

Review

Accuracy of Fitbit Devices: Systematic Review and Narrative Syntheses of Quantitative Data

Lynne M Feehan^{1,2}, PhD; Jasmina Geldman², MSc; Eric C Sayre², PhD; Chance Park²; Allison M Ezzat^{3,4}, MSc; Ju Young Yoo², BSc; Clayon B Hamilton^{1,2}, PhD; Linda C Li^{1,2}, PhD

¹Department of Physical Therapy, University of British Columbia, Vancouver, BC, Canada

²Arthritis Research Canada, Richmond, BC, Canada

³School of Population and Public Health, University of British Columbia, Vancouver, BC, Canada

⁴BC Children's Hospital Research Institute, Vancouver, BC, Canada

Corresponding Author:

Lynne M Feehan, PhD
Department of Physical Therapy
University of British Columbia
Friedman Building
2177 Wesbrook Mall
Vancouver, BC, V6T 1Z3
Canada
Phone: 1 604 822 7408
Email: lynnefeehan@gmail.com

Abstract

Background: Although designed as a consumer product to help motivate individuals to be physically active, Fitbit activity trackers are becoming increasingly popular as measurement tools in physical activity and health promotion research and are also commonly used to inform health care decisions.

Objective: The objective of this review was to systematically evaluate and report measurement accuracy for Fitbit activity trackers in controlled and free-living settings.

Methods: We conducted electronic searches using PubMed, EMBASE, CINAHL, and SPORTDiscus databases with a supplementary Google Scholar search. We considered original research published in English comparing Fitbit versus a reference- or research-standard criterion in healthy adults and those living with any health condition or disability. We assessed risk of bias using a modification of the Consensus-Based Standards for the Selection of Health Status Measurement Instruments. We explored measurement accuracy for steps, energy expenditure, sleep, time in activity, and distance using group percentage differences as the common rubric for error comparisons. We conducted descriptive analyses for frequency of accuracy comparisons within a $\pm 3\%$ error in controlled and $\pm 10\%$ error in free-living settings and assessed for potential bias of over- or underestimation. We secondarily explored how variations in body placement, ambulation speed, or type of activity influenced accuracy.

Results: We included 67 studies. Consistent evidence indicated that Fitbit devices were likely to meet acceptable accuracy for step count approximately half the time, with a tendency to underestimate steps in controlled testing and overestimate steps in free-living settings. Findings also suggested a greater tendency to provide accurate measures for steps during normal or self-paced walking with torso placement, during jogging with wrist placement, and during slow or very slow walking with ankle placement in adults with no mobility limitations. Consistent evidence indicated that Fitbit devices were unlikely to provide accurate measures for energy expenditure in any testing condition. Evidence from a few studies also suggested that, compared with research-grade accelerometers, Fitbit devices may provide similar measures for time in bed and time sleeping, while likely markedly overestimating time spent in higher-intensity activities and underestimating distance during faster-paced ambulation. However, further accuracy studies are warranted. Our point estimations for mean or median percentage error gave equal weighting to all accuracy comparisons, possibly misrepresenting the true point estimate for measurement bias for some of the testing conditions we examined.

Conclusions: Other than for measures of steps in adults with no limitations in mobility, discretion should be used when considering the use of Fitbit devices as an outcome measurement tool in research or to inform health care decisions, as there are seemingly a limited number of situations where the device is likely to provide accurate measurement.

(JMIR Mhealth Uhealth 2018;6(8):e10527) doi: [10.2196/10527](https://doi.org/10.2196/10527)

KEYWORDS

wearable activity tracker; accuracy; Fitbit; steps; sleep; energy expenditure; distance; time in activity; systematic review; fitness trackers; data accuracy; energy metabolism; review

Introduction

Commercially available wearable activity trackers have grown rapidly in popularity since their introduction just over a decade ago [1]. While the technologies behind them are quickly and continuously changing, in general they are small devices that are commonly worn on the wrist or attached to clothing. They aim to provide the user with real-time feedback on various aspects of daily activities, such as number of steps taken, energy expenditure, time spent asleep, and time spent in different levels of activity. They also typically provide personal goal-setting options, summary data, and visualizations through synchronization with interactive mobile- and computer-based apps, as well as opportunities to connect to social media and other health and fitness apps. These devices are aimed primarily at health- and fitness-conscious consumers and are designed to motivate and offer support to individuals to self-monitor and increase their daily physical activity.

Fitbit (Fitbit Inc, San Francisco, CA, USA), one of the most popular commercial wearable activity trackers, holds approximately 20% of the market share for wearable tracking devices, with more than 63 million devices sold worldwide in the last 10 years [2]. In 2017, the company sold 15 million devices and had 25 million active users [2]. The Classic model was introduced in 2009 as a clip-on device to be worn on the torso; new models of clip-on devices became commercially available in 2011 with the introduction of the Ultra, Zip, and One models. In 2013, Fitbit introduced a line of wristband activity trackers: Force, Flex (2), Charge (2, HR), and Alta (HR).

Fitbit devices use a microelectronic triaxial accelerometer to capture body motion in 3-dimensional space, with these motion data analyzed using proprietary algorithms to identify patterns of motion to identify daily steps taken, energy expenditure, sleep, distance covered, and time spent in different intensity of activities. Although designed as a consumer product to help motivate individuals to be physically active, Fitbit devices are becoming increasingly popular as measurement tools in physical activity and health promotion research and are also commonly used to inform patient–health professional interactions [3-7]. Between 2011 and 2017, a total of 171 clinical trials registered at ClinicalTrials.gov used Fitbit as an outcome measurement tool; 97 of those were registered in the last 3 years [8]. Most of the registered trials identified number of steps taken as the outcome of interest, followed, in order, by time in activity, sleep, energy expenditure, and distance covered.

Fitbit devices, and particularly the wrist-worn devices, have demonstrated dependability, durability, and acceptability [9,10]. A 2015 systematic review by Evenson et al examined the “validity and reliability of...[Fitbit devices] and their ability to estimate steps, distance, physical activity, energy expenditure,

and sleep” [11]. They concluded that Fitbit devices were moderately associated with criterion reference devices for measures of steps, sleep, and distance, with associations varying from poor to moderate with criterion reference devices for measures of energy expenditure and time in activity [11]. They also found that Fitbit had a high interdevice reliability for all outcome measures. In addition, the review provided some data for measurement accuracy; however, it did not comprehensively examine device measurement accuracy or study quality. Measurement accuracy, or how close to “true” the measured value is, is an important consideration, as Fitbit devices are being used as an outcome measurement tool in research and to inform health care decisions [12,13]. Therefore, the purpose of this review was to systematically examine and report the accuracy of measures derived from the triaxial accelerometry data in Fitbit devices—that is, measures of steps, energy expenditure, sleep, distance, and time in activity—when used by adults in controlled and free-living settings.

Methods

Search Strategy

We conducted an electronic literature search of the PubMed, EMBASE, CINAHL, and SPORTDiscus databases, with an additional supplementary search conducted via Google Scholar. Keywords within each database search included variations on the terms Fitbit AND Accuracy (accura*) OR Validity / Validation (valid*) OR Comparison / Comparative (compar*) OR Relationship (relation*) OR Association (associa*) OR Equivalence (equival*) OR Agreement ([Multimedia Appendix 1](#)). We applied a language filter to limit results to English and a date limitation from January 1, 2011 (Fitbit devices were not commercially available prior to this date) to October 31, 2017 (the end date for our search). We applied no further search limits or filters. We also hand searched reference lists of the included studies for potentially eligible studies.

Study Selection and Eligibility Criteria

We screened all citations, removed duplicates, and assessed the remaining titles and abstracts for potential eligibility. All potentially eligible citations were retrieved for full-text review by 2 independent reviewers (JYY, JG) with disagreements resolved through consensus. Initial inclusion criteria were original research studies, published in a full, short, or letter format in English in peer-reviewed journals. We verified journal peer-review status using a Web-based serial directory database search (ULRICHSWEB, ProQuest LLC, Ann Arbor, MI, USA). The studies also had to include or separately report data for adults (≥ 18 years old) and examine measurement accuracy for one or more of the following outcome domains: steps, energy expenditure, time in activity, distance, or sleep. Studies could be conducted in any controlled-testing (ie, using a standardized testing protocol) or free-living (ie, during usual daily activity)

setting and could include individuals living with any health, disease, or mobility or functional status. Studies examining accuracy in controlled settings had to compare a Fitbit measure against a predefined reference-standard criterion measure, whereas studies conducted in free-living settings had to compare the Fitbit measure against a predefined research-standard criterion measure (Multimedia Appendix 2). To be included in the final review, studies had to have extractable data for one or more of the following accuracy analyses: group mean or percentage differences, mean or median absolute percentage error (MAPE), or level-of-agreement analyses [14]. We did not contact authors if these data were not reported in the publication. We excluded studies (or comparisons) if the accuracy evaluations were conducted on 10 or fewer participants. We also excluded studies (or comparisons) if they examined heart rate accuracy, as heart rate measurement is not derived from accelerometry data.

Data Extraction

Data were extracted and checked for accuracy by a second independent reviewer (JYY, JG, AME, LMF) with discrepancies resolved through discussion and consensus. Data extracted included study, participant, and Fitbit device characteristics, as well as details about the study setting, outcomes examined, and reference criterion used (Multimedia Appendix 2). Group mean or percentage difference values for the Fitbit device and criterion groups were extracted for all accuracy comparisons reported in each study. If group percentage difference was not reported, we calculated group percentage error ($[\text{Fitbit}_{\text{mean}} - \text{Criterion}_{\text{mean}}] / \text{Criterion}_{\text{mean}} \times 100$) to allow for a common unit of measure (rubric) for comparison of accuracy measures within and across outcome domains (Multimedia Appendix 3). We also extracted reported MAPE or level-of-agreement accuracy data when available.

Risk-of-Bias Assessment

All articles were independently assessed for risk of bias by 2 independent reviewers (JG, CP) using a modification of the validation subscale from the checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments (Consensus-Based Standards for the Selection of Health Status Measurement Instruments [COSMIN]) [15]. All discrepancies were resolved by discussion and consensus, or by a third independent reviewer (LMF). Quality evaluation included 5 design or methodology components (percentage missing data, missing data management, adequate sample size, acceptable criterion comparison, design or methodological flaws) and one analysis component (acceptable accuracy analyses). We rated each dimension as excellent, good, fair, or poor quality based on a priori modifications to the COSMIN validation subscale scoring criteria appropriate for accuracy studies (Multimedia Appendix 4) [16].

Data Handling

We sorted each accuracy comparison into one of the following outcome domains: (1) steps, (2) energy expenditure, (3) sleep, (4) time in activity, or (5) distance. Within each domain, we coded individual accuracy comparisons to identify testing

parameters that may influence measurement accuracy, such as variations in the testing environment(s), placement of device(s), or variations in the type of ambulation or activity or task examined (Multimedia Appendix 5). All coding was independently reviewed by a second reviewer with discrepancies resolved through discussion and consensus (LMF, JG).

Syntheses

Given the diversity of outcomes reported and the variety of testing conditions under which accuracy measures were examined and reported across and within different studies, we were unable to conduct meta-analyses. As an alternative, and as recommended by the UK Economic and Social Research Council guidelines for conducting and reporting narrative syntheses, we conducted a narrative synthesis of quantitative data, where we explored measurement accuracy within each outcome domain (ie, steps, energy expenditure, sleep, time in activity, and distance) using group percentage difference as the common rubric for measurement error comparisons [17,18]. We performed descriptive analyses for frequency (number and percentage) of percentage error comparisons that were within and outside predefined cutoff points for measurement accuracy in controlled or free-living settings. We also explored potential trends for direction of measurement error (ie, potential measurement bias) by defining a point estimation for both mean and median percentage error, with negative values indicating a trend for Fitbit device underestimation compared with the criterion device. In addition, we explored measurement error dispersion by defining the range (maximum–minimum) for percentage error measures. Given the diversity of testing conditions, we conducted further secondary exploratory analyses for comparisons of steps and energy expenditure accuracy in controlled settings to examine the potential influence of different testing parameters such as variations in body placement, ambulation speed, or variations in the type of activity on measurement accuracy. We completed these secondary exploratory analyses only when there were 10 or more accuracy comparisons within each subgroup.

We provide summaries for all descriptive analyses in tabular formats. For selected secondary subanalyses, we also provide modified scatter plots depicting the distribution of accuracy comparisons for group percentage error, color coded by variations in testing parameters, to allow for visual interpretation of how measurement error may be influenced by variations in testing parameters.

