CORA Cork Open Research Archive
Cartlann Taighde Oscailte Chorcaí

Title	A comprehensive comparison of commercial wrist- worn trackers in a young cohort in a lab- environment
Authors	Tedesco, Salvatore;Sica, Marco;Garbay, Thomas;Barton, John;O'Flynn, Brendan
Publication date	2018-03-01
Original Citation	Tedesco, S., Sica, M., Garbay, T., Barton, J., O'Flynn, B. (2018) 'A Comprehensive Comparison of Commercial Wrist- Worn Trackers in a Young Cohort in a Lab- Environment', IEEE International Conference on Pervasive Computing and Communications Workshops, PerCom Workshops, Athens, Greece, 19-23 March. doi: 10.1109/PERCOMW.2018.8480386
Type of publication	Conference item
Link to publisher's version	https://ieeexplore.ieee.org/document/8480386 - 10.1109/PERCOMW.2018.8480386
Rights	© 2015 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.
Download date	2026-11-06 12:59:56
Item downloaded from	https://hdl.handle.net/10468/9776



University College Cork, Ireland
Coláiste na hOllscoile Corcaigh

A Comprehensive Comparison of Commercial Wrist-Worn Trackers in a Young Cohort in a Lab-Environment

Salvatore Tedesco, Marco Sica, Thomas Garbay, John Barton, Brendan O'Flynn
WSN Group, Micro & Nano Systems, Tyndall National Institute, University College Cork, Cork, Ireland
Email: salvatore.tedesco@tyndall.ie

***Abstract*—In today's society, the use of watch-based technology is growing steadily and is being used in a wide range of applications and on different aspects of the user's life, from sport and fitness measurement, to entertainment and healthcare evaluation. Considering the multiple application fields for smart-watch/wristbands and their potential adoption in precision medicine applications, it is thus critical to investigate the performance and accuracy of these devices in different potential scenarios of interest. This study investigated the performance and accuracy of a variety of commercially available activity trackers as regards the estimation of stepcount, distance, and heart rate in a number of walking/household/sedentary activities typical in everyday life, and recreated in a lab-environment in a study population of young adults. Results show that heart rate and stepcount measurements are accurate but unstandardized activities, such as common domestic or leisure tasks, may cause large errors in some devices. Finally, travelled distance can also represent a quantification challenge when climbing up/down-stairs. This preliminary work will support the next phase of the project whose goal is to evaluate elderly subjects in lab- and free-living environments in an ambient assisted living context.**

Keywords—Activity Trackers, Steps, Heart Rate, Distance, Statistical Analysis, Assisted Living

I. INTRODUCTION

ProACT (Integrated Technology Systems for ProACTive Patient Centre Care) [1] is an EU-funded Horizon 2020 project whose aim is to develop and evaluate an ecosystem to integrate a wide variety of new and existing technologies to improve and advance home-based integrated care for older adults with multimorbidity. The considered ecosystem will consist of multiple heterogeneous sensing solutions, from smart home systems to body-worn sensors, and in particular watches/wrist-based sensors, as their adoption has been growing massively due to their user-friendliness. Smartwatches and smart-wristbands are wearable devices designed to contain numerous sensors (such as motion sensors, microphone, light, temperature, audio, etc…), a substantial computing power which allows wide processing capabilities, and a battery life of weeks or even months depending on sensor modalities integrated in the system. These features increase the viability of these devices for realistic user studies. Currently, the user uptake of wrist-based technology is growing steadily and is normally used in a wide range of applications and on different aspects of the user's life, from sport and fitness, to entertainment and healthcare [2]. In particular, the healthcare service aims to gather valuable user information through the functioning of embedded inertial sensors (e.g. accelerometer, gyroscope) in order to define and evaluate user's physical activities in real-time. Among a considerable variety of wearable platforms, smartwatch technology has been an emergent solution in the past few years and has opened a new horizon in user-friendliness and suitability for people [3].

