Original Paper

Validity Evaluation of the Fitbit Charge2 and the Garmin vivosmart HR+ in Free-Living Environments in an Older Adult Cohort

Salvatore Tedesco¹, MSc; Marco Sica¹, MSc; Andrea Ancillao¹, PhD; Suzanne Timmons², MD; John Barton¹, MEng; Brendan O'Flynn¹, PhD

Corresponding Author:

Salvatore Tedesco, MSc Tyndall National Institute, University College Cork Lee Maltings, Prospect Row Cork, T12R5CP Ireland

Phone: 353 212346286

Email: salvatore.tedesco@tyndall.ie

Abstract

Background: Few studies have investigated the validity of mainstream wrist-based activity trackers in healthy older adults in real life, as opposed to laboratory settings.

Objective: This study explored the performance of two wrist-worn trackers (Fitbit Charge 2 and Garmin vivosmart HR+) in estimating steps, energy expenditure, moderate-to-vigorous physical activity (MVPA) levels, and sleep parameters (total sleep time [TST] and wake after sleep onset [WASO]) against gold-standard technologies in a cohort of healthy older adults in a free-living environment.

Methods: Overall, 20 participants (>65 years) took part in the study. The devices were worn by the participants for 24 hours, and the results were compared against validated technology (ActiGraph and New-Lifestyles NL-2000i). Mean error, mean percentage error (MPE), mean absolute percentage error (MAPE), intraclass correlation (ICC), and Bland-Altman plots were computed for all the parameters considered.

Results: For step counting, all trackers were highly correlated with one another (ICCs>0.89). Although the Fitbit tended to overcount steps (MPE=12.36%), the Garmin and ActiGraph undercounted (MPE 9.36% and 11.53%, respectively). The Garmin had poor ICC values when energy expenditure was compared against the criterion. The Fitbit had moderate-to-good ICCs in comparison to the other activity trackers, and showed the best results (MAPE=12.25%), although it underestimated calories burned. For MVPA levels estimation, the wristband trackers were highly correlated (ICC=0.96); however, they were moderately correlated against the criterion and they overestimated MVPA activity minutes. For the sleep parameters, the ICCs were poor for all cases, except when comparing the Fitbit with the criterion, which showed moderate agreement. The TST was slightly overestimated with the Fitbit, although it provided good results with an average MAPE equal to 10.13%. Conversely, WASO estimation was poorer and was overestimated by the Fitbit but underestimated by the Garmin. Again, the Fitbit was the most accurate, with an average MAPE of 49.7%.

Conclusions: The tested well-known devices could be adopted to estimate steps, energy expenditure, and sleep duration with an acceptable level of accuracy in the population of interest, although clinicians should be cautious in considering other parameters for clinical and research purposes.

(JMIR Mhealth Uhealth 2019;7(6):e13084) doi: 10.2196/13084

KEYWORDS

aging; fitness trackers; wristbands; older adults; wearable activity trackers; Fitbit; Garmin; energy expenditure; physical activity; sleep



¹Tyndall National Institute, University College Cork, Cork, Ireland

²Centre for Gerontology and Rehabilitation, University College Cork, Cork, Ireland

Introduction

Fitness trackers are popular devices used by athletes and the general public to monitor their physical activity levels, sport performance, and even their general health status in real time, with the latter having the potential to also predict the person's future health status [1]. Compared to other body positions, the wrist has been identified as the most suitable location for enhancing user acceptability and the user-friendliness of the device [2]. Common consumer-level, wrist-worn devices typically provide data on step count, distance traveled, number of floors climbed, and minutes of physical activity, as well as sport-related activity recognition, physiological measurements, energy expenditure, and sleep patterns. This information can promote a healthier lifestyle or an optimal training program through user-friendly visual feedback of current status or performance compared to set targets.

Fitness trackers based on motion sensors are also being used for monitoring biomechanical quantities of clinical interest, such as gait analysis applications [3,4], indirect estimation of ground reaction forces, and posture in general [5,6].

Although the number of studies investigating the validity and reliability of different fitness trackers is growing, the majority of the evidence is limited to young and middle-aged adult populations, mostly in good health [7-9]. Considering the multiple applications of wrist-based technology and its potential adoption in health care, and with an aging population, it is important to investigate the use of these devices in different populations, such as older people [10]. Although older adults perceive commercial trackers as useful and acceptable [11,12], older person-specific activity trackers are still limited [13].

Few studies have investigated the validity of mainstream wrist-based activity trackers in healthy older adults [14,15]. However, such investigations mainly involved a protocol structured around a number of daily activities simulated or recreated in a laboratory environment. Studies that investigated fitness trackers' performance when used by older people in their home environment, where older adults can perform their real daily routine, are scarce and mainly limited to step-counting features [16]. This study reviewed the validity and reliability of consumer-grade activity trackers in older community-dwelling adults through seven observational studies, of which only five studied free-living settings for a monitoring period of between 3 and 7 days.

For example, Paul et al [17] reported that the average steps per day measured over 7 days in a community-dwelling older adult population with a Fitbit and an ActiGraph showed excellent agreement, with the ActiGraph undercounting steps compared against participants' physical activity logs.

In another study, a Fitbit Flex and an ActiGraph were worn by a cohort of cardiac patients and their family members to measure steps and moderate-to-vigorous physical activity (MVPA) levels for 4 days [18]. It showed a significant correlation for step counts but lower values for MVPA, with the Fitbit Flex slightly overestimating both parameters.

Boeselt et al [19] compared a Polar A300 with a BodyMedia SenseWear in a cohort of patients with chronic obstructive pulmonary disease (mean age 66.4 years). Participants used the devices for three consecutive days, measuring steps, calories burned, daily activity time, and metabolic equivalents. The study showed a high correlation for step count and calories burned.

Farina and Lowry [20] compared the accuracy of step counts from two consumer-level activity monitors (Misfit Shine on both wrist and waist, and Fitbit Charge HR on the wrist) against two waist-worn reference devices (ActiGraph GT3X+ and New Lifestyle NL2000i) in healthy, community-dwelling older adults in free-living conditions over seven consecutive days. All consumer-level activity monitors positively correlated with reference devices. Compared to the ActiGraph GT3X+, the waist-worn Misfit Shine displayed the highest agreement, whereas the wrist-worn devices showed poorer performances.