We focused our interpretation of measurement accuracy based on predefined acceptable limits for measurement accuracy in controlled settings as a percentage difference of $\pm 3\%$ and acceptable limits for relative accuracy in free-living settings as a percentage difference of $\pm 10\%$ [19–22]. We completed all descriptive analyses and plots using SAS version 9.4 software (SAS Institute Inc).

In the review, we included accuracy studies not reporting data to allow for the examination of group percentage measurement error if they reported MAPE or level-of-agreement data. These studies were included in the syntheses of study characteristics and the risk-of-bias assessment. As well, we provide narrative summaries for how the reported accuracy from these studies

may or may not be consistent with our evaluation of percentage measurement error.

Results

Study Selection

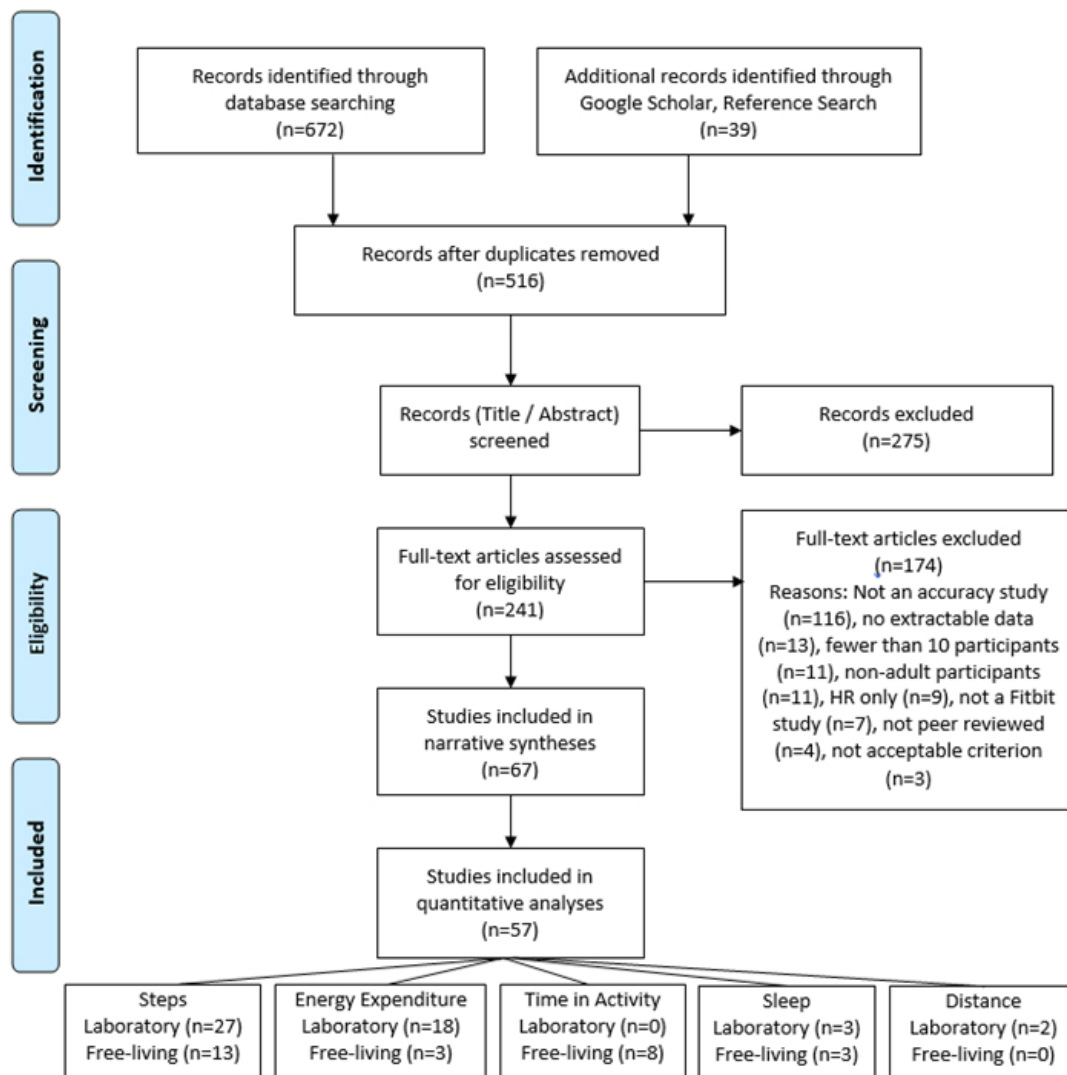
We identified 711 citations, of which we screened 516 titles and abstracts for potential eligibility after removing duplicates. Following screening, we excluded 275 titles, with the remaining 241 full-text reviewed. After full-text review, we subsequently excluded 174 articles. A total of 67 studies met the final inclusion criteria, with 57 providing adequate data for inclusion in the quantitative analyses. Of these, 40 studies investigated step count (laboratory: n=27; free-living settings: n=13), 21 addressed energy expenditure (laboratory: n=18; free-living settings: n=3), 8 examined time spent in different intensities of activity in free-living settings, 6 examined sleep measurements (laboratory: n=3; free-living settings: n=3), and 2 examined distance walked in a controlled testing environment (Figure 1) [23].

Study and Participant Characteristics

Publication dates varied from 1 publication in each of 2012 and in 2013, to 8 publications in each of 2014 and in 2015, with 49 studies published in or after 2016. Publications were from 11 countries across North America, Western Europe, South Asia, and Australia. The largest number of publications were from the United States (n=39), followed by Australia (n=9) and Canada (n=5). Of the 67 publications, 61 were full research articles, 5 were short reports, and 1 was a letter to the editor (Multimedia Appendix 6).

The 67 studies comprised a total of 2441 participants, with the mean number being 36 (SD 25), varying from 12 to 166. Of the 61 studies reporting age, the mean age of participants was 37 (SD 18) years, varying from 21 to 84 years. Of the 65 studies reporting sex, 53.95% (1251/2319) of the participants were female. A total of 55 studies included only healthy participants, with the remaining 12 including participants living with a variety of chronic diseases or mobility limitations, or both (Multimedia Appendix 6). Studies used several models of Fitbit devices, including the Ultra, Classic, Zip, or One worn on the torso (waist, hip, or chest), the Flex, Charge HR, Force, or Surge worn on the wrist, and the One worn on the ankle.

Figure 1. Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow chart.



Risk of Bias

We rated the vast majority of the 67 studies as excellent or good for study design, reporting of missing data, and use of an acceptable reference criterion measure (Multimedia Appendix 7). For 34 studies (43 accuracy comparisons), it was unclear how missing data were handled in the analyses (Multimedia Appendix 7). We did not exclude these accuracy comparisons from the descriptive analyses of percentage measurement error based on this criterion. However, we did exclude 21 accuracy comparisons in the descriptive analyses, as the Fitbit versus criterion group mean or percentage differences were not reported (Multimedia Appendix 7). Rather than excluding these accuracy comparisons (or studies) completely from the review, we provide a narrative summary for how the reported MAPE or level-of-agreement accuracy data may or may not be consistent with our exploration of percentage measurement error.

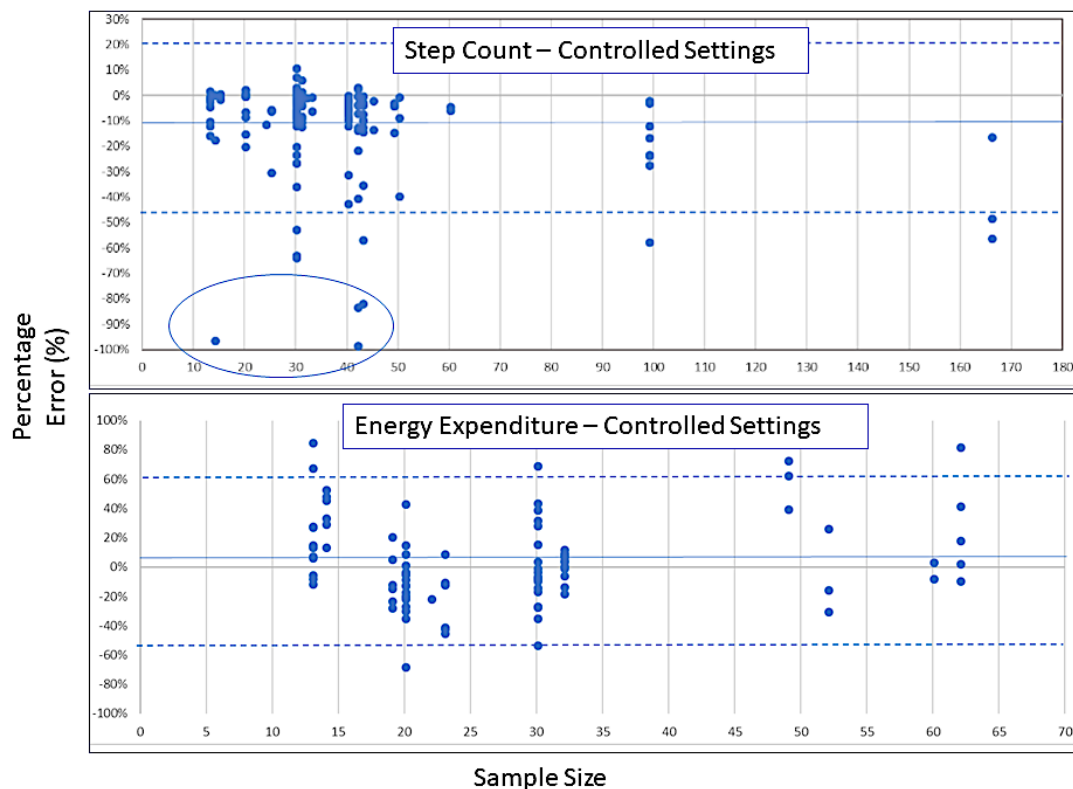
We also did not exclude 55 studies (85 accuracy comparisons) rated fair or poor for sample size (<50 participants), as there were only 2 studies with 100 or more participants (excellent rating) and 10 with 50 to 99 participants (good rating) (Multimedia Appendix 7). Rather, we excluded studies (or accuracy comparisons) with 10 or fewer participants. As well, for step count and energy expenditure in controlled settings, we explored the potential for bias based on sample size by exploring the dispersion of group percentage error across different sample sizes using modified scatter plots (Figure 2). In these exploratory

analyses, we saw no apparent systematic bias for measurement error, other than a slight tendency for extreme underestimation of steps in 4 comparisons from 2 studies with fewer than 50 participants. However, when we explored these extreme outliers further, we determined that they were likely true reflections of a greater tendency for underestimation of step count during very slow walking activities when the device was worn on the torso, rather than due to small sample size. Therefore, we included all percentage error accuracy comparisons, independent of sample size, in our descriptive analyses.

Step Count

A total of 27 studies (191 accuracy comparisons) examined Fitbit device step measurements compared with a reference-standard criterion of direct observation and counting of steps in a controlled setting (Multimedia Appendix 3) [12,24-49]. Of these, 21 studies recruited healthy adults with a mean age of 37.2 (SD 18.3) years; the remaining 6 recruited adults living with limited mobility or chronic disease with a mean age of 64.8 (SD 14.8) years. Fitbit devices were worn on the torso, wrist, or ankle. Across the 191 accuracy comparisons examining step count in controlled settings, 46% (n=88) were within a $\pm 3\%$ measurement error, 51% (n=97) were below a -3% measurement error, and 3% (n=6) were above a 3% measurement error, with an overall tendency for Fitbit devices to underestimate steps (estimated mean [median] difference of -9% [-3%]) (Multimedia Appendix 8 and Figure 2).

Figure 2. Percentage error distribution by sample size. Top: Step count in controlled settings. The blue oval indicates extreme outliers (n=4 comparisons). Bottom: Energy expenditure in controlled settings. Solid blue lines indicate mean error estimation. Dotted blue lines indicate 95% CI.



When we further explored factors potentially influencing step count accuracy, we observed that accuracy of step count in controlled settings seemed to vary with speed of ambulation (jog, normal, self-paced, slow, or very slow) [50,51], with body placement (torso, wrist, or ankle), and with variations in how the body moved during the activity (normal, constrained, variable, or exaggerated) (Multimedia Appendix 8). Constrained body motion throughout the task or activity could be due to, for example, a disease-related mobility limitation, walking with a walking aide, or pushing a stroller while walking. Levels of body motion could have varied while performing a series of different simulated household tasks or doing simulated agility-dependent sporting activities. Exaggerated motions could have occurred when the device was worn on the wrist during simulated household or sporting activities that involved exaggerated arm motions.

