Published in final edited form as:
J Aging Phys Act. ; 29(3): 529–535. doi:10.1123/japa.2020-0059.

Comparing Methods to Identify Wear-Time Intervals for Physical Activity With the Fitbit Charge 2

Sophie E. Claudel,
Social Determinants of Obesity and Cardiovascular Risk Laboratory, Cardiovascular Branch, Division of Intramural Research, National Heart, Lung, and Blood Institute, National Institutes of Health, Bethesda, MD, USA.

Kosuke Tamura,
Social Determinants of Obesity and Cardiovascular Risk Laboratory, Cardiovascular Branch, Division of Intramural Research, National Heart, Lung, and Blood Institute, National Institutes of Health, Bethesda, MD, USA.

James Troendle,
Social Determinants of Obesity and Cardiovascular Risk Laboratory, Cardiovascular Branch, Division of Intramural Research, National Heart, Lung, and Blood Institute, National Institutes of Health, Bethesda, MD, USA.

Marcus R. Andrews,
Social Determinants of Obesity and Cardiovascular Risk Laboratory, Cardiovascular Branch, Division of Intramural Research, National Heart, Lung, and Blood Institute, National Institutes of Health, Bethesda, MD, USA.

Joniqua N. Ceasar,
Social Determinants of Obesity and Cardiovascular Risk Laboratory, Cardiovascular Branch, Division of Intramural Research, National Heart, Lung, and Blood Institute, National Institutes of Health, Bethesda, MD, USA.

Valerie M. Mitchell,
Social Determinants of Obesity and Cardiovascular Risk Laboratory, Cardiovascular Branch, Division of Intramural Research, National Heart, Lung, and Blood Institute, National Institutes of Health, Bethesda, MD, USA.

Nithya Vijayakumar,
Social Determinants of Obesity and Cardiovascular Risk Laboratory, Cardiovascular Branch, Division of Intramural Research, National Heart, Lung, and Blood Institute, National Institutes of Health, Bethesda, MD, USA.

Tiffany M. Powell-Wiley
Social Determinants of Obesity and Cardiovascular Risk Laboratory, Cardiovascular Branch, Division of Intramural Research, National Heart, Lung, and Blood Institute, National Institutes of Health, Bethesda, MD, USA.

corresponding author. Powell-Wiley (tiffany.powell-wiley@nih.gov).

Intramural Research Program, National Institute on Minority Health and Health Disparities, National Institutes of Health, Bethesda, MD, USA.

Abstract

There is no established method for processing data from commercially available physical activity trackers. This study aims to develop a standardized approach to defining valid wear time for use in future interventions and analyses. Sixteen African American women (mean age = 62.1 years and mean body mass index = 35.5 kg/m^2) wore the Fitbit Charge 2 for 20 days. Method 1 defined a valid day as ≥10-hr wear time with heart rate data. Method 2 removed minutes without heart rate data, minutes with heart rate ≤ mean − 2 *SD*s below mean and ≤2 steps, and nighttime. Linear regression modeled steps per day per week change. Using Method 1 (n = 292 person-days), participants had 20.5 (*SD* = 4.3) hr wear time per day compared with 16.3 (*SD* = 2.2) hr using Method 2 (n = 282) (p < .0001). With Method 1, participants took 7,436 (*SD* = 3,543) steps per day compared with 7,298 (*SD* = 3,501) steps per day with Method 2 (p = .64). The proposed algorithm represents a novel approach to standardizing data generated by physical activity trackers. Future studies are needed to improve the accuracy of physical activity data sets.

Keywords

mobile health; physical activity intervention; steps

Insufficient physical activity (PA) is a modifiable risk factor for preventing cardiovascular disease (CVD) and obesity (Lim et al., 2012; Wahid et al., 2016). In the United States, African Americans have a higher burden of CVD and are much less likely than other racial/ethnic groups to meet PA guidelines (Clarke, Norris, & Schiller, 2017). Despite recent national survey data showing that 52% of Americans meet PA guidelines of 150 min of moderate to vigorous PA per week (i.e., 500 metabolic equivalent minutes per week) (Clarke et al., 2017), prior objective assessment of PA via accelerometers indicated that <10% of U.S. adults meet the guidelines (Troiano et al., 2008; Tucker, Welk, & Beyler, 2011). The discrepancy between self-reported and objectively measured PA through accelerometers highlights the need for improved adherence to PA guidelines (Troiano et al., 2008) and new strategies for increasing PA, such as mobile health interventions using wearable devices (i.e., commercially available PA trackers).

The nearly ubiquitous use of smartphones and other mobile technology has prompted research on the potential of these tools to deliver low-cost, effective interventions in the community setting. Prior work also demonstrated strong adherence to using commercial wearable devices (such as Fitbit, Fitbit, Inc. San Francisco, CA; Microsoft Band, Microsoft, Inc. Redmond, WA) across a variety of populations (Hartman, Nelson, & Weiner, 2018; Hermsen, Moons, Kerkhof, Wiekens, & De Groot, 2017; Phillips, Cadmus-Bertram, Rosenberg, Buman, & Lynch, 2018). The combination of activity tracking and app-based motivational messaging holds the potential to have a more powerful effect on behavior; however, more rigorous methodologic approaches to data processing are needed.