Finally, Burton et al [21] reported good reliability and validity for the Fitbit Flex and Fitbit Charge HR compared against a GENEactiv accelerometer in a free-living environment over 14 days. Step count, distance traveled, MVPA minutes, and sleep were measured. Good strength of agreement was found for total distance and steps (obtained with the fitness tracker) and the MVPA estimated by the GENEactiv.

It is evident that a comparative analysis of mainstream trackers worn by healthy older people in a more ecologically valid environment is needed. This study aims to investigate the reliability and accuracy of the wrist-based Fitbit Charge 2 and the Garmin vivosmart HR+ activity trackers in the estimation of daily step count, total calorie expenditure, MVPA, and sleep parameters within a home environment in a cohort of older adults.

Methods

Participants

This study was based on a sample of 20 healthy older people (9 males, 11 females). Volunteers were recruited via a general invitation email, posters, and word of mouth to exstaff at University College Cork (Cork, Ireland) and their relatives, and also through local social and voluntary groups that had older adults as members. They were informed of the study by the Centre for Gerontology and Rehabilitation in University College Cork.

For the cohort, the inclusion criteria were age 65 years and older, with no history of neurological or other disorders or disability that could affect the participant's movements, and in good general health. Before participation, volunteers received an oral and written explanation of the study protocol, and written consent was obtained. Sociodemographic information was collected on gender, age, weight, height, and dominant arm. The study received approval by the Clinical Research Ethics Committee at the University College Cork. Demographic information on the participants who completed the study protocol is presented in Table 1.



Table 1. Participant characteristics (N=20).

Characteristic	Males (n=9)	Females (n=11)
Age (years), mean (SD)	70.2 (2.9)	71.1 (3.1)
Height (cm), mean (SD)	176.8 (4.7)	161.1 (6.9)
Weight (kg), mean (SD)	81.6 (12.5)	65.9 (9.3)
People with right dominant arm, n	9	10
People with left dominant arm, n	0	1

Equipment

The following consumer-level and research-grade devices were selected for comparison:

- Fitbit Charge 2 (Fitbit Inc, San Francisco, CA, USA): a wrist-based device with a large organic light-emitting diode screen featuring heartrate monitoring and tracking of steps, distance, calories burned, floors climbed, active minutes, and sleep duration.
- Garmin vivosmart HR+ (Garmin, Olathe, KS, USA): a wrist-based device that monitors heart rate, calories burned, intensity of fitness activities, distance, time, and pace for indoor or outdoor activities.
- 3. ActiGraph GT9X-BT (ActiGraph LLC, Pensacola, FL, USA): a research-grade activity monitor based on motion sensors that provide raw data and numerous activity and sleep measures (eg, activity counts, energy expenditure, steps taken, activity/sedentary bouts, sleep latency/efficiency) via publicly available validated algorithms. The device can be worn on the waist, hip, wrist, ankle, or thigh.
- 4. New-Lifestyles NL-2000i Activity Monitor (New-Lifestyles Inc, Lee's Summit, MO, USA): a research-grade measurement tool for data collection based on a three-dimensional piezoelectric accelerometer. The device can provide details on steps, MVPA, total/active calories, and distance, and can store data for 7 to 14 days. Examples of studies that have validated the pedometer are available [20,22-24].

Experimental Protocol

Two consumer-based wrist-mounted brands were tested (Fitbit and Garmin), worn on the nondominant arm. The trackers' position on the wrist was randomized. The dominant side of the waist (midaxillary line) and the dominant wrist are reported to be optimal for monitoring energy expenditure, MVPA, and sleep in older adults [25-28]; therefore, two ActiGraph monitors were located in these positions as a reference for those parameters. Energy expenditure, MVPA levels, and steps were measured with the ActiGraph on the waist; sleep parameters were extrapolated from the ActiGraph on the wrist. The New-Lifestyles NL-2000i tracker was also worn on the dominant waist (midaxillary line) and was considered as a reference for step counting. Figure 1 illustrates the body positions of the different devices on a participant.

The algorithm adopted by the ActiGraph for estimating energy expenditure was based on the method designed by Crouter et al [29] and also considered in Patterson et al [26]. This method provides estimations expressed in metabolic equivalent (MET), which are later converted into total calories per day. Likewise, the algorithms of Troiano et al [30] and Cole-Kripke et al [31] were considered for estimating MVPA levels and sleep parameters, respectively, through the ActiGraph accelerometer [28,32]. By definition, MVPA level is the amount of time spent performing any activities requiring more than 3 METs, which according to Troiano et al [30] is defined by the ActiGraph by at least 2020 counts per minute. Finally, to guarantee a fair comparison for the different trackers, only the sleep parameters measured by both the Fitbit and Garmin were analyzed. Those parameters were the total sleep time (TST), and the wake after sleep onset (WASO). Participants were asked to complete a sleep diary as well, and the in-bed and out-bed information required by the Cole-Kripke method were input manually according to the values reported in the sleep diary.

The devices were attached on to the person in the morning for data collection and were returned to the researchers the following morning. All trackers were removed by the participants during bathing, whereas only the trackers on the waist were removed during sleep.

Nonwear periods were defined as 90 minutes or more with no activity counts [33]. A valid day was defined as 10 wearing hours or more in a 24-hour period [30].

Statistical Analysis

Descriptive statistics were run on the computed parameters. The following indicators were computed for each parameter and device: mean estimated value with related standard deviation (SD), mean bias with standard deviation, mean percentage error (MPE) with standard deviation, and mean absolute percentage error (MAPE). Intraclass correlation (ICC[2,1]) was performed for each tracker compared against all other devices and the criterion as well. The related 95% confidence intervals (CIs) were also computed. ICC values less than 0.5, between 0.5 and 0.75, between 0.75 and 0.9, and greater than 0.90 are indicative of poor, moderate, good, and excellent reliability, respectively [34]. Bland-Altman plots were also obtained for every parameter comparing all the possible permutations of trackers and the criterion. All statistical analyses were performed using MATLAB (MathWorks, Natick, MA, USA).