Within the different speeds of ambulation, measurement error was within $\pm 3\%$ more than 50% of the time for jogging (14/24) or normal (25/48) ambulation speeds. More than 50% of the time, measurement error was below -3% for self-paced (35/70), slow (12/23), and very slow (19/26) ambulation speeds. Within each ambulation speed, Fitbit tended to underestimate step counts (mean [median] error estimations varying from -24% [-12%] to -4% [-2%]) (Multimedia Appendix 8 and Figure 3).

Within the different body placements for the device, measurement error was within $\pm 3\%$ more than 50% of the time for comparisons with torso (65/114) or ankle (8/16) placement, whereas 70% (43/61) of the time measurement error was below -3% for wrist placement. Within each body placement, Fitbit tended to underestimate steps (estimated mean [median] errors varying from a -11% [-2%] to -3% [-1%]) (Multimedia Appendix 8 and Figure 4).

Within the variations of body motion during the activity, measurement error was within $\pm 3\%$ more than 50% (82/154) of the time during activities with normal body motion. Measurement error was below -3% more than 90% of the time for activities that involved constrained (19/24) or variable (10/10) body motion during the activity, with Fitbit tending to underestimate steps during these activities (estimated mean [median] errors varying from -35% [-26%] up to -21% [-12%]). Conversely, when the Fitbit device was worn on the wrist during exaggerated arm motion, 2 of the 3 comparisons were above 3% (Multimedia Appendix 8 and Figure 5).

We also observed that, within the different speeds of ambulation, step count accuracy appeared to be influenced further by the placement of the device on the body (Figures 3 and 4). For torso placement, measurement accuracy was within $\pm 3\%$ more than 60% of the time for normal (24/30), self-paced (28/44), and slow (7/11) ambulation speeds. Torso placement was lower than -3% more than 60% of the time for jogging (9/14) and more than 90% of the time for very slow (14/15) walking speeds. In addition, we observed that the underestimation of steps was largest during very slow walking when the device was worn on the torso, with 7 of these 15 comparisons having a measurement error lower than -25% . For ankle placement, 70% (11/16) of the accuracy comparisons were within $\pm 3\%$ measurement error for slow or very slow walking speeds. There were no accuracy comparisons for ankle placement at normal or jogging speeds and only 1 comparison for ankle placement during self-paced ambulation. For wrist placement, 90% (9/10) of time measurement error was within $\pm 3\%$ for jogging speeds and 75% (38/51) of the time it was lower than -3% for all other speeds.

Figure 3. Step count percentage error in controlled settings. Speed (jog, normal, self-paced, slow, very slow) by body placement (torso, wrist, ankle) of the Fitbit device. Dark lines indicate mean (horizontal). Dashed lines indicate median (horizontal). Gray shading indicates $\pm 3\%$ measurement error.

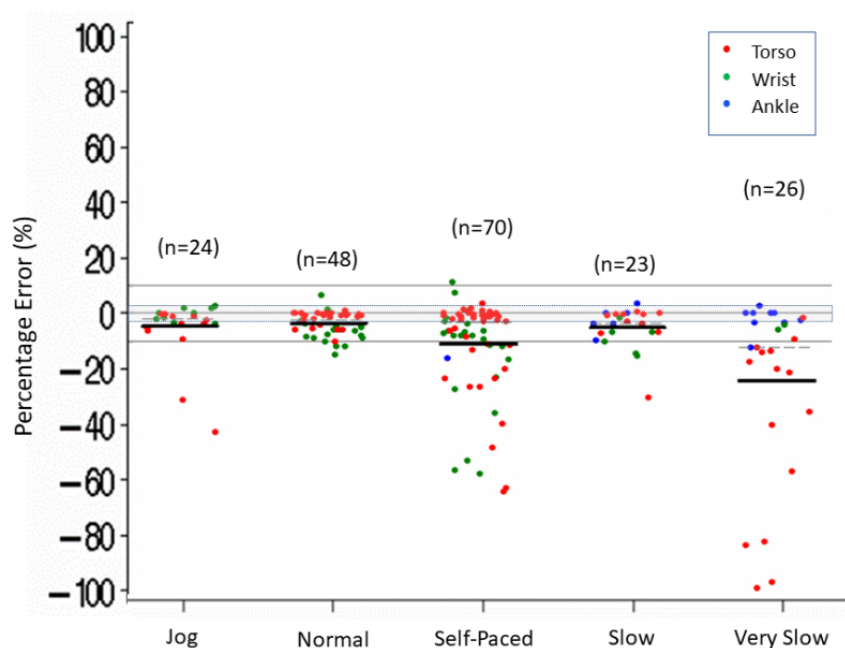


Figure 4. Step count percentage error in controlled settings. Body placement (torso, wrist, ankle) of the Fitbit device by speed (jog, normal, self-paced, slow, very slow). Dark lines indicate mean (horizontal). Dashed lines indicate median (horizontal). Gray shading indicates $\pm 3\%$ measurement error.

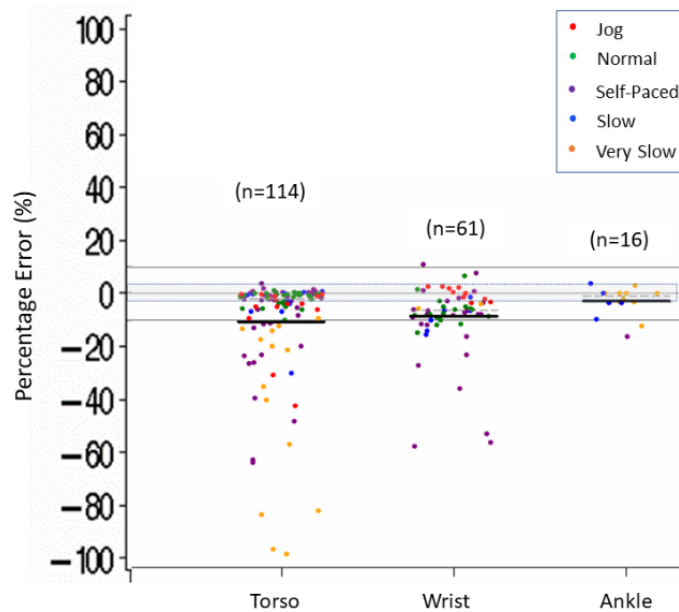
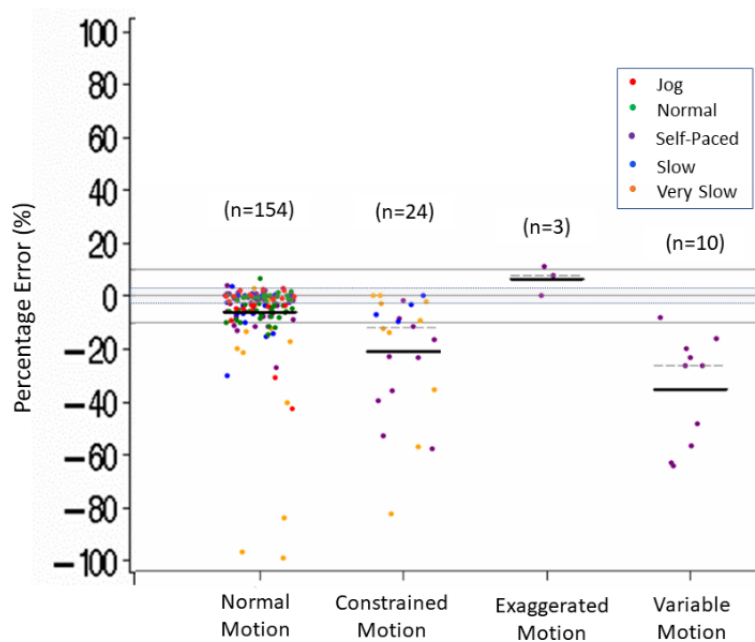


Figure 5. Step count percentage error in controlled settings. Body motion (normal, constrained, exaggerated, variable) by speed (jog, normal, self-paced, slow, very slow). Dark lines indicate mean (horizontal). Dashed lines indicate median (horizontal).



A total of 13 studies examined Fitbit accuracy for step count in free-living conditions (20 accuracy comparisons; [Multimedia Appendix 3](#)) [36,52-63]. Of these, 8 studies were conducted in healthy young adults; 5 were conducted on older adults, of whom 3 were healthy, active older adults and 2 had mobility limitations. Duration of wear varied from 1 to 14 days. Fitbit devices were compared with ActiGraph, activPAL, or Actical accelerometers, or an Omron or Shimmer pedometer.

Across the 20 accuracy comparisons examining step count in free-living settings, 55% (n=11) were within a $\pm 10\%$ measurement error, 30% (n=6) were below a -10% measurement error, and 15% (n=3) were above a 10% measurement error, with a tendency for Fitbit to overestimate steps in free-living

conditions relative to a research-grade criterion. When explored further, it appeared that measurement error for step count in free-living conditions varied depending on the reference criterion used, body placement of the device, and the age and mobility status of the study participants. Compared with ActiGraph or activPAL accelerometers, Fitbit step count was within a $\pm 10\%$ error for 6 of 8 torso comparisons and 3 of 5 wrist comparisons in healthy young adults, and in 1 comparison when worn on the torso in older adults with no mobility limitation. In 1 comparison in older adults with no mobility limitations, a Fitbit device worn on the torso overestimated steps by more than 35% relative to an Omron pedometer worn on the ankle.

In contrast, in 2 of 3 accuracy comparisons in older adults with mobility limitations, Fitbit step count error was approximately –25% lower than that of a Shimmer pedometer or an Actical accelerometer worn on the ankle when Fitbit was worn on the torso (Multimedia Appendix 9).

Our evaluation for step count accuracy in free-living settings was consistent with those of 5 other studies examining MAPE or level-of-agreement differences in daily step measures in healthy adults for Fitbit compared with an accelerometer worn on the torso or a pedometer worn on the ankle [42,64–67]. These studies reported Fitbit overestimations of median steps per day varying from 700 to 1800 steps or MAPE values greater than 10% when compared with an ActiGraph accelerometer or Omron or New Life pedometers. In contrast, 1 study showed similar measures for median steps per day for Fitbit compared with a Yamax pedometer (–55 steps/day) [67].

Energy Expenditure

A total of 18 studies (98 accuracy comparisons) examined Fitbit device energy expenditure measurement accuracy in controlled settings compared with a reference standard of direct (2 studies) or indirect (16 studies) calorimetry (Multimedia Appendix 3) [29,34,37,39,46,68–80]. All 18 studies recruited healthy adults. Fitbit devices were worn on the torso or wrist. Of the accuracy comparisons, 88 measured energy expenditure during an activity, while 10 measured energy expenditure at rest.

Findings indicated that, across the 88 activity comparisons, measurement error was rarely within $\pm 3\%$ (4% [n=4] within a $\pm 3\%$ error, 47% [n=41] below a –3% error, and 49% [n=43] above a 3% error). Overall, Fitbit showed a tendency to overestimate energy expenditure during activity (estimated

mean [median] error of 4% [2%]). Across the 10 comparisons at rest, 3 were within a $\pm 3\%$ measurement error, with 6 lower than –3% and 1 higher than 3%, with a tendency to underestimate energy expenditure (estimated mean [median] error of –3% [–6%]) (Multimedia Appendix 8 and Figure 6).

When we explored further by factors potentially influencing energy expenditure measurement accuracy, we observed that accuracy appeared to vary with speed of ambulation, with body placement, and with variations in body motion during the activity. In addition, energy expenditure accuracy appeared to be influenced by type of ambulation. Types of ambulation included continuous ambulation on an incline or a flat surface, as well as intermittent ambulation (stop-and-start ambulation) while performing common simulated household or sporting activities (Multimedia Appendix 8).

Within the different body placements, measurement error for energy expenditure was lower than –3% more than 60% (32/52) of the time with torso placement (estimated mean [median] error of –5% [–8%]) and greater than 3% more than 60% (24/36) of the time for wrist placement (estimated mean [median] error of 18% [9%]) (Multimedia Appendix 8 and Figure 7).