Whereas the literature on using accelerometry includes specific analytic methods to classify wear/nonwear time and intensity of PA (Choi, Capen Ward, Schnelle, & Buchowski, 2012; Choi, Liu, Matthews, & Buchowski, 2011; Koster et al., 2016; Lee et al., 2018), the same does not hold true for commercially available PA trackers. In a review of current literature assessing step counts in relation to health outcomes, Kraus et al. (2019) highlighted the need for further research on the methods used to analyze and present data from step-based PA studies. Furthermore, the authors called for increased attention to the relationship between individual characteristics, step counts, and health (Kraus et al., 2019). Similar concerns have been raised by other research groups (Alharbi, Straiton, Smith, Neubeck, & Gallagher, 2019; Lobelo et al., 2016; Welk et al., 2019) and were also highlighted by the 2018 U.S. Physical Activity Guideline Committee (U.S. Department of Health and Human Services, 2018). The committee challenged researchers to more fully explore analytic methods relating to the use of commercially available PA trackers as a prevalent data source (U.S. Department of Health and Human Services, 2018).

Fitbit and other similar devices are increasingly used in public health research as an alternative to traditional accelerometry as they demonstrate reliable step counts in adults without mobility limitations (Evenson, Goto, & Furberg, 2015; Feehan et al., 2018). Step counts have been widely used as a PA outcome and PA surveillance (Bassett, Toth, LaMunion, & Crouter, 2017; Feehan et al., 2018) despite little discussion of how to process raw data from commercially available devices for research purposes. As a result, the literature is replete with diverging methods for data management (Alharbi et al., 2019). For example, in a validation study of the heart rate (HR) feature of the Fitbit Charge HR device, Gorny, Liew, Tan, and Müller-Riemenschneider (2017) regarded every nonzero HR minute as a valid minute (Gorny et al., 2017). Sprint, Cook, and Schmitter-Edgecombe (2018) interpolated missing HR data for periods of time <30 min but did not comment on how step data were handled during these periods (Sprint et al., 2018). Many other studies did not report how raw Fitbit data were processed before analysis of step counts (Chen, Kuo, Pellegrini, & Hsu, 2016; Tedesco et al., 2019; Weatherall, Paprocki, Meyer, Kudel, & Witt, 2018). This raises concerns regarding the reliability of estimates and ability of other researchers to replicate results. Furthermore, given the methodologic diversity of existing studies, it is difficult to fully appreciate the potential for commercially available PA trackers to increase PA in the community setting. This is especially relevant to older African American women, who may be at high risk for CVD and are understudied in the literature.

Within the context of an ongoing community collaboration to improve cardiovascular health among African American women in the Washington, DC metropolitan area, the purpose of this study was twofold: (a) compare two methods for processing large-volume minute-by-minute data sets to classify wear time and (b) examine how individual characteristics are associated with African American women's PA to inform future interventions to increase PA.

Methods

Participants

The participants in this study were a convenience sample of African American women recruited in the context of an ongoing community-based participatory research project to improve cardiovascular health in the Washington, DC metropolitan area (Yingling et al., 2016, 2017). Briefly, the overall study pilot tested the potential application of mobile health technology for a technology-enabled, community-based PA intervention in a resource-limited community (Ceasar et al., 2019). Participants were given a novel mobile phone app and a commercially available PA tracker, the Fitbit Charge 2 (Fitbit, San Francisco, CA). Fitbit data from the 20-day pilot testing period are analyzed here. The current study involves women at risk of CVD who met the following inclusion criteria: (a) residence in lower income areas including Wards 5, 7, and 8 of Washington, DC or Prince George's county, MD; (b) age 18–85 years at time of enrollment; (c) smartphone ownership; (d) English proficiency; (e) physical ability to engage in study activities; and (f) body mass index (BMI) ≥ 25 kg/m^2 by self-reported height and weight. Twenty women were recruited, and 16 ultimately enrolled. The study was approved by the institutional review board of the National Heart, Lung, and Blood Institute (National Institutes of Health; NCT01927783), and all participants provided written informed consent. The funding sources had no role in the design, implementation, or publication of this research.

Data Collection

Participants met with the study team at a partnering church in Washington, DC for Fitbit deployment. The study team linked the Fitbits to each participant's Fitbit app using anonymous user accounts with a study-generated e-mail address and generic demographic and health information to protect participants' privacy. The participants were instructed to wear the Fitbit device at all times for the 20-day pilot study period with the exception of water-based activities (showering, bathing, swimming, etc.). They were individually instructed in the use, charging, and syncing of the Fitbit during the deployment session.

Each participants' Fitbit account was linked to a project analytics system using Fitabase (Small Steps Labs, LLC, San Diego, CA; available at www.fitabase.com). Therefore, activity could be monitored remotely by the study team. The Fitbit recorded the steps and HR achieved by each participant at the minute-by-minute level for a total of 320 person-days of raw data among 16 participants. Following study completion, all Fitbit data were downloaded from the Fitabase website for analysis. Thereafter, participants could keep the devices and were assisted in creating a personal account to use with the Fitbit app.

