG), total cholesterol (TC), high-density lipoprotein cholesterol (HDL-C) and low-density lipoprotein cholesterol (LDL-C), type 2 diabetes
2. Cardiovascular diseases: heart failure, myocardial infarction, stroke, systolic blood pressure (SBP), diastolic blood pressure (DBP), carotid-femoral pulse wave velocity (cfPWV)
3. Physical function and impairment: short physical performance battery (SPPB), gait speed, falls, fear of falling, frailty
4. Cognitive and psychological outcomes: cognitive function, depression (important health outcomes in older adults that are associated with physical activity).

### **Covariates:**

1. Demographic: age, sex, race/ethnicity, site, education
2. Lifestyle: smoking, alcohol consumption
3. Genetic risk factor: APOE
4. Physical function: short physical performance

All these covariates will be based on ARIC Visit 9. If covariates at visit 9 are missing, measurements in prior visits will be considered.

### **Data Analysis:**

For Aim 1, we will employ five distinct open-source step count algorithms with their own default parameters—ADEPT, Oak, Step Detection Threshold, Verisense, and Stepcount—on the same raw sub-second Actigraph physical activity data obtained from the ARIC Study Visit 9 to generate estimated step counts. To prevent extreme value influencing distributions, we plan to exclude individuals who have outlier step counts using adaptive trimmed-mean estimation for each algorithm (Léger and Romano, 1990; “Measuring Robustness,” 2018). Then, we will use histograms to visualize the distribution of step counts and use scatter plots to examine pairwise relationships. We will calculate Pearson correlation coefficients to assess the relationships between step counts, and use Bland-Altman plots to analyze the agreement between algorithms. We will visualize the distribution of steps by covariates such as age, gender, race, education, smoking, drinking, short physical function performance (SPPB) battery and APOE.

For Aim 2, we plan to investigate the associations between step counts and health outcomes mentioned in the recent review (details are shown in outcome part). We will use multivariable linear or logistic regression models to assess the associations of step counts with health outcomes. The choice of model will depend on the type of outcome. Model 1 will adjust for demographic variables, while Model 2 will additionally adjust for lifestyle and physical functions. Then, we evaluate non-linear trends by implementing the restricted cubic spline regression with knots placed at the 5th, 27.5th, 50th, 72.5th, and 95th percentiles. From the analyses we generate, we will compare the directionality and effect sizes of associations across algorithms. This will involve examining whether an increase in step counts is consistently associated with positive or negative health outcomes and the relative magnitude of these associations.

In sensitivity analyses, we will run the analysis including individuals who have extreme values of step counts. Also, to determine whether results differed by physical functioning, the same analyses will be repeated in the sample stratified by SPPB quartiles. To address potential bias from step undercounting in specific subpopulations, stratified analyses will be conducted based on sex and race, with normalization performed within each stratum. Means and standard deviations of each method will be reported for sex and race to examine any differences. Additionally, subjects will be randomly allocated into five groups, and the analyses will be repeated (akin to cross-validation) to validate the robustness of the findings.

#### **Anticipated methodologic limitations or challenges:**

- 1) At visit 9, participants in the study were 70 years or older. This limits the generalizability of our findings to the younger population.
- 2) The wrist-worn nature of accelerometry may lead to an overestimation of steps. However, some open-source algorithms, such as SDT, are designed to address this issue. This challenge may explain the differences in estimated step counts among various algorithms.
- 3) At the time of this analysis, only Actigraph measurements from visit 9 are available, limiting the study design to cross-sectional.
- 4) There are no ground truth step counts in ARIC, limiting the evaluation of generated step counts.

**7.a. Will the data be used for non-ARIC analysis or by a for-profit organization in this manuscript? (Non-ARIC analysis means that the authors are not regarded as ARIC investigators and the "ARIC author" is essentially just a facilitator rather than an integral part of the writing group.)** ☐ Yes ☒ No

**b. If Yes, is the author aware that the current derived consent file ICTDER05 must be used to exclude persons with a value RES\_OTH and/or RES\_DNA = "ARIC only" and/or "Not for Profit" ?** ☐ Yes ☐ No

(The file ICTDER is distributed to ARIC PIs annually, and contains the responses to consent updates related to stored sample use for research.)

**8.a. Will the DNA data be used in this manuscript?** ☐ Yes ☒ No

**8.b. If yes, is the author aware that either DNA data distributed by the Coordinating Center must be used, or the current derived consent file ICTDER05 must be used to exclude those with value RES\_DNA = "No use/storage DNA"?** ☐ Yes ☐ No

**9. The lead author or the "sponsoring" ARIC author of this manuscript proposal has reviewed the list of existing ARIC Study manuscript proposals and has found no overlap between this proposal and previously approved manuscript proposals either published or still in active status.** ARIC Investigators have access to the publications lists under the Study Members Area of the website at: <https://aric.csc.unc.edu/aric9/proposalsearch> [ARIC Website ☐ Publications ☐ Proposal Search]