Within the different speeds of ambulation, more than 50% of jogging (8/15) and normal (17/24) speed comparisons for energy expenditure were greater than 3% (estimated mean [median] errors varying from 7% (5%) to 18% (12%)). Conversely, more than 50% (25/39) of the self-paced ambulation comparisons were below –3% (estimated mean [median] error of –6% [–9%]). There were fewer than 10 comparisons for energy expenditure at slow and very slow ambulation speeds, with no apparent trend or pattern for measurement error noted (Multimedia Appendix 8 and Figure 8).

Figure 6. Energy expenditure percentage error in controlled settings. Activity versus rest by body placement (torso, wrist). Dark lines indicate mean (horizontal). Dashed lines indicate median (horizontal). Triangles indicate measurement by direct calorimetry. Gray shading indicates $\pm 3\%$ measurement error.

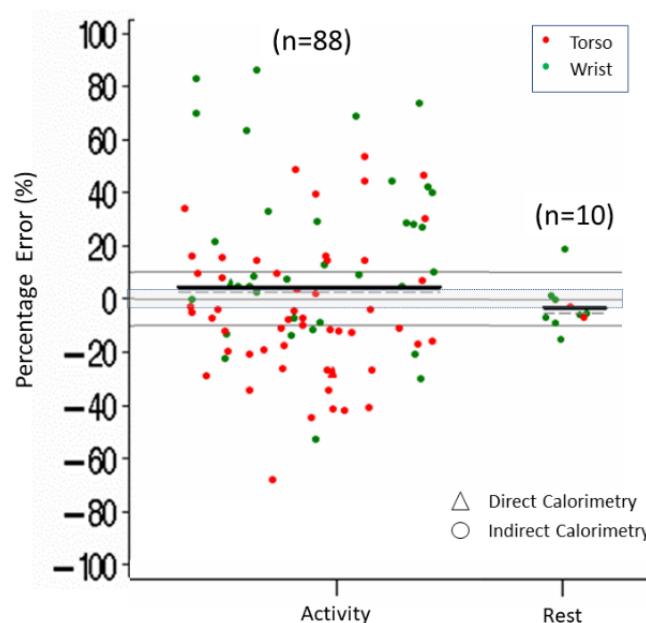


Figure 7. Energy expenditure percentage error in controlled settings. Body placement (torso, wrist) by speed (jog, normal, self-paced, slow, very slow). Dark lines indicate mean (horizontal). Dashed lines indicate median (horizontal). Triangles indicate measurement by direct calorimetry. Gray shading indicates $\pm 3\%$ measurement error.

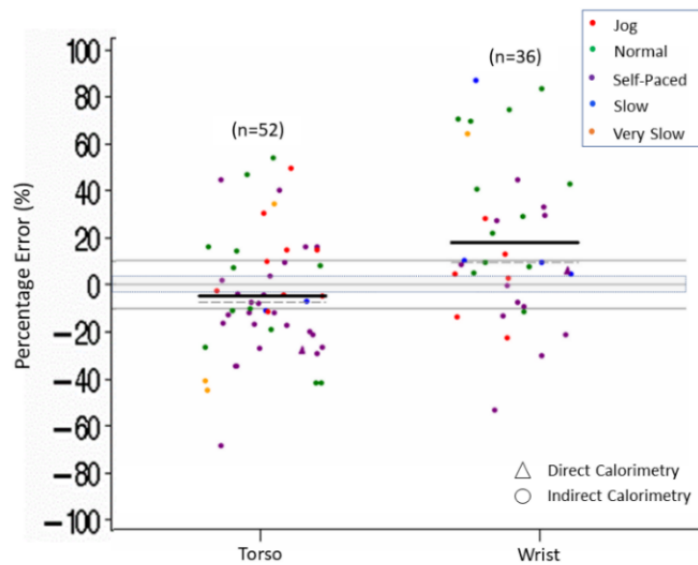
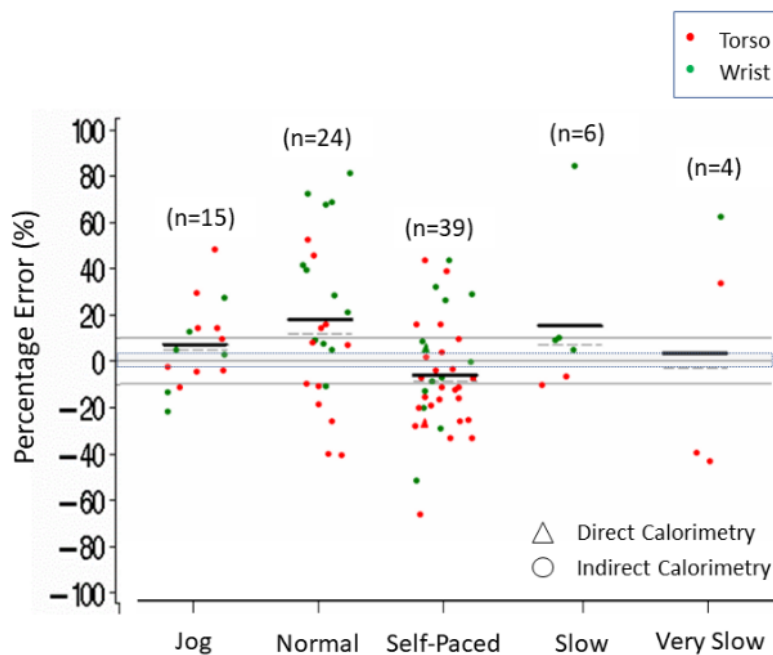


Figure 8. Energy expenditure percentage error in controlled settings. Speed (jog, normal, slow, self-paced, very slow) by body placement (torso, wrist). Dark lines indicate mean (horizontal). Dashed lines indicate median (horizontal). Triangles indicate measurement by direct calorimetry. Gray shading indicates $\pm 3\%$ measurement error.



Within the different body motion parameters, more than 50% (34/58) of activities with normal body motion had an energy expenditure error greater than 3% (estimated mean [median] difference of 12% [9%]). In contrast, more than 60% of the accuracy comparisons during activities with constrained (6/10) or variable (13/16) body motion activities had an energy expenditure error lower than -3% (estimated mean [median] difference varying from -14% [-15%] to -8% [-10%]). Similarly, 3 of the 4 comparisons with exaggerated motion also had a measurement error for energy expenditure lower than -3% (Multimedia Appendix 8 and Figure 9).

Within the different types of ambulation, more than 60% (35/53) of the continuous ambulation activities on flat surfaces had an error for energy expenditure greater than a 3% (estimated mean [median] difference of 17% [13%]). More than 60% of the time, the error was lower than -3% with continuous ambulation activities on an incline (7/11) or intermittent ambulation during simulated household or sporting activities (18/24) (estimated mean [median] errors varying from -19% [-21%] to -12% [-12%]) (Multimedia Appendix 8 and Figure 10).

Figure 9. Energy expenditure percentage error in controlled settings. Motion limitations (normal, constrained, exaggerated, variable) by speed (jog, normal, self-paced, slow, very slow). Dark lines indicate mean (horizontal). Dashed lines indicate median (horizontal). Triangles indicate measurement by direct calorimetry. Gray shading indicates $\pm 3\%$ measurement error.

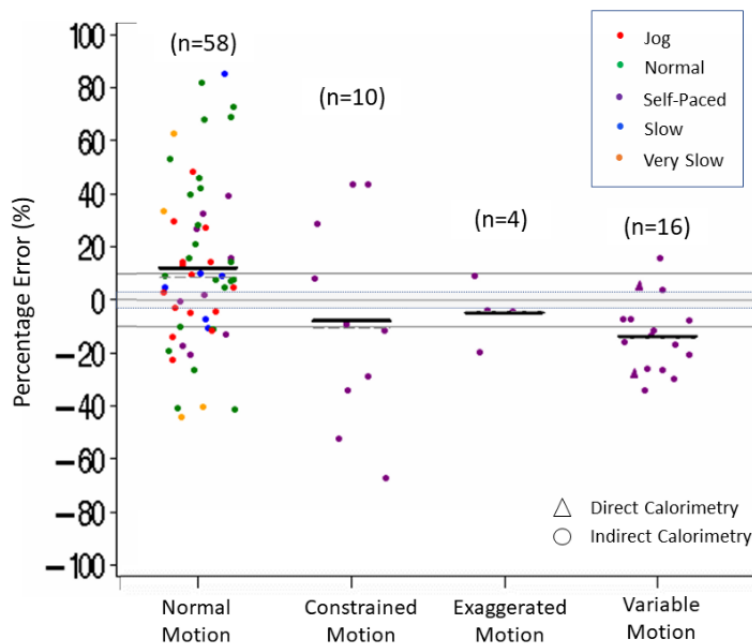
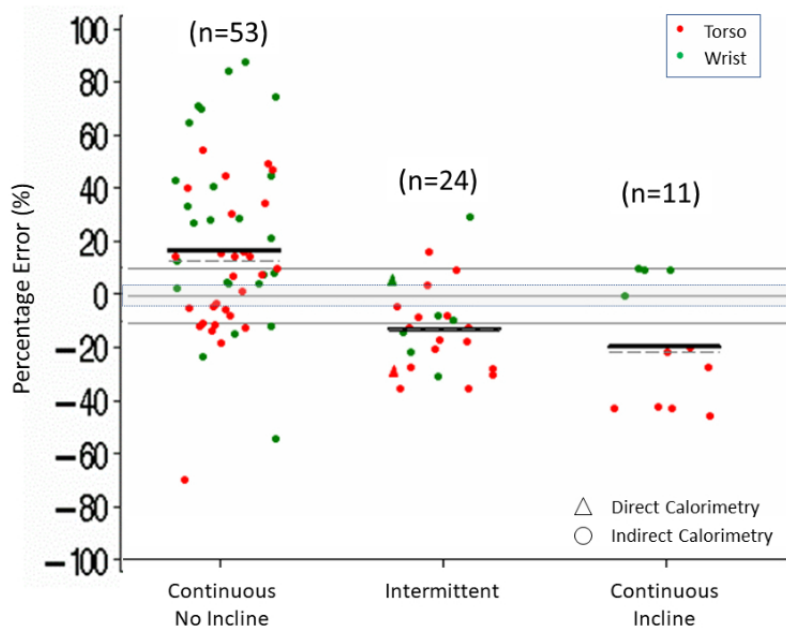


Figure 10. Energy expenditure percentage error in controlled settings. Type of ambulation (continuous no incline, continuous incline, intermittent) by body placement (torso, wrist). Dark lines indicate mean (horizontal). Dashed lines indicate median (horizontal). Triangles indicate measurement by direct calorimetry. Gray shading indicates $\pm 3\%$ measurement error.



A total of 3 studies examined Fitbit device accuracy for measures of energy expenditure in healthy adults in free-living conditions compared with doubly labelled water (1 accuracy comparison) or a SenseWear accelerometer (4 accuracy comparisons; [Multimedia Appendix 3](#)) [56,77,81]. Findings from 1 study showed that Fitbit worn on the wrist tended to slightly underestimate (-7%) energy expenditure over 15 days compared with doubly labelled water [77]. All 4 accuracy comparisons with a SenseWear accelerometer were lower than a -10% measurement error (estimated mean [median] difference of -15% [-15%]) ([Multimedia Appendix 9](#)). These findings are

consistent with those of 2 other accuracy studies reporting Fitbit underestimations of daily energy expenditure, with MAPE values varying from 16% to 30% for Fitbit devices compared with measurements from an ActiGraph or Actiheart accelerometer [58,70].

Time in Activity

A total of 8 studies (28 accuracy comparisons) examined Fitbit device measures in free-living settings for time spent in different intensities of activity compared with measures from an ActiGraph accelerometer worn on the torso or Actical accelerometer worn on the ankle ([Multimedia Appendix 3](#))

[52,53,56,57,59-62]. Of these studies, 5 were conducted on healthy young adults and 3 were conducted on older adults living with a variety of chronic diseases. The duration of wear varied from 2 to 9 days. Studies examined time spent in sedentary, light, moderate, vigorous, or moderate to vigorous physical activity during waking hours.

Notably, the Fitbit device and the reference criterion accelerometers across the studies used variable cutoff points for defining intensity levels of physical activity. Despite these differences, 3 of the 4 accuracy comparisons for sedentary time had an error lower than -10% when compared with ActiGraph (torso) or Actical (ankle) accelerometers. Compared with Actical (ankle) or ActiGraph (torso) accelerometers, more than 80% (21/24) of accuracy comparisons for time spent in light to vigorous activity time had a measurement error greater than 10% (estimated mean [median] difference varying from 44% [52%] to 632% [390%]) (Multimedia Appendix 9).