Determining Wear and Nonwear Time

Two separate methods for determining wear time were developed for comparison (Table 1). First, the standard method applied for defining valid days was designed for consistency with prior accelerometry analyses (Brewer, Swanson, & Ortiz, 2017; Cadmus-Bertram, Marcus, Patterson, Parker, & Morey, 2015; Rich et al., 2013; Tully, McBride, Heron, & Hunter, 2014). This method restricted the data set to days with at least 10 hr of HR monitored minutes (i.e., wear time). Second, a new method was developed to further restrict the data set to valid minutes. The restrictions were sequentially imposed as follows (please see Supplementary Material [available online], e.g., SAS code [SAS Institute, Cary, NC]):

a. Removing minutes with missing HR data. This represents an approximation of nonwear time, as the Fitbit device monitors HR in 1-s epochs. If a participant is wearing the device appropriately, the HR is captured and averaged over 1-min intervals. Prior studies have used the HR feature to assess wear time (Gorny et al., 2017), and appropriate wrist placement usually produces a detectable HR according to Fitbit (Fitbit, 2019). Therefore, minutes without HR data can be assumed as nonwear time (Collins, Yang, Trentadue, Gong, & Losina, 2019).

b. Removing invalid minutes, defined as minutes with both an HR less than 2 *SD*s below the population mean HR (the lower limit of the expected physiologic range; Avram et al., 2019) and two or fewer steps. These steps were considered invalid, as the Fitbit device is known to occasionally register certain hand and arm movements as steps (Chen et al., 2016). If these motions occur while the HR is below the expected physiologic range, it further suggests that the device is not accurately capturing true PA (i.e., steps; Avram et al., 2019). Spurious HR detection has also been noted when the device is placed in a backpack or purse (Kondama Reddy et al., 2018). The combination of low HR and low step rate was employed to ensure that sedentary time (a period of normal HR and few steps) would not be inadvertently removed from analysis.

c. Removing sleeping hours (11 p.m.–5 a.m.). Sleeping hours were excluded to ensure that sufficient hours with the potential for intentional PA were included when restricting valid days to those with >10 hr of wear time. Fitbit devices are known to occasionally register false-positive steps during activities of daily living (Chen et al., 2016; O'Connell, Ólaighin, & Quinlan, 2017), which may also occur with movement during sleep. Although the Fitbit device reports sleeping time, the validity of these data are unclear (Evenson et al., 2015). Therefore, nighttime hours were selected as presumed sleeping time. During focus groups with this participant population, the women reported early awakening times, which informed the research team to select 11 p.m.–5 a.m. as presumed sleep time. Sensitivity analysis was conducted using sleeping parameters of 12 a.m.–6 a.m. and assessing for any differences in outcomes compared with the sleeping parameters of 11 p.m.–5 a.m.

d. Removing invalid days (≤10 hr of HR monitored minutes). As noted above, this is standard practice in the accelerometry literature (Brewer et al., 2017; Cadmus-Bertram et al., 2015; Tully et al., 2014) and has been used with Fitbit data as well (Collins et al., 2019). Therefore, it was imposed in the new method as well.

PA Outcome

The PA outcome was defined as mean steps per day per week (n = 47 person-week level, for 16 individuals). Intensity of PA was not assessed due to concerns regarding the accuracy of energy expenditure as calculated by the proprietary Fitbit algorithm (Feehan et al., 2018;

Nuss et al., 2019; Shcherbina et al., 2017). For example, in a systematic review of 18 studies, Feehan et al. (2018) concluded that Fitbit devices underestimated activity intensity compared with calorimetry-based energy expenditure. Across multiple activity comparisons, the Fitbit was rarely within an error margin of ±3% and was subject to additional variability depending on the type of ambulation.(Feehan et al., 2018). Furthermore, the proprietary algorithm may change in the app background at any time without notice to users and, therefore, result in inconsistent calculations over the study period.

Covariates

Self-reported individual-level covariates included age (in years), BMI (in kilograms per meter squared), marital status (single/divorced/widowed/separated vs. married/living with partner), employment status (employed vs. unemployed), and educational attainment (below college education vs. college and above).

Statistical Analysis

Means and *SD*s were used to describe the individual-level characteristics and overall PA data. For each method of data restriction, mean steps per day and hours of wear time per day were compared using independent sample *t* tests.

Since the data are a multilevel structure (i.e., person-days are nested within weeks), an intraclass correlation coefficient was calculated for an intercept only model at the person-day level. The intraclass correlation coefficient was 34.3%, indicating higher between-person variability compared with within-person variability. Therefore, mixed-effects modeling was used to examine the associations of individual characteristics (age, BMI, income, marital status, education, or employment) with the PA outcome (average steps per person per day per week). A person-level random effect was included to account for within-person correlation. Analyses were conducted with the data set of standard restrictions and new restrictions, including the sensitivity analysis. Time-series data of average step counts per day at the participant level were graphed using LOESS curves. All statistical analyses were conducted in SAS (version 9.4; SAS Institute).

Results

Demographic Characteristics

The study sample consisted of 16 African American women with a mean age of 62.1 years (*SD* = 6.6 years; range = 52–74 years; Table 2). Mean BMI was 35.5 kg/m^2 (range = 25.6–54.6 kg/m^2) with 75.0% of women classified as obese (BMI ≥ 30 kg/m^2). Half (50.0%) of the sample was retired, with the remaining being employed full time (26.7%), part time (13.3%), or unemployed (12.5%). Educational attainment was varied, ranging from below high school to professional degree. Income information was only available for eight participants (50%), spanning a range of $20,000–$99,999 (by $10,000 increment categories). The study demonstrated high adherence to wearing the Fitbit device for the study period, which generated 20 days of raw minute-by-minute level step counts (after restrictions: using the standard method, 75% wore the Fitbit for ≥19 days; using the new method, 63% wore the Fitbit for ≥19 days).