Considering the multiple application fields for smart-watch/wristbands and their potential adoption in precision healthcare, in conjunction with an aging population, it is thus fundamental to investigate performance and accuracy of these devices in different scenarios when adopted by elderly people with multimorbidities [4]. Counting steps, for instance, can be challenging on wrist-mounted devices. In particular, movement artefacts (such as the arm swing) can cause under counting or over counting a step. Therefore, it is evident that wrist-worn devices need an activity/motion tracking algorithm which satisfies particular conditions [5-6]. Similar outcomes can be also obtained when investigating daily activities (e.g. walking, sitting, standing, lying, ascending/descending stairs) and personal-related exercises (e.g. brushing teeth, drinking, smoking, writing), and can be extended to another step-correlated variable, such as the travelled distance [7].

Another important parameter related to physical activity and energy expenditure is heart rate. Since heart rate is used to define the intensity category of physical activity and also estimate energy expenditure rate, measuring heart rate accurately is essential to quantify the amount of physical activity and highlight potential issues in the user's vital signs.

The present study investigates the performance and accuracy of different commercial activity trackers in the estimation of stepcount, distance walked, and heart rate in a number of walking/household/sedentary activities recreated in a lab-environment in a young adult cohort. This is a preliminary work which will help confirm results obtained in similar studies in literature and, support the next phase of the project which will evaluate elderly subjects in lab- and free-living environments. In particular, given the fact that commercially available activity trackers are targeted at young healthy subjects, the comparison between the two cohorts will provide a quantification of the inaccuracies and will be

extremely useful to understand where the gaps in commercial devices as regards their use in an aging population are and provide indications on how to fill those gaps for the cohort of interest, that of senior citizens. The manuscript is organized as follows: Section II gives a presentation of the methodology of the study, including a description of the trackers adopted, the subjects analyzed, and the experimental protocol. Results and their thorough discussion are illustrated in Section III. Conclusions and future works are finally drawn in Section IV.

II. Study Methodology

A specific study has been designed to conduct a deep and comprehensive performance evaluation for the sensors used. The study was designed by, and has been conducted at the Wearable Systems Lab, Tyndall National Institute, Cork. This section is devoted to the description of the study methodology.

A. Commercial Activity Trackers

Six consumer-level wrist-worn activity trackers have been selected:

- Philips Health Watch (PHW),
- Fitbit Charge 2 (FC),
- Garmin VivoSmart HR+ (GVS),
- Withings Pulse Ox (WP),
- Jawbone Up3 (JU),
- Xiaomi MiBand2 (XMB),

Moreover, also the research-grade activity tracker ActiGraph GT9X-BT (AG) has been adopted.

The trackers will be compared with some additional devices used for references as gold-standards, such as

- Pedometer Omron HJ-72OITC on the hip,
- Chest-strap Polar H7 on the chest.

Moreover, also the LumoLift (LL) attached on the chest was investigated. The chosen systems can measure different parameters, including steps, distance, calories, activity counts, energy expenditure, floor climbing, sleep etc.

B. Experimental Protocol

The study is based on a sample of twenty (n=20) subjects. Volunteers' recruitment was performed via a general e-mail and word of mouth to staff and students at the University College Cork. The inclusion criteria were to be adults (males or females) between 18 and 40, with no history of motor disorders, with good general health status, not pregnant, or with no permanent disability that affects their movements.

Participants were requested to use the wearable devices on both their wrists, hip and chest, and to perform basic activities in order to analyse specific biometric parameters

In the first scenario, the participants walk on a treadmill at 3, 4 and 5 km/h for 3 min each. Secondly, volunteers were asked to perform household activities again for 3 min each, e.g. walking while carrying a box, dusting and vacuuming and finally climb up and down three floors (6 flights of stairs overall, with 10 steps per flight). Finally, the sedentary activities were writing, computer use/typing, lying down, reading, and playing cards. The aim was to cover the typical household and sedentary activities performed as part of general activities of daily living. Vigorous exercises and related activities have been discarded as they may have posed a health risk to individuals in the study.

A randomization approach was performed to define the position of the trackers as they should be worn for each activity and individual subject, as well as the order of the exercises to be carried out in the tests. Each data collection was carried out in one sitting lasting 3 hours, consisting of three repetitions per activity in order to test all the considered trackers while wearing only maximum three wristbands at the same time. The volunteers could rest as long as they require between the exercises to avoid fatigue. Ethics approval for this study has been obtained from the University College Cork ethical committee.