Our observation of marked overestimation of time spent in higher-intensity activity was consistent with those of 2 other studies reporting Fitbit overestimations of moderate to vigorous physical activity in free-living settings compared with an ActiGraph accelerometer (MAPEs $>30\%$) [58,66]. In contrast to our finding of Fitbit underestimation of sedentary time during the day, 1 study reported Fitbit overestimation of combined night (sleep) and daytime sedentary time (MAPE $\sim 10\%$) when compared with an activPAL accelerometer worn on the thigh. [66].

Sleep

A total of 3 studies examined sleep in controlled settings (12 accuracy comparisons), comparing a Fitbit worn on the wrist against reference-standard polysomnography over 1 night of sleeping in a laboratory (Multimedia Appendix 3) [82-84]. All 3 studies included young adults, 2 comprising healthy participants and 1 comprising individuals living with depression. All 3 studies examined measures of sleep in a normal-mode setting, and all reported Fitbit overestimation of total sleep time and sleep efficiency by more than 10%. On the other hand, 1 study examined total sleep time and sleep efficiency in the sensitive sleep mode, reporting Fitbit underestimation of both by more than 15% [82]. One study also examined sleep-onset latency (minutes to initial sleep) and time awake after sleep onset in normal and sensitive sleep modes and reported measurement errors varying from 12% to 180%, with an opposite tendency for either over- or underestimations of these sleep parameters depending on the sleep-mode setting (Multimedia Appendix 8).

A total of 3 studies (5 accuracy comparisons) reported sleep measurement accuracy in healthy young adults in free-living settings comparing a Fitbit device worn on the wrist with a SenseWear or Actiwatch accelerometer also worn on the wrist (Multimedia Appendix 3) [56,81,85]. Duration of wear varied from 1 to 13 nights of home sleep. There were 4 comparisons for measures of time in bed, with all 4 reporting measurement errors within $\pm 10\%$ compared with a SenseWear accelerometer. One study also reported very similar measures for time in bed (-0.4%) by a Fitbit device compared with an Actiwatch accelerometer. One study also reported a slight overestimate

(6%) of sleeping minutes for Fitbit compared with an Actiwatch accelerometer [85] (Multimedia Appendix 9). These findings are consistent with those of 2 other studies reporting Fitbit overestimations of sleeping time compared with a portable sleep monitor (MAPE approximately 10%) or Actiwatch accelerometer (approximately 10 minutes per night) [66,86].

Distance

There were 2 studies (17 accuracy comparisons) examining Fitbit device distance measurement accuracy in a controlled setting in healthy young adults (Multimedia Appendix 3) [33,48]. Both studies reported a Fitbit tendency to overestimate distance at slower and self-paced ambulation speeds (varying from 5% [torso] to 25% [wrist]) and underestimate distance at brisk walking or jogging speeds (varying from -15% [wrist] to -5% [torso]). During normal speed ambulation, torso placement tended to overestimate distance (10%), and wrist placement tended to slightly underestimate distance (-3%). These findings are consistent with 1 additional study reporting Fitbit overestimations of distances at slower walking speeds varying from 5% to 15% and underestimations of distance by more than 10% during running activities (Multimedia Appendix 8) [87].

Discussion

Principal Findings

This review adds to the existing literature, as it is the first, to our knowledge, to systematically examine and report Fitbit device measurement accuracy in controlled and free-living settings for measures of step count, energy expenditure, sleep, time in activity, and distance in healthy adults or adults living with any health condition or disability.

Findings across many studies suggested that, approximately 50% of the time, Fitbit devices were likely to provide accurate measures (within $\pm 3\%$) of steps in controlled testing conditions, with an overall tendency to underestimate steps. Findings also indicated that step count accuracy was likely to improve if the device was worn on the torso during normal or self-paced walking activities, worn on the wrist during jogging activities, and worn on the ankle during slow or very slow walking activities. Findings from several studies examining step count in free-living settings also showed that, approximately 50% of the time, Fitbit devices were likely to provide relatively accurate (within $\pm 10\%$) measures of steps compared with research-grade accelerometers or pedometers when worn on the torso or wrist in healthy adults with no mobility limitations, with a tendency to overestimate steps in free-living settings.

Consistent findings across studies in controlled-testing settings indicated that Fitbit devices were also more likely to provide notable underestimations of step count during activities with very slow ambulation, particularly when worn on the torso, where body motion may be constrained by mobility limitations or walking while pushing a walker or stroller, and during activities that simulate household or sporting activities that involve stop-and-start ambulation throughout the task. Findings from a few studies in free-living conditions suggested that, compared with a research-grade accelerometer or pedometer worn at the ankle, a Fitbit device worn on the torso may

markedly overestimate steps in older adults with no mobility limitation and markedly underestimate steps in older adults with limited mobility.

There were also consistent findings from many studies examining energy expenditure in controlled settings, indicating that Fitbit devices were rarely likely to provide accurate measures of energy expenditure. Findings suggested that Fitbit was more likely to markedly overestimate energy expenditure when worn on the wrist and when walking at normal adult walking speeds on flat surfaces. On the contrary, Fitbit was more likely to underestimate energy expenditure when worn on the torso, with a tendency to markedly underestimate energy expenditure during inclined ambulation, during activities with constrained or variable body motion throughout the activity, and during simulated household or sporting activities that involve stop-and-start ambulation. Findings from 1 study for measures of energy expenditure in free-living settings suggested that Fitbit and doubly labelled water may provide similar measures of total energy expenditure over a 2-week period. However, findings from a few studies in free-living settings suggested that Fitbit devices may provide notable underestimations of daily energy expenditure compared with a SenseWear accelerometer.

A few studies examined Fitbit measurement accuracy for time spent in different intensity of activity in free-living settings. Across these studies, there was consistent evidence to suggest that, compared with research-grade accelerometers, Fitbit devices may underestimate sedentary time and progressively overestimate time in spent in activity as intensity of activity increases. Similarly, a few studies examined the accuracy for measures of sleep in controlled or free-living settings. Consistent evidence from these studies suggested that Fitbit may not provide accurate measures of sleep quality or quantity in a controlled-testing setting compared with polysomnography. However, there was some indication that Fitbit may provide relatively similar measures to SenseWear or Actiwatch accelerometers for time spent in bed and time sleeping in free-living settings. Finally, findings from 2 studies suggested that Fitbit may overestimate distance with slower walking speeds and progressively underestimate distance as walking speed increases.

Most of the studies included in this review were published in the last 2 years, with studies primarily examining measurement accuracy for models of Fitbit activity trackers introduced prior to 2015. The included studies mainly focused on step count and energy expenditure outcome measurement accuracy, with only a few of the studies examining measurement accuracy for sleep, distance, or time in activity. As well, the vast majority of studies included only healthy participants, with few including older adults, and fewer still including any adult living with disease or functional limitation. Overall, the quality of the included studies was excellent in terms of study design, reporting of missing data, and use of acceptable accuracy evaluations. However, some studies did not clearly identify how they may have handled missing data in their analyses, and few comprised more than 50 participants.

Most of the studies focused on measurement accuracy in controlled-testing environments comparing measurements from a Fitbit device against a reference-standard criterion. Standardized and controlled testing environments allow for evaluations of “true” measurement accuracy but do not necessarily reflect device measurement accuracy in uncontrolled or free-living settings, which are the intended environments for Fitbit activity tracker use. However, it is very difficult to measure true device accuracy in free-living conditions, as the reference-standard criterion measures generally cannot be used over a number of days while someone is conducting their usual daily activities. Therefore, the studies examining Fitbit device measurement accuracy in free-living conditions examined the accuracy of Fitbit device measures relative to an established research-grade criterion device measure of the same outcome when worn at the same time in free-living conditions.

For the purposes of this review, we defined satisfactory levels of measurement accuracy based on previously published standards for acceptable accuracy of step count in controlled ($\pm 3\%$) and free-living ($\pm 10\%$) settings [19-22]. Given that we were not able to identify published standards for accuracy of other outcome measures, we applied these same cutoff points for acceptable limits of measurement accuracy for all outcomes. However, we provide details of our descriptive analyses in the supplemental summary tables and offer visual representations for error estimations in the figures to allow for independent assessment of alternative definitions for acceptable limits for measurement accuracy by Fitbit devices.

Limitations

Our review has some potential limitations. These include the decision to include only data that were published in peer-reviewed journals and to exclude non-English studies. These decisions may have introduced a level of bias in our analyses and interpretation. In addition, we included all studies, independent of potential risk of bias. Moreover, the descriptive analyses and subsequent point estimations for percentage measurement error (ie, potential bias) gave equal weighting to accuracy comparisons with different sample sizes and variations in significance levels, which may misrepresent the true point estimate for measurement error for some of the testing conditions examined in this review [17,18]. Allowing for these potential limitations, and the limited number of studies examining Fitbit measurement accuracy for sleep, distance, and time spent in activity, we note that discretion should be exercised when considering our evaluations of the potential accuracy for these outcome domains. To address this gap in the literature, further high-quality research examining Fitbit measurement accuracy for sleep, time in activity, and distance is warranted.

We should identify as well that defining relative (in)accuracy of a Fitbit device in free-living settings does not define true measurement (in)accuracy, as neither the Fitbit device nor the reference device was compared to a reference-standard criterion. Rather, relative inaccuracy of a Fitbit defines only the likelihood that a Fitbit device will provide different values for measures of the same outcome when compared with a research-grade criterion in free-living conditions.

It is also important to clarify that we derived estimates of Fitbit device measurement accuracy in this review from studies that used different models of Fitbit, which might have different versions of firmware, software, and data processing algorithms. Since the design details for the devices and software are proprietary information, we were not able to determine whether and what modifications have been made by the company over time. Nonetheless, we indirectly explored the potential effect of differences in model design over time by using body placement for the device as a proxy, as the earlier models (eg, Classic, One, Zip, and Ultra) were worn on the torso, whereas the later models (eg, Flex, Charge, and Surge) were worn on the wrist. Therefore, some of the variability in error estimations with different body placement may be related in part to differences in device design or in analysis protocols over time.

Finally, our finding of potential limitations in Fitbit device measurement accuracy in a variety of testing conditions does not imply that Fitbit device measurement accuracy will remain static. Rather, it is very likely that accuracy will improve as technological advances in the firmware are implemented. As well, given the ability for Fitbit to tap into metadata from millions of users worldwide and apply advanced algorithms to better identify complex patterns of motion, it is likely that

evolving software upgrades will also lead to improved measurement accuracy. Furthermore, our findings do not negate the value of using Fitbit activity trackers in the manner for which the devices were intended, which is for self-monitoring of physical activity patterns and motivating individuals to achieve their physical activity goals [88-91].

Conclusion

Fitbit devices are most likely to provide accurate measures of steps in adults with no mobility limitations, when the device is worn on the torso while walking at normal or self-paced walking speeds. However, Fitbit devices are unlikely to provide accurate measures of energy expenditure. Limited evidence suggests that Fitbit activity trackers may not provide accurate measures for sleep, distance, or time spent in activity; however, further accuracy studies are warranted.

Implications

Other than for measures of steps in adults with no limitations in mobility, discretion should be used when considering the use of Fitbit devices as an outcome measurement tool in research or to inform health care decisions, as there are seemingly a limited number of situations where the device is likely to provide accurate measurement.

Acknowledgments

The authors would like to thank Dr Hui Xie, the Maureen and Milan Ilich and Merck Chair in Statistics for Arthritis and Musculoskeletal Diseases at the Arthritis Research Canada, Richmond, and the Faculty of Health Sciences, Simon Fraser University, Burnaby, BC, Canada, for his discussion and suggestions about the descriptive analyses and narrative syntheses of quantitative data conducted in this review.

This study was funded by “PRECISION: Preventing Complications from Inflammatory Skin, Joint and Bowel Conditions,” a Team Grant from the Canadian Institutes of Health Research, Canada (THC-316595).

Conflicts of Interest

None declared.

Multimedia Appendix 1

Search strategy example.

[\[PDF File \(Adobe PDF File\), 83KB-Multimedia Appendix 1\]](#)

Multimedia Appendix 2

Data extraction framework.

[\[PDF File \(Adobe PDF File\), 58KB-Multimedia Appendix 2\]](#)

Multimedia Appendix 3

Master data – percentage error values (all accuracy comparisons).