Comparison of Methods to Determine Wear Time for Counting Steps

The two methods (Table 1) were compared based on number of valid days, average hours of wear time per day, and average number of steps per day (Table 3). Using the new method with sleeping hours defined as 11 p.m.–5 a.m., there were 281 person-days available for analysis, or 88% of possible days (vs. 292 person-days, or 92% using the standard method). Participants wore the devices for an average of 16.3 ($SD = 2.2$) hr per day, which was statistically lower than the standard method (20.5 [$SD = 4.3$] hr per day, $p < .0001$). Using the new method, participants took an average of 7,298 ($SD = 3{,}501$) steps per day, which was not statistically different from the standard method (7,436 [$SD = 3{,}543$] steps per day, $p = .64$). Sensitivity analysis using sleeping hours 12 a.m.–6 a.m. yielded similar results, with 282 available person-days, 16.3 ($SD = 2.1$) hr of wear time ($p < .0001$ compared with the wear time calculated by the standard method), and 7,345 ($SD = 3{,}475$) steps per day ($p = .76$ compared with the steps per day based on the standard method). A graphical representation of average steps per day over the study period for each method is included in Supplementary Figure 1 (available online).

Mixed-Effects Modeling of Average Steps per Person per Day per Week

Fitbit data collected during the study period, which included simultaneous PA app use, were analyzed using mixed-effects modeling. Using step data generated by the standard method, mixed-effects modeling of average steps per day at the week level demonstrated no significant increase in steps across each study week (Table 4). Using data generated from the new method, mixed-effects modeling of average steps per day at the week level demonstrated a marginally significant increase in steps at Week 2. Individual characteristics were not significant in any of the models.

Discussion

Among this population of middle-aged African American women in the Washington, DC metropolitan area, comparing two methods for processing raw data from Fitbit devices yielded significantly different wear times. The novelty of this study is in the deliberate construction of a new analytic approach to use data from a consumer wearable device. Well-established algorithms have been developed for analysis of research-grade accelerometers, and research teams have continued to assess the development of wear-time criteria (Choi et al., 2012, 2011; Koster et al., 2016). For example, Choi et al. (2011) demonstrated the application of a new wear-time algorithm to decrease ActiGraph wear-time misclassification (Choi et al., 2011). The rise of consumer wearable devices and their use in research has been focused solely on validation studies rather than generating guidelines for analysis of wear time and other features (James et al., 2016). The problems of large volume, commercial, sensor-based data will only increase with the popularity of these devices and call for new methodologic approaches to these data sets (James et al., 2016). The findings presented in this study demonstrate that greater emphasis on analytic techniques is needed in research using commercially available PA data in community-dwelling adults.

Using a more comprehensive approach to determining valid minutes from a raw data set of Fitbit data, this study shows significant differences in hours of wear time compared with standard data processing methods. In contrast, the average number of steps per day across the study population was not statistically different after imposing the two different wear-time methods. As this study was limited by a small population over a short study period, differences in step count may be seen following a longer period of observation or with a greater number of participants. A recent paper compared the Fitbit Charge 2 with a research-grade accelerometer and included a short description of Fitbit data processing in the methods (Collins et al., 2019). The authors used two approaches, an HR-based algorithm to define wear time and a step-based algorithm to define nonwear time, which was then subtracted from a 24-hr period (Collins et al., 2019). The HR-based algorithm included all minutes with nonzero HR and nonzero steps as wear time (Collins et al., 2019). The step-based algorithm defined nonwear time as bouts with zero steps lasting >60 consecutive minutes (Collins et al., 2019). Although this does demonstrate Fitbit data processing using a systematic approach, it does not fully account for the larger underlying issues with raw Fitbit data, such as spurious steps and inaccurate HR detection, or shorter periods of nonwear time. Our algorithms more comprehensively address raw data quality and activity patterns using a combination of both HR and step parameters. Unfortunately, most studies using Fitbit data do not report methods determining wear time; therefore, we cannot directly compare our methods with much of the existing literature (Chen et al., 2016; Tedesco et al., 2019; Weatherall et al., 2018). The lack of detailed methodology in existing studies leads to the assumption that all step data are used for analysis, regardless of potential inaccurate counts during nonwear periods. Without attention to data processing, studies employing Fitbits and other commercial activity monitors may be vulnerable to skew.

This study is novel in its approach to defining wear time for a commercial wearable device, especially in an understudied population of urban African American women. Participants in this study took an average of approximately 7,300 steps per day over the entire study period. This is similar to our prior findings in the community wherein the baseline step count was 7,050 steps per day among African American women with overweight/obesity (Thomas et al., 2017). Overall, however, this is higher than previously reported. For example, a large pedometer-based study of U.S. adults found that African Americans took an average of 3,974 steps per day (vs. 5,086 steps per day among European Americans), whereas women of all races took an average 4,912 steps per day (vs. 5,340 steps per day among men of all races; Bassett, Wyatt, Thompson, Peters, & Hill, 2010). Additionally, individuals with obesity took approximately 1,500 steps fewer than those with normal BMI (Bassett et al., 2017). Although a recent, small study of young, overweight urban women (62% African American) reported an overall average of 14,143 steps per day, the unexpectedly high step count is likely attributable to the young age (mean = 26.5 years), highly walkable urban environment (Boston, MA), racial heterogeneity (38% European American), lower BMI (mean = 31.5 kg/m^2), and use of a waist-worn accelerometer rather than Fitbit device (Camhi et al., 2019). Despite prior work demonstrating significantly different step counts based on individual characteristics such as age and BMI (Sisson, Camhi, Tudor-Locke, Johnson, & Katzmarzyk, 2012), our mixed-effects modeling at the week level did not identify any significant individual characteristics in the relationship with steps. Sustained use of the study mobile app beyond the short pilot period may result in more appreciable increases in PA over time.