Figure 1. Placement of devices on a participant.



Results

Overall, 20 participants took part in the data collection. All participants were white of Irish and British ancestry. Data collection was carried out at the Tyndall National Institute between April 2018 and August 2018.

Table 2 shows the mean values measured, mean error, mean percentage error, and MAPE with related standard deviations for the activity monitors and each parameter. Likewise, Table 3 shows the related ICCs with the 95% CI for each tracker and every parameter. Figure 2 displays the MAPE.

The average wear time for each tracker among all participants was mean 963 (SD 102) minutes per day; thus, every monitored day for each participant was deemed a valid test day above the 10 wearing hours threshold. All monitored days were on weekdays.

For step counting, all the trackers were highly correlated with one another (ICCs>0.89). Although the Fitbit tended to overcount steps (MPE=12.36%), the Garmin and ActiGraph undercounted with a MPE of 9.36% and 11.53%, respectively. For the MAPE, the Garmin and ActiGraph were slightly more accurate with mean MAPEs of 12.89% and 14.23%, respectively. Therefore, all the considered activity trackers can accurately capture steps when worn on the nondominant wrist.

However, the Garmin had poor ICC values when comparing energy expenditure against the New-Lifestyles NL-2000i and the criterion. Likewise, similar results were shown when

comparing the New-Lifestyles NL-2000i and the criterion. Conversely, the Fitbit had moderate-to-good ICCs compared against the other activity trackers. All the tested activity monitors underestimated the amount of calories burned, with the Fitbit showing the best results with a MAPE of 12.25%.

For MVPA level estimation, the Fitbit and Garmin were highly correlated (ICC=0.96), whereas the New-Lifestyles NL-2000i showed poor correlation with these two devices. However, all monitors were moderately correlated against the criterion. The Fitbit and Garmin overestimated the minutes of MVPA activity (mean 12.63, SD 28.31 and mean 13.8, SD 35.7 minutes per day), whereas New-Lifestyles-2000i underestimated, although it showed the best MAPE results, at 45.45%, suggesting that consumer-grade activity trackers may not be reliable in estimating MVPA in older adults.

When analyzing the sleep parameters, the ICCs were poor for all cases, except when comparing the Fitbit to the criterion, which showed a moderate agreement. The TST was slightly overestimated with the Fitbit (mean 5.72, SD 49.11 minutes), although it provided good results with a mean MAPE equal to 10.13%. Conversely, the WASO estimation was poorer; it was overestimated by the Fitbit but underestimated by the Garmin. Again, the Fitbit was the most accurate, with a mean MAPE of 49.7%.

The Bland-Altman plots are shown in Figures 3-6 for steps, energy expenditure, MVPA, and sleep parameters, respectively, and summarized in Table 4.



Table 2. Mean values measured, mean error, mean percentagage error, and mean absolute percentage error with related standard deviation for each parameter and tracker (N=20).

Parameter	Mean (SD)	Mean error (SD)	MPE ^a (SD)	MAPE ^b (SD)
Steps	·		-	
Fitbit	10088.45 (5067.32)	698.60 (1491.27)	12.36 (18.26)	17.05 (13.71)
Garmin	8834.85 (5067.07)	-555.00 (928.11)	-9.36 (18.55)	12.90 (16.16)
ActiGraph	8375.70 (4716.39)	-1014.15 (846.81)	-11.53 (11.55)	14.23 (7.76)
Energy expenditure (cal)				
Fitbit	2324.25 (547.02)	-175.70 (293.18)	-6.98 (12.67)	12.26 (7.33)
Garmin	2434.65 (804.12)	-65.30 (682.94)	-2.09 (27.27)	20.05 (18.04)
NL2000i	2102.30 (256.93)	-397.65 (306.13)	-14.61 (10.41)	16.70 (6.3)
MVPA ^c (min)				
Fitbit	44.32 (45.83)	12.63 (28.31)	31.78 (103.51)	75.74 (75.32)
Garmin	39.10 (57.18)	13.80 (35.70)	6.42 (109.75)	91.98 (47.13)
NL2000i	20.30 (19.03)	-11.70 (18.45)	-41.43 (35.37)	45.45 (29.67)
TST ^d (min)				
Fitbit	389.83 (59.33)	5.72 (49.11)	2.27 (13.66)	10.14 (9.12)
Garmin	442.83 (48.64)	55.39 (48.07)	15.85 (14.56)	16.86 (13.3)
WASO ^e (min)				
Fitbit	49.21 (16.5)	0.21 (22.15)	25.02 (84.72)	49.73 (72.03)
Garmin	13.14 (8.93)	-35.86 (22.13)	-66.89 (28.61)	66.89 (28.61)

 $^{^{\}mathrm{a}}\mathrm{MPE}$: mean percentage error.



^bMAPE: mean absolute percentage error.

^cMVPA: moderate-to-vigorous physical activity.

^dTST: total sleep time.

^eWASO: wake after sleep onset.

Table 3. Intraclass correlation (ICC) and 95% CI for each parameter and tracker (N=20).

Parameter	ICC (95% CI)				
	Fitbit	Garmin	ActiGraph (waist)	NL2000i	Criterion
Step					
Fitbit	a	0.93 (0.65, 0.98)	0.89 (0.39, 0.97)	_	0.95 (0.87, 0.98)
Garmin	_	_	0.98 (0.94, 0.99)	_	0.98 (0.93, 0.99)
ActiGraph (waist)	_	_	_	_	0.97 (0.59, 0.99)
Energy expenditure					
Fitbit	_	0.64 (0.29, 0.84)	_	0.66 (0.21, 0.86)	0.80 (0.48, 0.92)
Garmin	_	_	_	0.32 (-0.07, 0.65)	0.48 (0.05, 0.76)
NL2000i	_	_	_	_	0.45 (-0.10, 0.78)
MVPA ^b					
Fitbit	_	0.96 (0.86, 0.99)	_	0.41 (-0.01, 0.72)	0.69 (0.35, 0.87)
Garmin	_	_		0.46 (-0.11, 0.82)	0.68 (0.18, 0.91)
NL2000i	_	_	_	_	0.60 (0.18, 0.83)
TST ^c					
Fitbit	_	0.43 (-0.10, 0.77)	_	_	0.67 (0.30, 0.86)
Garmin	_	_	_	_	0.42 (-0.10, 0.76)
WASO ^d					
Fitbit	_	< 0.01 (-0.08, 0.18)	_	_	0.32 (-0.28, 0.72)
Garmin	_	_	_	_	0.01 (-0.10, 0.25)

^aNot applicable.