[\[PDF File \(Adobe PDF File\), 72KB-Multimedia Appendix 3\]](#)

Multimedia Appendix 4

Modified Consensus-Based Standards for the Selection of Health Status Measurement Instruments (COSMIN) criteria.

[\[PDF File \(Adobe PDF File\), 117KB-Multimedia Appendix 4\]](#)

Multimedia Appendix 5

Data coding (steps and energy expenditure in controlled settings).

[\[PDF File \(Adobe PDF File\), 83KB-Multimedia Appendix 5\]](#)

Multimedia Appendix 6

Study characteristics.

[\[PDF File \(Adobe PDF File\), 242KB-Multimedia Appendix 6\]](#)

Multimedia Appendix 7

Accuracy evaluations reported and risk-of-bias assessment.

[\[PDF File \(Adobe PDF File\), 179KB-Multimedia Appendix 7\]](#)

Multimedia Appendix 8

Controlled settings: accuracy – measurement error.

[\[PDF File \(Adobe PDF File\), 164KB-Multimedia Appendix 8\]](#)

Multimedia Appendix 9

Free-living settings: relative accuracy – measurement error.

[\[PDF File \(Adobe PDF File\), 131KB-Multimedia Appendix 9\]](#)

References

1. Wearable tech: leveraging canadian innovation to improve health. 2014. MaRS Mark Insights URL: <https://www.marsdd.com/wp-content/uploads/2015/02/MaRSReport-WearableTech.pdf> [accessed 2018-03-14] [WebCite Cache ID 6xuxgfaM0]
2. Fitbit Reports \$571M Q4'17 and \$1.616B FY'17 Revenue.: fitbit.com; 2018 Feb 26. URL: <https://investor.fitbit.com/press/press-releases/press-release-details/2018/Fitbit-Reports-571M-Q417-and-1616B-FY17-Revenue/default.aspx> [accessed 2018-08-01] [WebCite Cache ID 6xuzBvDvV]
3. Bunn JA, Navalta JW, Fountaine CJ, Reece JD. Current state of commercial wearable technology in physical activity monitoring 2015-2017. *Int J Exerc Sci* 2018;11(7):503-515 [FREE Full text] [Medline: 29541338]
4. Mercer K, Li M, Giangregorio L, Burns C, Grindrod K. Behavior change techniques present in wearable activity trackers: a critical analysis. *JMIR Mhealth Uhealth* 2016 Apr 27;4(2):e40 [FREE Full text] [doi: 10.2196/mhealth.4461] [Medline: 27122452]
5. Mishra A, Nieto A, Kitsiou S. Systematic review of mHealth interventions involving Fitbit activity tracking devices. 2017 Presented at: IEEE International Conference on Healthcare Informatics; Aug 23-26, 2017; Park City, UT, USA. [doi: 10.1109/ICHI.2017.42]
6. Noah B, Keller MS, Mosadeghi S, Stein L, Johl S, Delshad S, et al. Impact of remote patient monitoring on clinical outcomes: an updated meta-analysis of randomized controlled trials. *npj Digit Med* 2018 Jan 15;1(1). [doi: 10.1038/s41746-017-0002-4]
7. Phillips SM, Cadmus-Bertram L, Rosenberg D, Buman MP, Lynch BM. Wearable technology and physical activity in chronic disease: opportunities and challenges. *Am J Prev Med* 2018 Jan;54(1):144-150. [doi: 10.1016/j.amepre.2017.08.015] [Medline: 29122356]
8. US National Library of Medicine. ClinicalTrials.gov. URL: <https://clinicaltrials.gov/> [accessed 2018-08-01] [WebCite Cache ID 71M4vKwW8]
9. Vooijs M, Alpay LL, Snoeck-Stroband JB, Beerthuizen T, Siemonsma PC, Abbink JJ, et al. Validity and usability of low-cost accelerometers for internet-based self-monitoring of physical activity in patients with chronic obstructive pulmonary disease. *Interact J Med Res* 2014;3(4):e14 [FREE Full text] [doi: 10.2196/ijmr.3056] [Medline: 25347989]
10. Rosenberg D, Kadokura EA, Bouldin ED, Miyawaki CE, Higano CS, Hartzler AL. Acceptability of Fitbit for physical activity tracking within clinical care among men with prostate cancer. *AMIA Annu Symp Proc* 2016;2016:1050-1059 [FREE Full text] [Medline: 28269902]
11. Evenson KR, Goto MM, Furberg RD. Systematic review of the validity and reliability of consumer-wearable activity trackers. *Int J Behav Nutr Phys Act* 2015 Dec 18;12:159 [FREE Full text] [doi: 10.1186/s12966-015-0314-1] [Medline: 26684758]
12. Wong CK, Mentis HM, Kuber R. The bit doesn't fit: Evaluation of a commercial activity-tracker at slower walking speeds. *Gait Posture* 2018 Jan;59:177-181. [doi: 10.1016/j.gaitpost.2017.10.010] [Medline: 29049964]

13. O'Connell S, ÓLaighin G, Quinlan LR. When a step is not a step! Specificity analysis of five physical activity monitors. *PLoS One* 2017;12(1):e0169616 [FREE Full text] [doi: [10.1371/journal.pone.0169616](https://doi.org/10.1371/journal.pone.0169616)] [Medline: [28085918](https://pubmed.ncbi.nlm.nih.gov/28085918/)]
14. Dixon PM, Saint-Maurice PF, Kim Y, Hibbing P, Bai Y, Welk GJ. A primer on the use of equivalence testing for evaluating measurement agreement. *Med Sci Sports Exerc* 2018 Apr;50(4):837-845. [doi: [10.1249/MSS.0000000000001481](https://doi.org/10.1249/MSS.0000000000001481)] [Medline: [29135817](https://pubmed.ncbi.nlm.nih.gov/29135817/)]
15. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, et al. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. *Qual Life Res* 2010 May;19(4):539-549 [FREE Full text] [doi: [10.1007/s11136-010-9606-8](https://doi.org/10.1007/s11136-010-9606-8)] [Medline: [20169472](https://pubmed.ncbi.nlm.nih.gov/20169472/)]
16. Terwee CB, Mokkink LB, Knol DL, Ostelo RWJG, Bouter LM, de Vet HCW. Rating the methodological quality in systematic reviews of studies on measurement properties: a scoring system for the COSMIN checklist. *Qual Life Res* 2012 May;21(4):651-657 [FREE Full text] [doi: [10.1007/s11136-011-9960-1](https://doi.org/10.1007/s11136-011-9960-1)] [Medline: [21732199](https://pubmed.ncbi.nlm.nih.gov/21732199/)]
17. Popay J, Roberts H, Sowden A, Petticrew M, Arai L, Rodgers M, et al. Guidance on the Conduct of Narrative Synthesis in Systematic Reviews. A Product From the ESRC Methods Programme. Version 1: April 2006. Lancaster, UK: Lancaster University; Apr 2006.
18. Campbell M, Katikireddi S, Sowden A, McKenzie J, Thomson H. Improving Conduct and Reporting of Narrative Synthesis of Quantitative Data (ICONS-Quant): protocol for a mixed methods study to develop a reporting guideline. *BMJ Open* 2018 Feb;8(2):e020064 [FREE Full text]
19. Bassett DR, Mahar MT, Rowe DA, Morrow JR. Walking and measurement. *Med Sci Sports Exerc* 2008 Jul;40(7 Suppl):S529-S536. [doi: [10.1249/MSS.0b013e31817c699c](https://doi.org/10.1249/MSS.0b013e31817c699c)] [Medline: [18562970](https://pubmed.ncbi.nlm.nih.gov/18562970/)]
20. Schneider PL, Crouter SE, Lukajic O, Bassett DR. Accuracy and reliability of 10 pedometers for measuring steps over a 400-m walk. *Med Sci Sports Exerc* 2003 Oct;35(10):1779-1784. [doi: [10.1249/01.MSS.0000089342.96098.C4](https://doi.org/10.1249/01.MSS.0000089342.96098.C4)] [Medline: [14523320](https://pubmed.ncbi.nlm.nih.gov/14523320/)]
21. Schneider PL, Crouter SE, Bassett DR. Pedometer measures of free-living physical activity: comparison of 13 models. *Med Sci Sports Exerc* 2004 Feb;36(2):331-335. [doi: [10.1249/01.MSS.0000113486.60548.E9](https://doi.org/10.1249/01.MSS.0000113486.60548.E9)] [Medline: [14767259](https://pubmed.ncbi.nlm.nih.gov/14767259/)]
22. Tudor-Locke C, Sisson SB, Lee SM, Craig CL, Plotnikoff RC, Bauman A. Evaluation of quality of commercial pedometers. *Can J Public Health* 2006;97 Suppl 1:S10-5, S10. [Medline: [16676833](https://pubmed.ncbi.nlm.nih.gov/16676833/)]
23. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Int J Surg* 2010;8(5):336-341 [FREE Full text] [doi: [10.1016/j.ijisu.2010.02.007](https://doi.org/10.1016/j.ijisu.2010.02.007)] [Medline: [20171303](https://pubmed.ncbi.nlm.nih.gov/20171303/)]
24. Balto JM, Kinnett-Hopkins DL, Motl RW. Accuracy and precision of smartphone applications and commercially available motion sensors in multiple sclerosis. *Mult Scler J Exp Transl Clin* 2016;2:2055217316634754 [FREE Full text] [doi: [10.1177/2055217316634754](https://doi.org/10.1177/2055217316634754)] [Medline: [28607720](https://pubmed.ncbi.nlm.nih.gov/28607720/)]
25. Battenberg AK, Donohoe S, Robertson N, Schmalzried TP. The accuracy of personal activity monitoring devices. *Semin Arthroplasty* 2017 Jun;28(2):71-75. [doi: [10.1053/j.sart.2017.07.006](https://doi.org/10.1053/j.sart.2017.07.006)]
26. Beevi FHA, Miranda J, Pedersen CF, Wagner S. An evaluation of commercial pedometers for monitoring slow walking speed populations. *Telemed J E Health* 2016 May;22(5):441-449. [doi: [10.1089/tmj.2015.0120](https://doi.org/10.1089/tmj.2015.0120)] [Medline: [26451900](https://pubmed.ncbi.nlm.nih.gov/26451900/)]
27. Chen M, Kuo C, Pellegrini CA, Hsu M. Accuracy of wristband activity monitors during ambulation and activities. *Med Sci Sports Exerc* 2016 Dec;48(10):1942-1949. [doi: [10.1249/MSS.0000000000000984](https://doi.org/10.1249/MSS.0000000000000984)] [Medline: [27183123](https://pubmed.ncbi.nlm.nih.gov/27183123/)]
28. Chow JJ, Thom JM, Wewege MA, Ward RE, Parmenter BJ. Accuracy of step count measured by physical activity monitors: the effect of gait speed and anatomical placement site. *Gait Posture* 2017 Sep;57:199-203. [doi: [10.1016/j.gaitpost.2017.06.012](https://doi.org/10.1016/j.gaitpost.2017.06.012)] [Medline: [28666177](https://pubmed.ncbi.nlm.nih.gov/28666177/)]
29. Diaz KM, Krupka DJ, Chang MJ, Shaffer JA, Ma Y, Goldsmith J, et al. Validation of the Fitbit One® for physical activity measurement at an upper torso attachment site. *BMC Res Notes* 2016 Apr 12;9:213 [FREE Full text] [doi: [10.1186/s13104-016-2020-8](https://doi.org/10.1186/s13104-016-2020-8)] [Medline: [27068022](https://pubmed.ncbi.nlm.nih.gov/27068022/)]
30. Floegel TA, Florez-Pregonero A, Hekler EB, Buman MP. Validation of consumer-based hip and wrist activity monitors in older adults with varied ambulatory abilities. *J Gerontol A Biol Sci Med Sci* 2017 Feb;72(2):229-236. [doi: [10.1093/gerona/glw098](https://doi.org/10.1093/gerona/glw098)] [Medline: [27257217](https://pubmed.ncbi.nlm.nih.gov/27257217/)]
31. Fokkema T, Kooiman TJM, Krijnen WP. Reliability and validity of ten consumer activity trackers depend on walking speed. *Med Sci Sports Exerc* 2017 Apr;49(4):793-800. [doi: [10.1249/MSS.0000000000001146](https://doi.org/10.1249/MSS.0000000000001146)] [Medline: [28319983](https://pubmed.ncbi.nlm.nih.gov/28319983/)]
32. Fulk GD, Combs SA, Danks KA, Nirider CD, Raja B, Reisman DS. Accuracy of 2 activity monitors in detecting steps in people with stroke and traumatic brain injury. *Phys Ther* 2014 Feb;94(2):222-229 [FREE Full text] [doi: [10.2522/ptj.20120525](https://doi.org/10.2522/ptj.20120525)] [Medline: [24052577](https://pubmed.ncbi.nlm.nih.gov/24052577/)]
33. Huang Y, Xu J, Yu B, Shull PB. Validity of FitBit, Jawbone UP, Nike+ and other wearable devices for level and stair walking. *Gait Posture* 2016 Dec;48:36-41. [doi: [10.1016/j.gaitpost.2016.04.025](https://doi.org/10.1016/j.gaitpost.2016.04.025)] [Medline: [27477705](https://pubmed.ncbi.nlm.nih.gov/27477705/)]
34. Imboden MT, Nelson MB, Kaminsky LA, Montoye AH. Comparison of four Fitbit and Jawbone activity monitors with a research-grade ActiGraph accelerometer for estimating physical activity and energy expenditure. *Br J Sports Med* 2017 May 08. [doi: [10.1136/bjsports-2016-096990](https://doi.org/10.1136/bjsports-2016-096990)] [Medline: [28483930](https://pubmed.ncbi.nlm.nih.gov/28483930/)]