Strengths and Limitations

This study has several strengths. First, the study was conducted in a resource-limited community among racial/ethnic minority participants. The study also demonstrated high adherence to wearing the Fitbit device for the study period, with most participants wearing the device daily. After data processing, valid person-days available for analysis represented 92% (standard method) and 88% (new method) of the raw data. Use of the Fitbit device also allowed for objective measurement of steps and, therefore, reduced the social desirability bias introduced by self-report surveys (Adams et al., 2005).

However, there were also several limitations to be noted. The sample size was small, exclusively women, geographically constrained to a single metropolitan area, and composed of a single race/ethnicity. Therefore, our findings may not be valid in other populations of African American women, such as younger women or those in rural areas. As there was no run-in period, it was not possible to capture baseline PA. BMI measures relied on self-reported height and weight, which introduces potential social desirability bias as previously shown among women using national data (Burke & Carman, 2017). As sleeping hours were determined by the research team, they may not reflect true sleeping hours of each participant. However, our sensitivity analysis showed no difference when using a shifted definition of sleeping time. Future studies should consider using sleep diaries. An inherent limitation of this study is the lack of a gold standard method for data processing to use as a comparison. Therefore, although it is possible that our proposed methods may increase accuracy through the application of more rigorous data processing, there is also a risk of unnecessarily restricting data sets and, thereby, reducing data quality. Further studies are needed to continue augmenting our understanding of the data produced by commercially available PA trackers and the influence of differential data processing methods.

Conclusions

There is a growing need for standardized methods to classify wear time in commercially available PA trackers. Commercially available devices, unlike traditional accelerometers, are not able to be analyzed through “counts”; therefore, additional metrics are needed to facilitate appropriate classification of registered movements (such as accurate assessment of wear time). In our analysis of two methods to assess Fitbit wear time within the context of a PA pilot intervention, we found that our new method of data processing yielded significantly different total wear time than the conventional method. Although step counts and mixed-effects modeling were not different between the two approaches, the proposed method may improve accuracy of future data sets generated by larger studies. Additional studies are needed to understand the impact of new methods of data processing, as there is currently no gold standard for comparison.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by the Division of Intramural Research of the National Heart, Lung, and Blood Institute and the Intramural Research Program of the National Institute on Minority Health and Health Disparities (National Institutes of Health). This research was also made possible by the National Institutes of Health (NIH) Medical Research Scholars Program, a public–private partnership supported jointly by the NIH and generous contributions to the Foundation for the NIH from the Doris Duke Charitable Foundation, Genentech, the American Association for Dental Research, the Colgate-Palmolive Company, Elsevier, alumni of the student research programs, and other supporters via contributions to the Foundation for the National Institutes of Health. K.T. is supported by the 2019 Lenfant Biomedical Fellowship Award from the NHLBI. 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. This study was registered at www.ClinicalTrials.gov (NCT01927783). The authors declare they have no competing interests. T.M.P.-W. and K.T. conceived of the study. S.E.C., K.T., T.M.P.-W., J.N.C., V.M.M., and M.R.A. assisted with data collection. S.E.C. and J.T. performed statistical analysis. K.T., S.E.C., and T.M.P.-W. drafted the initial manuscript. All authors edited and approved of the final written manuscript.

References

Adams SA, Matthews CE, Ebbeling CB, Moore CG, Cunningham JE, Fulton J, & Hebert JR (2005). The effect of social desirability and social approval on self-reports of physical activity. American Journal of Epidemiology, 161(4), 389–398. doi:10.1093/aje/kwi054 [PubMed: 15692083]

Alharbi M, Straiton N, Smith S, Neubeck L, & Gallagher R (2019). Data management and wearables in older adults: A systematic review. Maturitas, 124, 100–110. doi:10.1016/j.maturitas.2019.03.012 [PubMed: 30910279]

Avram R, Tison GH, Aschbacher K, Kuhar P, Vittinghoff E, Butzner M, … Olgin J (2019). Real-world heart rate norms in the Health eHeart study. NPJ Digital Medicine, 2(1), 58. doi:10.1038/s41746-019-0134-9 [PubMed: 31304404]

Bassett DR, Toth LP, LaMunion SR, & Crouter SE (2017). Step counting: A review of measurement considerations and health-related applications. Sports Medicine, 47(7), 1303–1315. doi:10.1007/s40279-016-0663-1 [PubMed: 28005190]

Bassett DR, Wyatt HR, Thompson H, Peters JC, & Hill JO (2010). Pedometer-measured physical activity and health behaviors in U.S. adults. Medicine & Science in Sports & Exercise, 42(10), 1819–1825. doi:10.1249/MSS.0b013e3181dc2e54 [PubMed: 20305579]

Brewer W, Swanson BT, & Ortiz A (2017). Validity of Fitbit's active minutes as compared with a research-grade accelerometer and self-reported measures. BMJ Open Sport & Exercise Medicine, 3(1), e000254. doi:10.1136/bmjsem-2017-000254