 $^{^{\}mathrm{b}}\mathrm{MVPA}$: moderate-to-vigorous physical activity.

^cTST: total sleep time.

 $^{^{\}rm d}W\!ASO$: wake after sleep onset.

Figure 2. Mean absolute percentage error (MAPE) with standard deviation for each parameter and tracker. EE: energy expenditure; MVPA: moderate-to-vigorous physical activity; TST: total sleep time; WASO: wake after sleep onset.

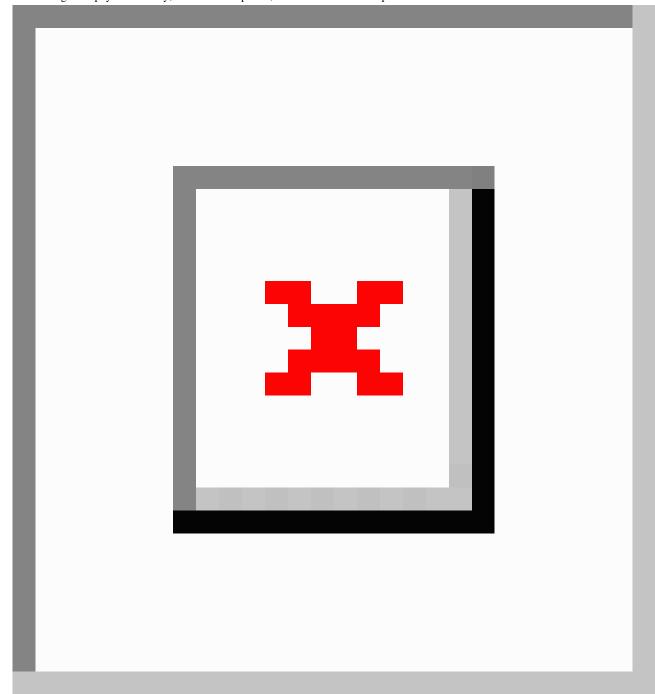




Figure 3. Bland-Altman plots for steps.

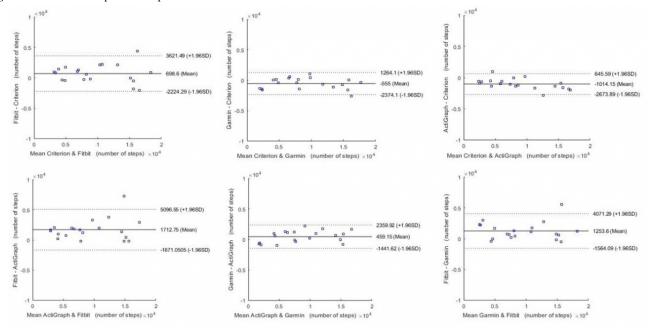


Figure 4. Bland-Altman plots for energy expenditure.

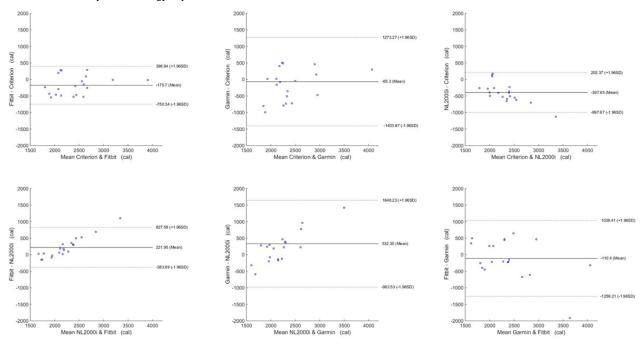




Figure 5. Bland-Altman plots for moderate-to-vigorous physical activity.

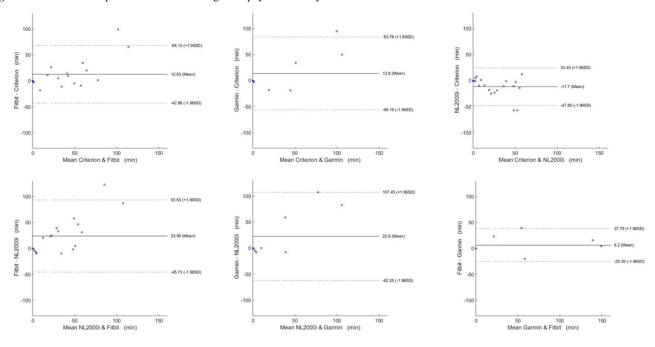


Figure 6. Bland-Altman plots for sleep parameters. Top row: total sleep time; bottom row: wake after sleep onset.

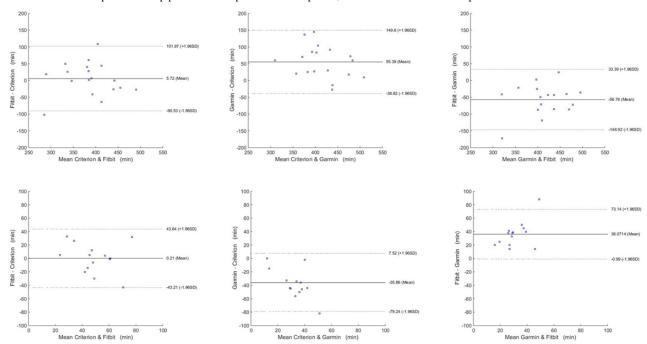




Table 4. Bland-Altman plots summary for each parameter and tracker (N=20).