35. Klassen TD, Simpson LA, Lim SB, Louie DR, Parappilly B, Sakakibara BM, et al. "Stepping up" activity poststroke: ankle-positioned accelerometer can accurately record steps during slow walking. *Phys Ther* 2016 Mar;96(3):355-360 [FREE Full text] [doi: [10.2522/ptj.20140611](https://doi.org/10.2522/ptj.20140611)] [Medline: [26251478](https://pubmed.ncbi.nlm.nih.gov/26251478/)]
36. Kooiman TJM, Dontje ML, Sprenger SR, Krijnen WP, van der Schans CP, de Groot M. Reliability and validity of ten consumer activity trackers. *BMC Sports Sci Med Rehabil* 2015;7:24 [FREE Full text] [doi: [10.1186/s13102-015-0018-5](https://doi.org/10.1186/s13102-015-0018-5)] [Medline: [26464801](https://pubmed.ncbi.nlm.nih.gov/26464801/)]
37. Nelson MB, Kaminsky LA, Dickin DC, Montoye AHK. Validity of consumer-based physical activity monitors for specific activity types. *Med Sci Sports Exerc* 2016 Aug;48(8):1619-1628. [doi: [10.1249/MSS.0000000000000933](https://doi.org/10.1249/MSS.0000000000000933)] [Medline: [27015387](https://pubmed.ncbi.nlm.nih.gov/27015387/)]
38. Modave F, Guo Y, Bian J, Gurka MJ, Parish A, Smith MD, et al. Mobile device accuracy for step counting across age groups. *JMIR Mhealth Uhealth* 2017 Jun 28;5(6):e88 [FREE Full text] [doi: [10.2196/mhealth.7870](https://doi.org/10.2196/mhealth.7870)] [Medline: [28659255](https://pubmed.ncbi.nlm.nih.gov/28659255/)]
39. Montes J, Young JC, Tandy R, Navalta JW. Fitbit Flex: energy expenditure and step count evaluation. *J Exerc Physiol Online* 2017 Oct;20(5):134-140 [FREE Full text]
40. O'Connell S, ÓLaighin G, Kelly L, Murphy E, Beirne S, Burke N, et al. These shoes are made for walking: sensitivity performance evaluation of commercial activity monitors under the expected conditions and circumstances required to achieve the international daily step goal of 10,000 steps. *PLoS One* 2016;11(5):e0154956 [FREE Full text] [doi: [10.1371/journal.pone.0154956](https://doi.org/10.1371/journal.pone.0154956)] [Medline: [27167121](https://pubmed.ncbi.nlm.nih.gov/27167121/)]
41. Park W, Lee VJ, Ku B, Tanaka H. Effect of walking speed and placement position interactions in determining the accuracy of various newer pedometers. *J Exerc Sci Fit* 2014 Jun;12(1):31-37. [doi: [10.1016/j.jesf.2014.01.003](https://doi.org/10.1016/j.jesf.2014.01.003)]
42. Paul SS, Tiedemann A, Hassett LM, Ramsay E, Kirkham C, Chagpar S, et al. Validity of the Fitbit activity tracker for measuring steps in community-dwelling older adults. *BMJ Open Sport Exerc Med* 2015;1(1):e000013 [FREE Full text] [doi: [10.1136/bmjsem-2015-000013](https://doi.org/10.1136/bmjsem-2015-000013)] [Medline: [27900119](https://pubmed.ncbi.nlm.nih.gov/27900119/)]
43. Phillips LJ, Petroski GF, Markis NE. A comparison of accelerometer accuracy in older adults. *Res Gerontol Nurs* 2015;8(5):213-219. [doi: [10.3928/19404921-20150429-03](https://doi.org/10.3928/19404921-20150429-03)] [Medline: [25942386](https://pubmed.ncbi.nlm.nih.gov/25942386/)]
44. Schaffer SD, Holzapfel SD, Fulk G, Bosch PR. Step count accuracy and reliability of two activity tracking devices in people after stroke. *Physiother Theory Pract* 2017 Oct;33(10):788-796. [doi: [10.1080/09593985.2017.1354412](https://doi.org/10.1080/09593985.2017.1354412)] [Medline: [28777710](https://pubmed.ncbi.nlm.nih.gov/28777710/)]
45. Simpson LA, Eng JJ, Klassen TD, Lim SB, Louie DR, Parappilly B, et al. Capturing step counts at slow walking speeds in older adults: comparison of ankle and waist placement of measuring device. *J Rehabil Med* 2015 Oct 05;47(9):830-835 [FREE Full text] [doi: [10.2340/16501977-1993](https://doi.org/10.2340/16501977-1993)] [Medline: [26181670](https://pubmed.ncbi.nlm.nih.gov/26181670/)]
46. Stackpool CM, Porcari JP, Mikat RP, Gillette C, Foster C. The accuracy of various activity trackers in estimating steps taken and energy expenditure. *J Fit Res* 2014 Dec;3(3):32-48.
47. Sushames A, Edwards A, Thompson F, McDermott R, Gebel K. Validity and reliability of fitbit flex for step count, moderate to vigorous physical activity and activity energy expenditure. *PLoS One* 2016;11(9):e0161224 [FREE Full text] [doi: [10.1371/journal.pone.0161224](https://doi.org/10.1371/journal.pone.0161224)] [Medline: [27589592](https://pubmed.ncbi.nlm.nih.gov/27589592/)]
48. Takacs J, Pollock CL, Guenther JR, Bahar M, Napier C, Hunt MA. Validation of the Fitbit One activity monitor device during treadmill walking. *J Sci Med Sport* 2014 Sep;17(5):496-500. [doi: [10.1016/j.jsams.2013.10.241](https://doi.org/10.1016/j.jsams.2013.10.241)] [Medline: [24268570](https://pubmed.ncbi.nlm.nih.gov/24268570/)]
49. Treacy D, Hassett L, Schurr K, Chagpar S, Paul SS, Sherrington C. Validity of different activity monitors to count steps in an inpatient rehabilitation setting. *Phys Ther* 2017 May 01;97(5):581-588. [doi: [10.1093/ptj/pzx010](https://doi.org/10.1093/ptj/pzx010)] [Medline: [28339904](https://pubmed.ncbi.nlm.nih.gov/28339904/)]
50. Steffen TM, Hacker TA, Mollinger L. Age- and gender-related test performance in community-dwelling elderly people: Six-Minute Walk Test, Berg Balance Scale, Timed Up & Go Test, and gait speeds. *Phys Ther* 2002 Feb;82(2):128-137. [Medline: [11856064](https://pubmed.ncbi.nlm.nih.gov/11856064/)]
51. Fritz S, Lusardi M. White paper: "walking speed: the sixth vital sign". *J Geriatr Phys Ther* 2009;32(2):46-49. [Medline: [20039582](https://pubmed.ncbi.nlm.nih.gov/20039582/)]
52. Alharbi M, Bauman A, Neubeck L, Gallagher R. Validation of Fitbit-Flex as a measure of free-living physical activity in a community-based phase III cardiac rehabilitation population. *Eur J Prev Cardiol* 2016 Sep;23(14):1476-1485. [doi: [10.1177/2047487316634883](https://doi.org/10.1177/2047487316634883)] [Medline: [26907794](https://pubmed.ncbi.nlm.nih.gov/26907794/)]
53. Brewer W, Swanson BT, Ortiz A. Validity of Fitbit's active minutes as compared with a research-grade accelerometer and self-reported measures. *BMJ Open Sport Exerc Med* 2017;3(1):e000254 [FREE Full text] [doi: [10.1136/bmjsem-2017-000254](https://doi.org/10.1136/bmjsem-2017-000254)] [Medline: [29018543](https://pubmed.ncbi.nlm.nih.gov/29018543/)]
54. Dominick GM, Winfree KN, Pohlig RT, Papas MA. Physical activity assessment between consumer- and research-grade accelerometers: a comparative study in free-living conditions. *JMIR Mhealth Uhealth* 2016 Sep 19;4(3):e110 [FREE Full text] [doi: [10.2196/mhealth.6281](https://doi.org/10.2196/mhealth.6281)] [Medline: [27644334](https://pubmed.ncbi.nlm.nih.gov/27644334/)]
55. Farina N, Lowry RG. The validity of consumer-level activity monitors in healthy older adults in free-living conditions. *J Aging Phys Act* 2018 Jan 01;26(1):128-135. [doi: [10.1123/japa.2016-0344](https://doi.org/10.1123/japa.2016-0344)] [Medline: [28595019](https://pubmed.ncbi.nlm.nih.gov/28595019/)]
56. Ferguson T, Rowlands AV, Olds T, Maher C. The validity of consumer-level, activity monitors in healthy adults worn in free-living conditions: a cross-sectional study. *Int J Behav Nutr Phys Act* 2015;12:42 [FREE Full text] [doi: [10.1186/s12966-015-0201-9](https://doi.org/10.1186/s12966-015-0201-9)] [Medline: [25890168](https://pubmed.ncbi.nlm.nih.gov/25890168/)]