☒ Yes ☐ No

**10. What are the most related manuscript proposals in ARIC (authors are encouraged to contact lead authors of these proposals for comments on the new proposal or collaboration)?**

U0431: Physical Function and Subsequent Risk of Cardiovascular Events in Older Adults: The Atherosclerosis Risk in Communities Study. (Hu)

1715: Physical activity and incidence of cardiovascular disease in African Americans. (Elizabeth)

3627: Patterns of Leisure-time Physical Activity and Sedentary Behavior with Carotid Artery Atherosclerosis Morphology: the ARIC Carotid-MRI Study. (Keith)

3426: Association of physical activity with the incidence of atrial fibrillation among the elderly in the Atherosclerosis Risk In Communities (ARIC) cohort. (Grace)

**11.a. Is this manuscript proposal associated with any ARIC ancillary studies or does it use current [or ongoing] ancillary study data (this includes ACHIEVE)?** ☐ Yes ☒ No ☐ Skip to question 12

**11.b. If yes to 11.a., is the proposal**

☐ **A. primarily the result of an ancillary study**

☐ **B. primarily based on ARIC data with ancillary data playing a minor role (usually control variables)**



**11.c. If yes to 11.a., list number\* \_\_**

\*ancillary studies are listed by number

[https://aric.csc.unc.edu/aric9/researchers/ancillary\\_studies/approved\\_ancillary\\_studies](https://aric.csc.unc.edu/aric9/researchers/ancillary_studies/approved_ancillary_studies) [ARIC Website □ Ancillary Studies □ Approved Ancillary Studies]

**12a. Manuscript preparation is expected to be completed in one to three years. If a manuscript is not submitted for ARIC review at the end of the 3-years from the date of the approval, the manuscript proposal will expire.**

**12b. The NIH instituted a Public Access Policy in April, 2008** which ensures that the public has access to the published results of NIH funded research. It is **your responsibility to upload manuscripts to PubMed Central** whenever the journal does not and be in compliance with this policy. Four files about the public access policy from <http://publicaccess.nih.gov/> are posted in [https://aric.csc.unc.edu/aric9/publications/policies\\_forms\\_and\\_guidelines](https://aric.csc.unc.edu/aric9/publications/policies_forms_and_guidelines) [ARIC Website □ Publications □ Publication Policies, Forms, and Guidelines]. [http://publicaccess.nih.gov/submit\\_process\\_journals.htm](http://publicaccess.nih.gov/submit_process_journals.htm) shows you which journals automatically upload articles to PubMed central.

**References: \_**

- Ao, Z., He, H., Shi, H., Liu, H., 2024. Step count and multiple health outcomes: An umbrella review. *J. Evid.-Based Med.* n/a, 1–18. <https://doi.org/10.1111/jebm.12596>
- Choi, L., Liu, Z., Matthews, C.E., Buchowski, M.S., 2011. Validation of Accelerometer Wear and Nonwear Time Classification Algorithm. *Med. Sci. Sports Exerc.* 43, 357–364. <https://doi.org/10.1249/MSS.0b013e3181ed61a3>
- del Pozo Cruz, B., Ahmadi, M., Naismith, S.L., Stamatakis, E., 2022a. Association of Daily Step Count and Intensity With Incident Dementia in 78 430 Adults Living in the UK. *JAMA Neurol.* 79, 1059–1063. <https://doi.org/10.1001/jamaneurol.2022.2672>
- del Pozo Cruz, B., Ahmadi, M.N., Lee, I.-M., Stamatakis, E., 2022b. Prospective Associations of Daily Step Counts and Intensity With Cancer and Cardiovascular Disease Incidence and Mortality and All-Cause Mortality. *JAMA Intern. Med.* 182, 1139–1148. <https://doi.org/10.1001/jamainternmed.2022.4000>
- Ducharme, S.W., Lim, J., Busa, M.A., Aguiar, E.J., Moore, C.C., Schuna, J.M., Barreira, T.V., Staudenmayer, J., Chipkin, S.R., Tudor-Locke, C., 2021. A Transparent Method for Step Detection using an Acceleration Threshold. *J. Meas. Phys. Behav.* 4, 311–320. <https://doi.org/10.1123/jmpb.2021-0011>
- Femiano, R., Werner, C., Wilhelm, M., Eser, P., 2022. Validation of open-source step-counting algorithms for wrist-worn tri-axial accelerometers in cardiovascular patients. *Gait Posture* 92, 206–211. <https://doi.org/10.1016/j.gaitpost.2021.11.035>
- Gu, F., Khoshelham, K., Shang, J., Yu, F., Wei, Z., 2017. Robust and Accurate Smartphone-Based Step Counting for Indoor Localization. *IEEE Sens. J.* 17, 3453–3460. <https://doi.org/10.1109/JSEN.2017.2685999>
- Karas, M., Bai, J., Strączkiewicz, M., Harezlak, J., Glynn, N.W., Harris, T., Zipunnikov, V., Crainiceanu, C., Urbanek, J.K., 2019. Accelerometry data in health research: challenges and opportunities. *Stat. Biosci.* 11, 210–237. <https://doi.org/10.1007/s12561-018-9227-2>