Parameter	Fitbit	Garmin	ActiGraph (waist)	NL2000i	Criterion
Step	•				
Fitbit					
Mean error (SD)	a	1253.60 (1437.60)	1712.75 (1726.43)	_	698.60 (1491.27)
95% LoA ^b	_	-1564.09, 4071.29	-1671.05, 5096.55	_	-2224.29, 3621.49
Garmin					
Mean error (SD)	_	_	459.15 (969.78)	_	-555 (928.11)
95% LoA	_	_	-1441.62, 2359.92	_	-2374.10, 1264.10
ActiGraph (waist)					
Mean error (SD)	_	_	_	_	-1014.15 (846.81)
95% LoA	_	_	_	_	-2673.89, 645.59
Energy expenditure (cal)					
Fitbit					
Mean error (SD)	_	-110.40 (586.13)	_	221.95 (309.00)	-175.7 (293.18)
95% LoA	_	-1259.21, 1038.41	_	-383.69, 827.59	-750.34, 398.94
Garmin					
Mean error (SD)	_	_	_	332.35 (671.37)	-65.3 (682.94)
95% LoA	_	_	_	-983.53, 1648.23	-1403.87, 1273.27
NL2000i					
Mean error (SD)	_	_	_	_	-397.65 (306.13)
95% LoA	_	_	_	_	-997.67, 202.37
MVPA ^c (min)					
Fitbit					
Mean error (SD)	_	6.20 (16.12)	_	23.95 (35.55)	12.63 (28.31)
95% LoA	_	-25.39, 37.79	_	-45.73, 93.63	-42.86, 68.13
Garmin					
Mean error (SD)	_	_	_	22.60 (43.29)	13.80 (35.70)
95% LoA	_	_		-62.25, 107.45	-56.18, 83.78
NL2000i					
Mean error (SD)	_	_	_	_	-11.70 (18.45)
95% LoA	_	_	_	_	-47.85, 24.45
TST ^d (min)					
Fitbit					
Mean error (SD)	_	-56.76 (45.99)	_	_	5.72 (49.11)
95% LoA	_	-146.92, 33.39	_	_	-90.53, 101.97
Garmin					
Mean error (SD)	_	_	_	_	55.39 (48.07)
95% LoA	_	_	_	_	-38.82, 149.60
WASO ^e (min)					
Fitbit					
Mean error (SD)	_	36.07 (18.91)	_	_	0.21 (22.15)
95% LoA	_	-0.99, 73.14	_		-43.21, 43.64



Parameter	Fitbit	Garmin	ActiGraph (waist)	NL2000i	Criterion
Garmin	·	·	·	•	·
Mean error (SD)	_	_	_	_	-35.86 (22.13)
95% LoA	_	_	_	_	-79.24, 7.52

^aNot applicable.

Discussion

This investigation is one of the first studies to investigate the reliability and accuracy of two consumer-level, wrist-based activity trackers in the estimation of daily step count, total calories expenditure, MVPA, and sleep parameters in a home environment for a cohort of healthy older adults.

Results show that the mainstream monitors may be adopted to estimate steps, energy expenditure, and some sleep parameters (eg, TST) with a certain level of accuracy in a healthy older adult population in free-living settings, whereas other variables (MVPA and WASO) may show excessively large errors.

The measured mean values are largely consistent with data reported studying the same population of interest in other studies adopting non-consumer-level technologies for energy expenditure [26], steps and MVPA [35], and sleep analysis [36].

Regarding step-counting performance, all the trackers presented a good strength of agreement among one another and against the reference device. Also, this study confirms previous findings [17,18] indicating a slight overcounting by the Fitbit device and an undercounting by the ActiGraph. In absolute terms, as shown by the MAPE, there is no significant difference between the Fitbit and the Garmin monitors when monitoring steps.

Similar considerations could be drawn for energy expenditure; however, only the Fitbit shows moderate-to-good agreement with the other trackers. The Fitbit underestimated energy expenditure in our study, confirming findings illustrated in the review by Feehan et al [37], which cited a number of studies in which Fitbits worn on the wrist in free-living settings slightly underestimated METs by 7% (when compared against doubly labeled water), and showed a -10% measurement error (against the SenseWear), and provided MAPE values varying from 16% to 30% when compared with measurements from an ActiGraph or Actiheart accelerometer. However, most of the studies reviewed considered healthy adults and not older adults; thus, the lower MAPE values reported in our study (mean 12.25%, SD 7.33%) may be due to the generally limited amount of moderate-to-vigorous activity performed by older adults. Fitbit showed the narrowest limit of agreement among the trackers, which indicated the device could underestimate the amount of calories per day up to 750.34 kcal and overestimate up to 398.94 kcal.

Due to its many health benefits reported, MVPA may represent an important aspect in people's life and, with aging, this may become even more useful to guarantee independent living and prevention of noncommunicable diseases [35]. The current national MVPA recommendations consider a threshold of 30 minutes per day or 150 minutes per week. Therefore, a correct and reliable estimation of MVPA bouts helps support behavior change techniques applied to sedentary older adults. All the trackers considered (Fitbit, Garmin on the wrist, and New-Lifestyles NL-2000i on the waist) were moderately correlated with the reference. However, the Fitbit and Garmin showed an excellent strength of agreement between each other. The Fitbit and Garmin tended to overestimate MVPA with an average error of 12.63 minutes per day and 13.8 minutes per day, consistent with previously reported results [18], whereas the New-Lifestyles NL-2000i underestimated by 11.7 minutes per day. However, due to the limited moderate-to-vigorous activities performed, MAPE values are large, especially for the wrist-worn devices. The MAPE was mean 75.73% (SD 75.31%) and mean 91.98% (SD 47.13%) for the Fitbit and Garmin, respectively, confirming the large overestimation errors observed in Fitbit devices estimating MVPA in free-living settings compared with an ActiGraph accelerometer in healthy young adults and older adults living with a variety of chronic diseases (MAPEs >30%) [37]. The waist-worn device showed slightly better results both in terms of MAPE and limits of agreement (-47.85 to 24.45).