57. Gomersall SR, Ng N, Burton NW, Pavey TG, Gilson ND, Brown WJ. Estimating physical activity and sedentary behavior in a free-living context: a pragmatic comparison of consumer-based activity trackers and ActiGraph accelerometry. *J Med Internet Res* 2016 Sep 07;18(9):e239. [Medline: [27604226](#)]
58. Hargens TA, Deyarmin KN, Snyder KM, Mihalik AG, Sharpe LE. Comparison of wrist-worn and hip-worn activity monitors under free living conditions. *J Med Eng Technol* 2017 Apr;41(3):200-207. [doi: [10.1080/03091902.2016.1271046](#)] [Medline: [28078908](#)]
59. Hui J, Heyden R, Bao T, Accettone N, McBay C, Richardson J, et al. Validity of the Fitbit One for measuring activity in community-dwelling stroke survivors. *Physiother Can* 2018;70(1):81-89. [doi: [10.3138/ptc.2016-40.ep](#)] [Medline: [29434422](#)]
60. Middelweerd A, Van Der Ploeg HP, Van Halteren A, Twisk JWR, Brug J, Te Velde SJ. A validation study of the Fitbit One in daily life using different time intervals. *Med Sci Sports Exerc* 2017 Jun;49(6):1270-1279. [doi: [10.1249/MSS.0000000000001225](#)] [Medline: [28511193](#)]
61. Reid RER, Insogna JA, Carver TE, Comptour AM, Bewski NA, Sciortino C, et al. Validity and reliability of Fitbit activity monitors compared to ActiGraph GT3X+ with female adults in a free-living environment. *J Sci Med Sport* 2017 Jun;20(6):578-582. [doi: [10.1016/j.jsams.2016.10.015](#)] [Medline: [27887786](#)]
62. Thorup CB, Andreasen JJ, Sørensen EE, Grønkjær M, Dinesen BI, Hansen J. Accuracy of a step counter during treadmill and daily life walking by healthy adults and patients with cardiac disease. *BMJ Open* 2017 Dec 31;7(3):e011742 [FREE Full text] [doi: [10.1136/bmjopen-2016-011742](#)] [Medline: [28363918](#)]
63. Van Blarigan EL, Kenfield SA, Tantum L, Cadmus-Bertram LA, Carroll PR, Chan JM. The Fitbit One physical activity tracker in men with prostate cancer: validation study. *JMIR Cancer* 2017 Apr 18;3(1):e5 [FREE Full text] [doi: [10.2196/cancer.6935](#)] [Medline: [28420602](#)]
64. An HS, Jones GC, Kang SK, Welk GJ, Lee JM. How valid are wearable physical activity trackers for measuring steps? *Eur J Sport Sci* 2017 Apr;17(3):360-368. [doi: [10.1080/17461391.2016.1255261](#)] [Medline: [27912681](#)]
65. Chu AHY, Ng SHX, Paknezhad M, Gauterin A, Koh D, Brown MS, et al. Comparison of wrist-worn Fitbit Flex and waist-worn ActiGraph for measuring steps in free-living adults. *PLoS One* 2017;12(2):e0172535 [FREE Full text] [doi: [10.1371/journal.pone.0172535](#)] [Medline: [28234953](#)]
66. Rosenberger ME, Buman MP, Haskell WL, McConnell MV, Carstensen LL. 24 hours of sleep, sedentary behavior, and physical activity with nine wearable devices. *Med Sci Sports Exerc* 2016 Mar;48(3):457-465. [doi: [10.1249/MSS.0000000000000778](#)] [Medline: [26484953](#)]
67. Tully MA, McBride C, Heron L, Hunter RF. The validation of Fitbit Zip™ physical activity monitor as a measure of free-living physical activity. *BMC Res Notes* 2014;7:952 [FREE Full text] [doi: [10.1186/1756-0500-7-952](#)] [Medline: [25539733](#)]
68. Adam NJ, Spierer DK, Gu J, Bronner S. Comparison of steps and energy expenditure assessment in adults of Fitbit Tracker and Ultra to the Actical and indirect calorimetry. *J Med Eng Technol* 2013 Oct;37(7):456-462. [doi: [10.3109/03091902.2013.831135](#)] [Medline: [24007317](#)]
69. Bai Y, Welk GJ, Nam YH, Lee JA, Lee J, Kim Y, et al. Comparison of consumer and research monitors under semistructured settings. *Med Sci Sports Exerc* 2016 Jan;48(1):151-158. [doi: [10.1249/MSS.0000000000000727](#)] [Medline: [26154336](#)]
70. Chowdhury EA, Western MJ, Nightingale TE, Peacock OJ, Thompson D. Assessment of laboratory and daily energy expenditure estimates from consumer multi-sensor physical activity monitors. *PLoS One* 2017;12(2):e0171720 [FREE Full text] [doi: [10.1371/journal.pone.0171720](#)] [Medline: [28234979](#)]
71. Dannecker KL, Sazonova NA, Melanson EL, Sazonov ES, Browning RC. A comparison of energy expenditure estimation of several physical activity monitors. *Med Sci Sports Exerc* 2013 Nov;45(11):2105-2112 [FREE Full text] [doi: [10.1249/MSS.0b013e318299d2eb](#)] [Medline: [23669877](#)]
72. Dondzila C, Garner D. Comparative accuracy of fitness tracking modalities in quantifying energy expenditure. *J Med Eng Technol* 2016 Aug;40(6):325-329. [doi: [10.1080/03091902.2016.1197978](#)] [Medline: [27280592](#)]
73. Dooley EE, Golaszewski NM, Bartholomew JB. Estimating accuracy at exercise intensities: a comparative study of self-monitoring heart rate and physical activity wearable devices. *JMIR Mhealth Uhealth* 2017 Mar 16;5(3):e34 [FREE Full text] [doi: [10.2196/mhealth.7043](#)] [Medline: [28302596](#)]
74. Gusmer RJ, Bosch TA, Watkins AN, Ostrem JD, Dengel DR. Comparison of FitBit® Ultra to ActiGraph™ GT1M for assessment of physical activity in young adults during treadmill walking. *Open Sports Med J* 2014 Apr 04;8(1):11-15. [doi: [10.2174/1874387001408010011](#)]
75. Lee J, Kim Y, Welk GJ. Validity of consumer-based physical activity monitors. *Med Sci Sports Exerc* 2014 Sep;46(9):1840-1848. [doi: [10.1249/MSS.0000000000000287](#)] [Medline: [24777201](#)]
76. Montoye AHK, Mitrzyk JR, Molesky MJ. Comparative accuracy of a wrist-worn activity tracker and a smart shirt for physical activity assessment. *Meas Phys Educ Exerc Sci* 2017 Jun 08;21(4):201-211. [doi: [10.1080/1091367X.2017.1331166](#)]
77. Murakami H, Kawakami R, Nakae S, Nakata Y, Ishikawa-Takata K, Tanaka S, et al. Accuracy of wearable devices for estimating total energy expenditure: comparison with metabolic chamber and doubly labeled water method. *JAMA Intern Med* 2016 Dec 01;176(5):702-703. [doi: [10.1001/jamainternmed.2016.0152](#)] [Medline: [26999758](#)]

78. Price K, Bird SR, Lythgo N, Raj IS, Wong JYL, Lynch C. Validation of the Fitbit One, Garmin Vivofit and Jawbone UP activity tracker in estimation of energy expenditure during treadmill walking and running. *J Med Eng Technol* 2017 Apr;41(3):208-215. [doi: [10.1080/03091902.2016.1253795](https://doi.org/10.1080/03091902.2016.1253795)] [Medline: [27919170](https://pubmed.ncbi.nlm.nih.gov/27919170/)]
79. Sasaki JE, Hickey A, Mavilia M, Tedesco J, John D, Kozey KS, et al. Validation of the Fitbit wireless activity tracker for prediction of energy expenditure. *J Phys Act Health* 2015 Feb;12(2):149-154. [doi: [10.1123/jpah.2012-0495](https://doi.org/10.1123/jpah.2012-0495)] [Medline: [24770438](https://pubmed.ncbi.nlm.nih.gov/24770438/)]
80. Wallen MP, Gomersall SR, Keating SE, Wisløff U, Coombes JS. Accuracy of heart rate watches: implications for weight management. *PLoS One* 2016;11(5):e0154420 [FREE Full text] [doi: [10.1371/journal.pone.0154420](https://doi.org/10.1371/journal.pone.0154420)] [Medline: [27232714](https://pubmed.ncbi.nlm.nih.gov/27232714/)]
81. Brooke SM, An H, Kang S, Noble JM, Berg KE, Lee J. Concurrent validity of wearable activity trackers under free-living conditions. *J Strength Cond Res* 2017 Apr;31(4):1097-1106. [doi: [10.1519/JSC.0000000000001571](https://doi.org/10.1519/JSC.0000000000001571)] [Medline: [27465631](https://pubmed.ncbi.nlm.nih.gov/27465631/)]
82. Cook JD, Prairie ML, Plante DT. Utility of the Fitbit Flex to evaluate sleep in major depressive disorder: a comparison against polysomnography and wrist-worn actigraphy. *J Affect Disord* 2017 Aug 01;217:299-305. [doi: [10.1016/j.jad.2017.04.030](https://doi.org/10.1016/j.jad.2017.04.030)] [Medline: [28448949](https://pubmed.ncbi.nlm.nih.gov/28448949/)]
83. Mantua J, Gravel N, Spencer RMC. Reliability of sleep measures from four personal health monitoring devices compared to research-based actigraphy and polysomnography. *Sensors (Basel)* 2016 Dec 05;16(5) [FREE Full text] [doi: [10.3390/s16050646](https://doi.org/10.3390/s16050646)] [Medline: [27164110](https://pubmed.ncbi.nlm.nih.gov/27164110/)]
84. Montgomery-Downs HE, Insana SP, Bond JA. Movement toward a novel activity monitoring device. *Sleep Breath* 2012 Sep;16(3):913-917. [doi: [10.1007/s11325-011-0585-y](https://doi.org/10.1007/s11325-011-0585-y)] [Medline: [21971963](https://pubmed.ncbi.nlm.nih.gov/21971963/)]
85. Lee H, Lee H, Moon J, Lee T, Kim M, In H, et al. Comparison of wearable activity tracker with actigraphy for sleep evaluation and circadian rest-activity rhythm measurement in healthy young adults. *Psychiatry Investig* 2017 Mar;14(2):179-185 [FREE Full text] [doi: [10.4306/pi.2017.14.2.179](https://doi.org/10.4306/pi.2017.14.2.179)] [Medline: [28326116](https://pubmed.ncbi.nlm.nih.gov/28326116/)]
86. Dickinson DL, Cazier J, Cech T. A practical validation study of a commercial accelerometer using good and poor sleepers. *Health Psychol Open* 2016 Jul;3(2):2055102916679012 [FREE Full text] [doi: [10.1177/2055102916679012](https://doi.org/10.1177/2055102916679012)] [Medline: [28815052](https://pubmed.ncbi.nlm.nih.gov/28815052/)]
87. Wahl Y, Düking P, Droszez A, Wahl P, Mester J. Criterion-validity of commercially available physical activity tracker to estimate step count, covered distance and energy expenditure during sports conditions. *Front Physiol* 2017;8:725 [FREE Full text] [doi: [10.3389/fphys.2017.00725](https://doi.org/10.3389/fphys.2017.00725)] [Medline: [29018355](https://pubmed.ncbi.nlm.nih.gov/29018355/)]
88. Li LC, Sayre EC, Xie H, Falck RS, Best JR, Liu-Ambrose T, et al. Efficacy of a community-based technology-enabled physical activity counselling program for people with knee osteoarthritis: a proof-of-concept study. *J Med Internet Res* 2018 Apr 30;20(4):e159. [doi: [10.2196/jmir.8514](https://doi.org/10.2196/jmir.8514)] [Medline: [29712630](https://pubmed.ncbi.nlm.nih.gov/29712630/)]
89. Li LC, Feehan LM, Shaw C, Xie H, Sayre EC, Aviña-Zubeita A, et al. A technology-enabled counselling program versus a delayed treatment control to support physical activity participation in people with inflammatory arthritis: study protocol for the OPAM-IA randomized controlled trial. *BMC Rheumatol* 2017 Nov 28;1(1). [doi: [10.1186/s41927-017-0005-4](https://doi.org/10.1186/s41927-017-0005-4)]
90. Li LC, Sayre EC, Xie H, Clayton C, Feehan LM. A community-based physical activity counselling program for people with knee osteoarthritis: feasibility and preliminary efficacy of the Track-OA Study. *JMIR Mhealth Uhealth* 2017 Jun 26;5(6):e86 [FREE Full text] [doi: [10.2196/mhealth.7863](https://doi.org/10.2196/mhealth.7863)] [Medline: [28652228](https://pubmed.ncbi.nlm.nih.gov/28652228/)]
91. Henriksen A, Haugen MM, Woldaregay AZ, Muzny M, Hartvigsen G, Hopstock LA, et al. Using fitness trackers and smartwatches to measure physical activity in research: analysis of consumer wrist-worn wearables. *J Med Internet Res* 2018 Mar 22;20(3):e110 [FREE Full text] [doi: [10.2196/jmir.9157](https://doi.org/10.2196/jmir.9157)] [Medline: [29567635](https://pubmed.ncbi.nlm.nih.gov/29567635/)]

Abbreviations

COSMIN: Consensus-Based Standards for the Selection of Health Status Measurement Instruments

MAPE: mean or median absolute percentage error

Edited by G Eysenbach; submitted 07.04.18; peer-reviewed by E Lyons, K Diaz, A Henriksen, B Price; comments to author 23.04.18; revised version received 05.06.18; accepted 23.07.18; published 09.08.18

Please cite as:

Feehan LM, Geldman J, Sayre EC, Park C, Ezzat AM, Yoo JY, Hamilton CB, Li LC
Accuracy of Fitbit Devices: Systematic Review and Narrative Syntheses of Quantitative Data
JMIR Mhealth Uhealth 2018;6(8):e10527
URL: <http://mhealth.jmir.org/2018/8/e10527/>
doi: [10.2196/10527](https://doi.org/10.2196/10527)
PMID: [30093371](https://pubmed.ncbi.nlm.nih.gov/30093371/)

©Lynne M Feehan, Jasmina Geldman, Eric C Sayre, Chance Park, Allison M Ezzat, Ju Young Yoo, Clayon B Hamilton, Linda C Li. Originally published in JMIR Mhealth and Uhealth (<http://mhealth.jmir.org>), 09.08.2018. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mhealth and uhealth, is properly cited. The complete bibliographic information, a link to the original publication on <http://mhealth.jmir.org/>, as well as this copyright and license information must be included.