Burke MA, & Carman KG (2017). You can be too thin (but not too tall): Social desirability bias in self-reports of weight and height. Economics and Human Biology, 27, 198–222. doi:10.1016/j.ehb.2017.06.002 [PubMed: 28768226]

Cadmus-Bertram L, Marcus BH, Patterson RE, Parker BA, & Morey BL (2015). Use of the Fitbit to measure adherence to a physical activity intervention among overweight or obese, postmenopausal women: Self-monitoring trajectory during 16 weeks. JMIR mHealth and uHealth, 3(4), e96. doi:10.2196/mhealth.4229 [PubMed: 26586418]

Camhi SM, Troped PJ, Garvey M, Hayman LL, Must A, Lichtenstein AH, & Crouter SE (2019). Associations between Walk Score and objective measures of physical activity in urban overweight and obese women. PLoS One, 14(3), e0214092. doi:10.1371/journal.pone.0214092 [PubMed: 30921337]

Ceasar JN, Claudel SE, Andrews MR, Tamura K, Mitchell V, Brooks A, … Powell-Wiley TM (2019). Community member engagement in the development of a mobile health technology-enabled physical activity and cardiovascular health intervention (step it up): Pilot Focus Group Study. JMIR Formative Research, 3(1), e10944. [PubMed: 30684422]

Chen MD, Kuo CC, Pellegrini CA, & Hsu MJ (2016). Accuracy of wristband activity monitors during ambulation and activities. Medicine & Science in Sports & Exercise, 48(10), 1942–1949. doi:10.1249/MSS.0000000000000984 [PubMed: 27183123]

Choi L, Capen Ward S, Schnelle JF, & Buchowski MS (2012). Assessment of wear/nonwear time classification algorithms for triaxial accelerometer. Medicine & Science in Sports & Exercise, 44(10), 2009–2016. doi:10.1249/MSS.0b013e318258cb36 [PubMed: 22525772]

Choi L, Liu Z, Matthews CE, & Buchowski MS (2011). Validation of accelerometer wear and nonwear time classification algorithm. Medicine & Science in Sports & Exercise, 43(2), 357–364. doi:10.1249/MSS.0b013e3181ed61a3 [PubMed: 20581716]

Clarke TC, Norris T, & Schiller JS (2017). Early release of selected estimates based on data from the 2016 national health interview survey. National Center for Health Statistics, 1–120. Retrieved from https://www.cdc.gov/nchs/data/nhis/earlyrelease/earlyrelease201705.pdf

Collins JE, Yang HY, Trentadue TP, Gong Y, & Losina E (2019). Validation of the Fitbit charge 2 compared to the ActiGraph GT3X+ in older adults with knee osteoarthritis in free-living conditions. PLoS One, 14(1), e0211231. doi:10.1371/journal.pone.0211231 [PubMed: 30699159]

Evenson KR, Goto MM, & Furberg RD (2015). Systematic review of the validity and reliability of consumer-wearable activity trackers. International Journal of Behavioral Nutrition and Physical Activity, 12(1), 159. doi:10.1186/s12966-015-0314-1

Feehan LM, Geldman J, Sayre EC, Park C, Ezzat AM, Young Yoo J, … Li LC (2018). Accuracy of Fitbit devices: Systematic review and narrative syntheses of quantitative data. JMIR mHealth and uHealth, 6(8), e10527. doi:10.2196/10527. [PubMed: 30093371]

Fitbit. (2019). Help article: How do I track my heart rate with my Fitbit device? Retrieved from https://help.fitbit.com/articles/en_US/Help_article/1565

Gorny AW, Liew SJ, Tan CS, & Müller-Riemenschneider F (2017). Fitbit charge HR wireless heart rate monitor: Validation study conducted under free-living conditions. JMIR mHealth and uHealth, 5(10), e157. doi:10.2196/mhealth.8233 [PubMed: 29055881]

Hartman SJ, Nelson SH, & Weiner LS (2018). Patterns of Fitbit use and activity levels throughout a physical activity intervention: Exploratory analysis from a randomized controlled trial. JMIR mHealth and uHealth, 6(2), e29. doi:10.2196/mhealth.8503 [PubMed: 29402761]

Hermsen S, Moons J, Kerkhof P, Wiekens C, & De Groot M (2017). Determinants for sustained use of an activity tracker: Observational study. JMIR mHealth and uHealth, 5(10), e164. doi:10.2196/mhealth.7311 [PubMed: 29084709]

James P, Jankowska M, Marx C, Hart JE, Berrigan D, Kerr J, … Laden F (2016). “Spatial Energetics”: Integrating data from GPS, accelerometry, and GIS to address obesity and inactivity. American Journal of Preventive Medicine, 51(5), 792–800. [PubMed: 27528538]

Kondama Reddy R, Pooni R, Zaharieva DP, Senf B, El Youssef J, Dassau E, … Jacobs PG (2018). Accuracy of wrist-worn activity monitors during common daily physical activities and types of structured exercise: Evaluation study. JMIR mHealth and uHealth, 6(12), e10338. [PubMed: 30530451]

Koster A, Shiroma EJ, Caserotti P, Matthews CE, Chen KY, Glynn NW, & Harris TB (2016). Comparison of sedentary estimates between activPAL and hip- and wrist-worn ActiGraph. Medicine & Science in Sports & Exercise, 48(8), 1514–1522. doi:10.1249/MSS.0000000000000924 [PubMed: 27031744]