Finally, aging also impacts sleep, and changes occur in sleep patterns with aging (for example, decrease in the amount of slow wave sleep, increases in non-rapid eye movement sleep, increase in the number of spontaneous arousals, changes in the normal circadian sleep cycle) [38]. Moreover, older adults are more prone to develop sleep-related respiratory disorders, which are associated with cardiovascular disease, metabolic disorders, and impaired neurocognition [38]. Thus, low-cost, unobtrusive, and effective sleep monitoring devices such as consumer-level activity trackers are ideal for providing insightful details on the normal changes in sleeping patterns with advancing age. Between the Fitbit and the Garmin, the Fitbit was moderately correlated with the ActiGraph worn on the wrist, and only for the estimation of the sleeping time. TST was overestimated by a mean 5.72 minutes per day with a MAPE equal to 10.13% (SD 9.12%), which are largely consistent with findings reported in other studies adopting Fitbit devices to investigate sleep measurement accuracy in healthy young adults in free-living settings (MAPE approximately 10%) [37]. The Garmin showed larger errors with a MAPE of 16.8% (SD 13.3%). In contrast, WASO measurements were poorly correlated against the ActiGraph for both devices. Although the lowest mean error



^bLoA: limits of agreement.

^cMVPA: moderate-to-vigorous physical activity.

^dTST: total sleep time.

^eWASO: wake after sleep onset.

was 0.21 minutes per day for the Fitbit, MAPE was large (mean 49.7%, SD 72%) due to the generally limited amount of time spent awake overnight. Conversely, the Garmin significantly underestimated the measurements. Limits of agreement were similar for both trackers for both parameters. These performances may not be suitable for clinical-grade investigations because they require accurate measurements for supporting the decision-making process. For example, WASO is typically adopted as a criterion for discriminating insomnia and normal-sleeper groups (the general threshold is WASO ≥31 minutes per day occurring at least three times per week for at least 6 months) [39]; thus, the WASO estimation inaccuracy may hinder the adoption of mainstream wristband devices for clinical assessments in populations expected to have abnormal sleep patterns [40].

It is worth clarifying that there is no universally accepted definition of an acceptable degree of error for physical activity wearable devices. Some studies recommend that an acceptable measurement error under controlled conditions or for research purposes is within $\pm 3\%$ [41,42] and under free-living conditions is within $\pm 10\%$ [41,42]. Other studies recommend that mean errors of less than 20% have acceptable validity for clinical purposes [43]. This investigation considers the validity criteria between the tested and criterion physical activity measures for clinical purposes when the mean error is less than 20%. Results suggest that the tested devices could be adopted to estimate steps, energy expenditure, and sleep duration with an acceptable level of accuracy in the population of interest, whereas clinicians should be cautious in considering other parameters (eg, MVPA, awakenings) for clinical and research purposes. Although performance estimation is modest in some variables, it may still be adequate for guidance purposes. For instance, the ever-growing acceptance of wearable technologies by older people may push the adoption of wrist-worn trackers in behavior change investigations [11,44].

This study was limited to healthy older adults. As a consequence, it is difficult to indicate if these findings are generalizable to less active older adults or impaired or hospitalized older adults. Indeed, as shown in the literature, step-counting accuracy in people using a walking aid in a

laboratory-structured protocol represents a challenge for all consumer-level trackers as evidenced by large MAPE values. Moreover, the small number of studied participants and the reduced intervention duration may also limit the generalizability of these findings. Thus, further studies would be needed to investigate activity trackers' performance in a large cohort and also in nonhealthy populations.

Although the most common commercial trackers were considered in this study, it is difficult to indicate if results may translate to other consumer monitors on the market, due to the different algorithms they may employ.

This analysis was limited to some health parameters, whereas other variables, which may be of interest in older adults, could not be taken into account due to the lack of a gold-standard for nonlaboratory settings. Some examples are sedentary bouts, light activity bouts, the amount of time spent in different postures, distance traveled, speed, additional sleep measures (eg, sleep efficiency, sleep latency), and physiological measures, such as continuous heart rate measurements, blood oxygen saturation levels, galvanic skin response, blood pressure, or photoplethysmography, and these should be further investigated in future studies.

This study explored the performance of two wrist-worn trackers (Fitbit Charge2 and Garmin vivosmart HR+) estimating steps, energy expenditure, MVPA levels, and sleep parameters against gold-standard technologies in a free-living environment in a cohort of healthy participants aged 65 years and older.

This study confirmed that the wrist-worn devices are effective in estimating steps, energy expenditure, and some sleep parameters with a certain level of accuracy in healthy older adults (lower MAPE values: 12.89% for step counting with the Garmin, 12.25% for energy expenditure with the Fitbit, and 10.13% for TST estimation with the Fitbit). The results were coherent with previous studies, and the observed accuracy was acceptable for monitoring everyday activities. However, clinicians should be cautious in considering other parameters (eg, MVPA levels and WASO) for clinical and research purposes.

Acknowledgments

The authors would like to thank all the participants who voluntarily took part in the study. This publication was from research supported by EU H2020-funded project ProACT under grant agreement No 689996. Aspects of this work have been supported in part by a research grant from Science Foundation Ireland (SFI) and it is cofunded under the European Regional Development Fund under Grant Number 13/RC/2077. Aspects of this work have been supported in part by INTERREG NPA funded project SenDOC.

Conflicts of Interest

None declared.