Karas, M., Strączkiewicz, M., Fadel, W., Harezlak, J., Crainiceanu, C.M., Urbanek, J.K., 2021a. Adaptive empirical pattern transformation (ADEPT) with application to walking stride segmentation. *Biostatistics* 22, 331–347. <https://doi.org/10.1093/biostatistics/kxz033>

Karas, M., Urbanek, J., Crainiceanu, C., Harezlak, J., Fadel, W., 2021b. Labeled raw accelerometry data captured during walking, stair climbing and driving. <https://doi.org/10.13026/51H0-A262>

Karas, M., Urbanek, J., Harezlak, J., Fadel, W., 2021c. adepdata: Accelerometry Data Sets.

Léger, C., Romano, J.P., 1990. Bootstrap adaptive estimation: The trimmed-mean example. *Can. J. Stat.* 18, 297–314. <https://doi.org/10.2307/3315837>

Maylor, B.D., Edwardson, C.L., Dempsey, P.C., Patterson, M.R., Plekhanova, T., Yates, T., Rowlands, A.V., 2022. Stepping towards More Intuitive Physical Activity Metrics with Wrist-Worn Accelerometry: Validity of an Open-Source Step-Count Algorithm. *Sensors* 22, 9984. <https://doi.org/10.3390/s22249984>

Measuring Robustness, 2018. , in: *Robust Statistics*. John Wiley & Sons, Ltd, pp. 51–86. <https://doi.org/10.1002/9781119214656.ch3>

Muschelli, J., 2024. Segments Accelerometry Data into Walking Using the Python `forest` Module.

Paluch, A.E., Bajpai, S., Bassett, D.R., Carnethon, M.R., Ekelund, U., Evenson, K.R., Galuska, D.A., Jefferis, B.J., Kraus, W.E., Lee, I.-M., Matthews, C.E., Omura, J.D., Patel, A.V., Pieper, C.F., Rees-Punia, E., Dallmeier, D., Klenk, J., Whincup, P.H., Dooley, E.E., Gabriel, K.P., Palta, P., Pompeii, L.A., Chernofsky, A., Larson, M.G., Vasan, R.S., Spartano, N., Ballin, M., Nordström, P., Nordström, A., Anderssen, S.A., Hansen, B.H., Cochrane, J.A., Dwyer, T., Wang, J., Ferrucci, L., Liu, F., Schrack, J., Urbanek, J., Saint-Maurice, P.F., Yamamoto, N., Yoshitake, Y., Newton, R.L., Yang, S., Shiroma, E.J., Fulton, J.E., 2022. Daily steps and all-cause mortality: a meta-analysis of 15 international cohorts. *Lancet Public Health* 7, e219–e228. [https://doi.org/10.1016/S2468-2667\(21\)00302-9](https://doi.org/10.1016/S2468-2667(21)00302-9)

Schrack, J.A., Cooper, R., Koster, A., Shiroma, E.J., Murabito, J.M., Rejeski, W.J., Ferrucci, L., Harris, T.B., 2016. Assessing Daily Physical Activity in Older Adults: Unraveling the Complexity of Monitors, Measures, and Methods. *J. Gerontol. Ser. A* 71, 1039–1048. <https://doi.org/10.1093/gerona/glw026>

Sheng, M., Yang, J., Bao, M., Chen, T., Cai, R., Zhang, N., Chen, H., Liu, M., Wu, X., Zhang, B., Liu, Y., Chao, J., 2021. The relationships between step count and all-cause mortality and cardiovascular events: A dose–response meta-analysis. *J. Sport Health Sci.* 10, 620–628. <https://doi.org/10.1016/j.jshs.2021.09.004>

Small, S.R., Chan, S., Walmsley, R., Fritsch, L. von, Acquah, A., Mertes, G., Feakins, B.G., Creagh, A., Strange, A., Matthews, C.E., Clifton, D.A., Price, A.J., Khalid, S., Bennett, D., Doherty, A., 2023. Development and Validation of a Machine Learning Wrist-worn Step Detection Algorithm with Deployment in the UK Biobank. <https://doi.org/10.1101/2023.02.20.23285750>

Straczekiewicz, M., Huang, E.J., Onnela, J.-P., 2023a. A “one-size-fits-most” walking recognition method for smartphones, smartwatches, and wearable accelerometers. *Npj Digit. Med.* 6, 1–16. <https://doi.org/10.1038/s41746-022-00745-z>

Straczekiewicz, M., Keating, N.L., Thompson, E., Matulonis, U.A., Campos, S.M., Wright, A.A., Onnela, J.-P., 2023b. Open-Source, Step-Counting Algorithm for Smartphone Data Collected in Clinical and Nonclinical Settings: Algorithm Development and Validation Study. *JMIR Cancer* 9, e47646. <https://doi.org/10.2196/47646>

**To view publications materials, click "Log in" at the top right of the ARIC website. Click "Forgot Password" if you are experiencing issues with logging in.**