Kraus WE, Janz KF, Powell KE, Campbell WW, Jakicic JM, Troiano RP, … Piercy KL (2019). Daily step counts for measuring physical activity exposure and its relation to health. Medicine & Science in Sports & Exercise, 51(6), 1206–1212. doi:10.1249/MSS.0000000000001932 [PubMed: 31095077]

Lee I-M, Shiroma EJ, Evenson KR, Kamada M, LaCroix AZ, & Buring JE (2018). Accelerometer-measured physical activity and sedentary behavior in relation to all-cause mortality: The Women’s Health Study. Circulation, 137(2), 203–205. doi:10.1161/CIRCULATIONAHA.117.031300 [PubMed: 29109088]

Lim SS, Vos T, Flaxman AD, Danaei G, Shibuya K, Adair-Rohani H, … Ezzati M (2012). A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990–2010: A systematic analysis for the Global Burden of Disease Study 2010. The Lancet, 380(9859), 2224–2260. doi:10.1016/S0140-6736(12)61766-8

Lobelo F, Kelli HM, Chernetsky Tejedor S, Pratt M, McConnell MV, Martin SS, & Welk GJ (2016). The Wild Wild West: A framework to integrate mHealth software applications to support physical activity assessment, counseling and interventions for cardiovascular disease risk reduction. Progress in Cardiovascular Diseases, 58(6), 584–594. [PubMed: 26923067]

Nuss KJ, Thomson EA, Courtney JB, Comstock A, Reinwald S, Blake S, … Li K (2019). Assessment of accuracy of overall energy expenditure measurements for the Fitbit Charge HR 2 and Apple Watch. American Journal of Health Behavior, 43(3), 498–505. doi:10.5993/AJHB.43.3.5 [PubMed: 31046881]

O'Connell S, Ólaighin G, & Quinlan LR (2017). When a step is not a step! Specificity analysis of five physical activity monitors. PLoS One, 12(1), 1–14.

Phillips SM, Cadmus-Bertram L, Rosenberg D, Buman MP, & Lynch BM (2018). Wearable technology and physical activity in chronic disease: Opportunities and challenges. American Journal of Preventive Medicine, 54(1), 144–150. [PubMed: 29122356]

Rich C, Geraci M, Griffiths L, Sera F, Dezateux C, & Cortina-Borja M (2013). Quality control methods in accelerometer data processing: Defining minimum wear time. PLoS One, 8(6), e67206. doi:10.1371/journal.pone.0067206 [PubMed: 23826236]

Shcherbina A, Mikael Mattsson C, Waggott D, Salisbury H, Christle JW, Hastie T, … Ashley EA (2017). Accuracy in wrist-worn, sensor-based measurements of heart rate and energy expenditure in a diverse cohort. Journal of Personalized Medicine, 7(2), 3. doi:10.3390/jpm7020003

Sisson SB, Camhi SM, Tudor-Locke C, Johnson WD, & Katzmarzyk PT (2012). Characteristics of step-defined physical activity categories in U.S. adults. American Journal of Health Promotion, 26(3), 152–159. doi:10.4278/ajhp.100326-QUAN-95 [PubMed: 22208412]

Sprint G, Cook DJ, & Schmitter-Edgecombe M (2018). Unsupervised detection and analysis of changes in everyday physical activity data. Intelligent Systems Reference Library, 137, 97–122.

Tedesco S, Sica M, Ancillao A, Timmons S, Barton J, & O'Flynn B (2019). Accuracy of consumer-level and research-grade activity trackers in ambulatory settings in older adults. PLoS One, 14(5), e0216891. doi:10.1371/journal.pone.0216891 [PubMed: 31112585]

Thomas S, Yingling L, Adu-Brimpong J, Mitchell V, Ayers CR, Wallen GR, … Powell-Wiley TM (2017). Mobile health technology can objectively capture physical activity (PA) targets among African-American women within resource-limited communities—The Washington, D.C. Cardiovascular Health and Needs Assessment. Journal of Racial and Ethnic Health Disparities, 4(5), 876–883. doi:10.1007/s40615-016-0290-4

Troiano RP, Berrigan D, Dodd KW, Mâsse LC, Tilert T, & Mcdowell M (2008). Physical activity in the United States measured by accelerometer. Medicine & Science in Sports & Exercise, 40(1), 181–188. doi:10.1249/mss.0b013e31815a51b3 [PubMed: 18091006]

Tucker JM, Welk GJ, & Beyler NK (2011). Physical activity in U.S. adults: Compliance with the physical activity guidelines for Americans. American Journal of Preventive Medicine, 40(4), 454–461. doi:10.1016/j.amepre.2010.12.016 [PubMed: 21406280]

Tully MA, McBride C, Heron L, & Hunter RF (2014). The validation of Fitbit Zip™ physical activity monitor as a measure of free-living physical activity. BMC Research Notes, 7(1), 952. doi:10.1186/1756-0500-7-952 [PubMed: 25539733]

U.S. Department of Health and Human Services. (2018). 2018 Physical Activity Guidelines Advisory Committee scientific report. Washington, DC: Author. Retrieved from https://health.gov/sites/default/files/2019-09/PAG_Advisory_Committee_Report.pdf

Wahid A, Manek N, Nichols M, Kelly P, Foster C, Webster P, … Scarborough P (2016). Quantifying the association between physical activity and cardiovascular disease and diabetes: A systematic review and meta-analysis. Journal of the American Heart Association, 5(9). doi:10.1161/JAHA.115.002495