References

- 1. Kamišalić A, Fister I, Turkanović M, Karakatič S. Sensors and functionalities of non-invasive wrist-wearable devices: a review. Sensors (Basel) 2018 May 25;18(6):1714 [FREE Full text] [doi: 10.3390/s18061714] [Medline: 29799504]
- 2. Alley S, Schoeppe S, Guertler D, Jennings C, Duncan MJ, Vandelanotte C. Interest and preferences for using advanced physical activity tracking devices: results of a national cross-sectional survey. BMJ Open 2016 Dec 07;6(7):e011243 [FREE Full text] [doi: 10.1136/bmjopen-2016-011243] [Medline: 27388359]



- 3. Tao W, Liu T, Zheng R, Feng H. Gait analysis using wearable sensors. Sensors (Basel) 2012 Feb;12(2):2255-2283 [FREE Full text] [doi: 10.3390/s120202255] [Medline: 22438763]
- 4. Tedesco S, Urru A, Clifford A, O'Flynn B. Experimental validation of the Tyndall Portable Lower-limb Analysis System with wearable inertial sensors. Procedia Engineer 2016;147:208-213. [doi: 10.1016/j.proeng.2016.06.215]
- 5. Fortune E, Lugade VA, Kaufman KR. Posture and movement classification: the comparison of tri-axial accelerometer numbers and anatomical placement. J Biomech Eng 2014 May;136(5):051003 [FREE Full text] [doi: 10.1115/1.4026230] [Medline: 24337255]
- 6. Ancillao A, Tedesco S, Barton J, O'Flynn B. Indirect measurement of ground reaction forces and moments by means of wearable inertial sensors: a systematic review. Sensors (Basel) 2018 Aug 05;18(8):E2564 [FREE Full text] [doi: 10.3390/s18082564] [Medline: 30081607]
- 7. Xie J, Wen D, Liang L, Jia Y, Gao L, Lei J. Evaluating the validity of current mainstream wearable devices in fitness tracking under various physical activities: comparative study. JMIR Mhealth Uhealth 2018 Apr 12;6(4):e94 [FREE Full text] [doi: 10.2196/mhealth.9754] [Medline: 29650506]
- 8. Rosenberger ME, Buman MP, Haskell WL, McConnell MV, Carstensen LL. Twenty-four hours of sleep, sedentary behavior, and physical activity with nine wearable devices. Med Sci Sports Exerc 2016 Mar;48(3):457-465 [FREE Full text] [doi: 10.1249/MSS.0000000000000778] [Medline: 26484953]
- 9. Tedesco S, Sica M, Garbay T, Barton J, O'Flynn B. A comprehensive comparison of commercial wrist-worn trackers in a young cohort in a lab-environment. 2018 Mar 19 Presented at: IEEE PERCOM Workshops 2018; March 19-23, 2018; Athens, Greece. [doi: 10.1109/PERCOMW.2018.8480386]
- 10. Malwade S, Abdul S, Uddin M, Nursetyo A, Fernandez-Luque L, Zhu X, et al. Mobile and wearable technologies in healthcare for the ageing population. Comput Methods Programs Biomed 2018 Jul;161:233-237 [FREE Full text] [doi: 10.1016/j.cmpb.2018.04.026] [Medline: 29852964]
- 11. Ehn M, Eriksson LC, Åkerberg N, Johansson A. Activity monitors as support for older persons' physical activity in daily life: qualitative study of the users' experiences. JMIR Mhealth Uhealth 2018 Feb 01;6(2):e34 [FREE Full text] [doi: 10.2196/mhealth.8345] [Medline: 29391342]
- 12. Puri A, Kim B, Nguyen O, Stolee P, Tung J, Lee J. User acceptance of wrist-worn activity trackers among community-dwelling older adults: mixed method study. JMIR Mhealth Uhealth 2017 Nov 15;5(11):e173 [FREE Full text] [doi: 10.2196/mhealth.8211] [Medline: 29141837]
- 13. Tedesco S, Barton J, O'Flynn B. A review of activity trackers for senior citizens: research perspectives, commercial landscape and the role of the insurance industry. Sensors (Basel) 2017 Jun 03;17(6):1277 [FREE Full text] [doi: 10.3390/s17061277] [Medline: 28587188]
- 14. Phillips LJ, Petroski GF, Markis NE. A comparison of accelerometer accuracy in older adults. Res Gerontol Nurs 2015 May;8(5):213-219. [doi: 10.3928/19404921-20150429-03] [Medline: 25942386]
- 15. 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] [Medline: 27257217]
- 16. Straiton N, Alharbi M, Bauman A, Neubeck L, Gullick J, Bhindi R, et al. The validity and reliability of consumer-grade activity trackers in older, community-dwelling adults: a systematic review. Maturitas 2018 Jun;112:85-93. [doi: 10.1016/j.maturitas.2018.03.016] [Medline: 29704922]
- 17. 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 Jul;1(1):e000013 [FREE Full text] [doi: 10.1136/bmjsem-2015-000013] [Medline: 27900119]
- 18. 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] [Medline: 26907794]
- 19. Boeselt T, Spielmanns M, Nell C, Storre JH, Windisch W, Magerhans L, et al. Validity and usability of physical activity monitoring in patients with chronic obstructive pulmonary disease (COPD). PLoS One 2016 Jun;11(6):e0157229 [FREE Full text] [doi: 10.1371/journal.pone.0157229] [Medline: 27305105]
- 20. 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] [Medline: 28595019]
- 21. Burton E, Hill KD, Lautenschlager NT, Thøgersen-Ntoumani C, Lewin G, Boyle E, et al. Reliability and validity of two fitness tracker devices in the laboratory and home environment for older community-dwelling people. BMC Geriatr 2018 Dec 03;18(1):103 [FREE Full text] [doi: 10.1186/s12877-018-0793-4] [Medline: 29724191]
- 22. Grant PM, Dall PM, Mitchell SL, Granat MH. Activity-monitor accuracy in measuring step number and cadence in community-dwelling older adults. J Aging Phys Act 2008 Apr;16(2):201-214. [Medline: 18483442]
- 23. 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] [Medline: 28085918]
- 24. 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]