Weatherall J, Paprocki Y, Meyer TM, Kudel I, & Witt EA (2018). Sleep tracking and exercise in patients with type 2 diabetes mellitus (Step-D): Pilot study to determine correlations between Fitbit data and patient-reported outcomes. JMIR mHealth and uHealth, 6(6), e131. doi:10.2196/mhealth.8122 [PubMed: 29871856]

Welk GJ, Bai Y, Lee JM, Godino JOB, Saint-Maurice PF, & Carr L (2019). Standardizing analytic methods and reporting in activity monitor validation studies. Medicine & Science in Sports & Exercise, 51(8), 1767–1780. doi:10.1249/MSS.0000000000001966 [PubMed: 30913159]

Yingling LR, Brooks AT, Wallen GR, Peters-Lawrence M, McClurkin M, Cooper-McCann R, … Powell-Wiley TM (2016). Community engagement to optimize the use of web-based and wearable technology in a Cardiovascular Health and Needs Assessment Study: A mixed methods approach. JMIR mHealth and uHealth, 4(2), e38. doi:10.2196/mhealth.4489 [PubMed: 27113680]

Yingling LR, Mitchell V, Ayers CR, Peters-Lawrence M, Wallen GR, Brooks AT, … Powell-Wiley T (2017). Adherence with physical activity monitoring wearable devices in a community-based population: Observations from the Washington, D.C., Cardiovascular Health and Needs Assessment. Translational Behavioral Medicine, 7(4), 719–730. doi:10.1007/s13142-016-0454-0 [PubMed: 28097627]

Table 1

Description of the Standard and New Methods for Defining Wear Time

Variable	Standard method	New method
Method description	≥10 hr of wear time per day	≥10 hr of valid minutes per day. Restriction of minute-by-minute data through removal of invalid minutes, sleeping time, and unreliable steps
Data processing components		
Presence of heart rate monitored minute	Heart rate detected by Fitbit device	Heart rate detected by Fitbit device
Heart rate threshold during minutes with few detected steps (≤2 steps)	None	≥2 *SD*s from below the mean
Sleep restrictions	None	Removal of sleeping hours (11 p.m.–5 a.m. vs. 12 p.m.–6 a.m.)
Valid day definition	≥10 hr of wear time per day, with wear time defined by measurable heart rate	≥10 hr of wear time per day, with wear time defined by measurable heart rate

Table 2

Individual Sociodemographic and Health-Related Characteristics (*n* = 16)

Age (years), mean (*SD*)	62.1 (6.6)
Employment status, *n* (%)	
Full-time employment	4 (26.7)
Part-time employment	2 (13.3)
Retired	8 (50.0)
Unemployed	2 (12.5)
Income (*n* = 8), *n* (%)	
<$19,999	0
$20,000–49,999	3 (18.8)
$50,000–79,999	3 (18.8)
$80,000–99,999	2 (12.5)
>$100,000	0
Education, *n* (%)	
Some college, or below	4 (25.0)
Technical degree	2 (13.3)
College degree	7 (43.8)
Graduate/professional degree	3 (20.0)
Marital status, *n* (%)	
Single/divorced/widowed	12 (75.0)
Married	4 (25.0)
Weight status	
BMI (kg/m^2), mean (*SD*)	35.5 (8.29)
Obese (BMI ≥ 30 kg/m^2), *n* (%)	12 (75.0)
Overweight/obese (BMI ≥ 25 kg/m^2), *n* (%)	16 (100.0)

Note. BMI = body mass index.

Table 3

Comparison of Physical Activity Metrics Using the Two Methods

Physical activity metric	Standard method[a]	New method[b]	*p* value[c]
Number of valid days	292	281	
Total wear time per day, hr	20.5 (4.3)	16.3 (2.2)	**<.0001**
Steps per day, mean	7,436 (3,543)	7,298 (3,501)	.64

[a] ≥10 hr of wear time detected by heart rate sensor.

[b] ≥10 hr of valid minutes (determined by heart rate and waking hours) as described in Table 2.

[c] Mean steps per day and hours of wear time per day were compared using independent sample *t* tests.

Significance for bold values is $p \leq .05$

Table 4

Mixed-Effects Modeling of Individual Participant Characteristics and Average Steps per Person-Day, by Week (n = 47, Person-Weeks)

Variable	Estimate	SE	*p*
Standard method[a]			
Intercept	17,281	8,047	.06
Week 2	679	342	.06
Week 3	192	342	.58
Age (years)	−106	98	.31
BMI (kg/m^2)	−86	74	.27
College education (ref = below college)	−355	1,078	.75
Married (ref = single/widow/divorced)	−1,618	1,154	.19
Employed (ref = retired/unemployed)	−37	1,256	.98
New method[b]			
Intercept	17,580	7,274	**.04**
Week 2	703	342	**.05**
Week 3	280	342	.42
Age (years)	−110	91	.26
BMI (kg/m^2)	−93	69	.21
College education (ref = below college)	−453	1,001	.66
Married (ref = single/widow/divorced)	−1,672	1,072	.15
Employed (ref = retired/unemployed)	136	1,167	.91

Note. BMI = body mass index.

[a] ≥10 hr of wear time detected by heart rate sensor.

[b] ≥10 hr of valid minutes (determined by heart rate and waking hours).

Significance for bold values is $p \leq .05$