- 25. Imboden MT, Nelson NB, 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 2018 Jul;52(13):844-850. [doi: 10.1136/bjsports-2016-096990] [Medline: 28483930]
- 26. Patterson M, Wang W, Ortiz A. Dietary and physical activity outcomes determine energy balance in US adults aged 50-74 years. J Aging Phys Act 2018 Oct 01;26(4):561-569. [doi: 10.1123/japa.2017-0304] [Medline: 29283747]
- 27. Kim M, Yoshida H, Sasai H, Kojima N, Kim H. Association between objectively measured sleep quality and physical function among community-dwelling oldest old Japanese: A cross-sectional study. Geriatr Gerontol Int 2015 Aug;15(8):1040-1048. [doi: 10.1111/ggi.12396] [Medline: 25312049]
- 28. Johnson ST, Thiel D, Al Sayah F, Mundt C, Qiu W, Buman MP, et al. Objectively measured sleep and health-related quality of life in older adults with type 2 diabetes: a cross-sectional study from the Alberta's Caring for Diabetes Study. Sleep Health 2017 Dec;3(2):102-106. [doi: 10.1016/j.sleh.2016.12.002] [Medline: 28346155]
- Crouter SE, Kuffel E, Haas JD, Frongillo EA, Bassett DRJ. Refined two-regression model for the ActiGraph accelerometer. Med Sci Sports Exerc 2010 May;42(5):1029-1037 [FREE Full text] [doi: 10.1249/MSS.0b013e3181c37458] [Medline: 20400882]
- 30. Troiano RP, Berrigan D, Dodd KW, Mâsse LC, Tilert T, McDowell M. Physical activity in the United States measured by accelerometer. Med Sci Sports Exerc 2008 Jan;40(1):181-188. [doi: 10.1249/mss.0b013e31815a51b3] [Medline: 18091006]
- 31. Cole RJ, Kripke DF, Gruen W, Mullaney DJ, Gillin JC. Automatic sleep/wake identification from wrist activity. Sleep 1992 Oct;15(5):461-469. [Medline: 1455130]
- 32. Qin J, Barbour KE, Nevitt MC, Helmick CG, Hootman JM, Murphy LB, et al. Objectively measured physical activity and risk of knee osteoarthritis. Med Sci Sports Exerc 2018 Dec;50(2):277-283. [doi: 10.1249/MSS.0000000000001433] [Medline: 28976494]
- 33. Song J, Semanik P, Sharma L, Chang RW, Hochberg MC, Mysiw WJ, et al. Assessing physical activity in persons with knee osteoarthritis using accelerometers: data from the osteoarthritis initiative. Arthritis Care Res (Hoboken) 2010 Dec;62(12):1724-1732 [FREE Full text] [doi: 10.1002/acr.20305] [Medline: 20806273]
- 34. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. J Chiropr Med 2016 Jun;15(2):155-163 [FREE Full text] [doi: 10.1016/j.jcm.2016.02.012] [Medline: 27330520]
- 35. Aspvik NP, Viken H, Zisko N, Ingebrigtsen JE, Wisløff U, Stensvold D. Are older adults physically active enough a matter of assessment method? The Generation 100 Study. PLoS One 2016;11(11):e0167012 [FREE Full text] [doi: 10.1371/journal.pone.0167012] [Medline: 27893785]
- 36. Marino M, Li Y, Rueschman MN, Winkelman JW, Ellenbogen JM, Solet JM, et al. Measuring sleep: accuracy, sensitivity, and specificity of wrist actigraphy compared to polysomnography. Sleep 2013 Nov 01;36(11):1747-1755 [FREE Full text] [doi: 10.5665/sleep.3142] [Medline: 24179309]
- 37. Feehan LM, Geldman J, Sayre EC, Park C, Ezzat AM, Yoo JY, et al. Accuracy of Fitbit devices: systematic review and narrative syntheses of quantitative data. JMIR Mhealth Uhealth 2018 Aug 09;6(8):e10527 [FREE Full text] [doi: 10.2196/10527] [Medline: 30093371]
- 38. Edwards BA, O'Driscoll DM, Ali A, Jordan AS, Trinder J, Malhotra A. Aging and sleep: physiology and pathophysiology. Semin Respir Crit Care Med 2010 Oct;31(5):618-633 [FREE Full text] [doi: 10.1055/s-0030-1265902] [Medline: 20941662]
- 39. Lineberger MD, Carney CE, Edinger JD, Means MK. Defining insomnia: quantitative criteria for insomnia severity and frequency. Sleep 2006 Apr;29(4):479-485. [Medline: <u>16676781</u>]
- 40. Kang SG, Kang JM, Ko KP, Park SC, Mariani S, Weng J. Validity of a commercial wearable sleep tracker in adult insomnia disorder patients and good sleepers. J Psychosom Res 2017 Dec;97:38-44. [doi: 10.1016/j.jpsychores.2017.03.009] [Medline: 28606497]
- 41. 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] [Medline: 14767259]
- 42. 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(1):S10-S15. [doi: 10.17269/cjph.97.1544]
- 43. 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] [Medline: 14523320]
- 44. McMahon SK, Lewis B, Oakes JM, Wyman JF, Guan W, Rothman AJ. Assessing the effects of interpersonal and intrapersonal behavior change strategies on physical activity in older adults: a factorial experiment. Ann Behav Med 2017 Jun;51(3):376-390. [doi: 10.1007/s12160-016-9863-z] [Medline: 28188585]

Abbreviations

MAPE: mean absolute percentage error **MVPA:** moderate-to-vigorous physical activity

TST: total sleep time

WASO: wake after sleep onset



ICC: intraclass correlationMET: metabolic equivalent

Edited by G Eysenbach; submitted 11.12.18; peer-reviewed by B Kim, M Ehn, M Alharbi; comments to author 30.01.19; revised version received 08.02.19; accepted 09.02.19; published 19.06.19

Please cite as:

Tedesco S, Sica M, Ancillao A, Timmons S, Barton J, O'Flynn B

Validity Evaluation of the Fitbit Charge2 and the Garmin vivosmart HR+ in Free-Living Environments in an Older Adult Cohort JMIR Mhealth Uhealth 2019;7(6):e13084

URL: https://mhealth.jmir.org/2019/6/e13084/

doi: <u>10.2196/13084</u> PMID: <u>31219048</u>

©Salvatore Tedesco, Marco Sica, Andrea Ancillao, Suzanne Timmons, John Barton, Brendan O'Flynn. Originally published in JMIR Mhealth and Uhealth (http://mhealth.jmir.org), 19.06.2019. 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.