C. Parameters Analysis

The wearable devices measured specific biometric parameters which can be extrapolated from their commercial front-ends/apps. Those metrics include steps taken, distance, and heart rate. For comparison, stepcount was obtained by all the trackers, distance by each tracker except Philips Health Watch, ActiGraph GT9X-BT, and LumoLift, and heart rate by only Philips Health Watch, Fitbit Charge 2, Garmin VivoSmart HR+, Xiaomi MiBand2.

In the first scenario, involving walking on a treadmill, reference stepcount and heart rate are provided by the gold-standard Omron pedometer and chest-strap Polar H7, respectively, while the actual distance is measured directly on the treadmill.

The household activities were divided in two parts. The first one considered dusting and vacuuming, and only heart rate (reference: Polar H7) and stepcount was measured in this case. Those activities were video-recorded in order to provide a reference for the number of steps taken, as the Omron pedometer has not been validated in similar conditions. The second part consisted of the climb up and down the stairs and the walking while carrying a box activities. Again, heart rate and steps reference were the Polar H7 and the Omron pedometer, respectively. On the other side, the reference distance was obtained in two ways:

- using a well-known formula (shown below) available in literature [8], for the climb up/down-stairs, e.g.

$$\text{Distance} = \text{sqrt}(\text{stairs_depth}^2 + \text{stairs_height}^2) \times \text{number_stairs} \times \pi \times r \times \text{number_turning} \quad (1)$$

where r is defined as the turning radius for each flight of stairs.

- For the walking with a box, using the standard calibration procedure adopted for calibrating the Omron pedometer, which consists of multiplying the number of steps by the average stride length estimated before the test and obtained for each volunteer by measuring the travelled distance after 10 steps and divide it by 10.

As per the sedentary scenario, only step counting was considered. No reference was adopted in this case, since, ideally, the number of steps taken is known *a priori* to be zero. The heart rate was always measured within 1 minute after the end of the performed activity. Unfortunately, no continuous measurements during activities were possible with these

devices at the current stage. Indeed, a correct heart rate measurement required the subjects to be static and do not move for a certain amount of time. Due to this limitation, this investigation considered only one static measurement at the end of the activity.

For each subject written consents were obtained as well as socio-demographic information collected on gender, age, weight, height, BMI and dominant arm. The subjects were divided between 11 men (mean age: 28.3 ± 4.1, mean height: 176 ± 5.73 cm, mean weight: 74 ± 6.8 kg) and 9 women (mean age: 25.9 ± 4.9, mean height: 163.9 ± 8.2 cm, mean weight: 57 ± 5.9 kg). All the subjects, except one, were right-handed.

All the parameters calculated were analysed via well-established statistical tests, helpful to understand trends and correlations. Statistical descriptors included mean error and relative percentage error, MAPE (Mean Absolute Percentage Error), MAD (Mean Absolute Deviation), and correlation coefficient, which are obtained for each activity, for each parameter, and for each tracker. Moreover, every gathered data distribution was subject to the Shapiro-Wilk test to validate the normality assumption of the distribution. Therefore, according to those results, two-sample paired t-test's p-value or Wilcoxon signed-rank test's T-value are also reported.

III. Results and Discussion

In each session, each volunteer performed all the described activities in a randomized order to guarantee no bias due to fatigue. For each activity, the statistical descriptors described in Section II were obtained comparing the tracker values with the gold-standard related to the specific activity. The results are summarized in tables below in order to correlate all the trackers. Results are also shown for the systems being worn on both left and right wrist. Finally, if the p-value is indicated with an asterisk, it has been estimated through a Wilcoxon signed-rank test, while they are presented in bold if lower than the 0.05 significance threshold.

Even though every device requires to input configuration information (height, weight, gender) for each patient before every test, it is unclear how the devices' proprietary algorithms adopt such details due to restricted access to these methods.

Results from data analysis, divided for the three parameters, are described below.

A. Steps

Steps results are shown for the treadmill, household and sedentary scenarios, in Table I-II and Fig. 1, respectively. Considering the walking on a treadmill case, it is evident that all the trackers present similar results for both wrists at every speed. The best results were obtained by the LumoLift which can then be assumed as a gold-standard in following studies for this scenario. Moreover, performance in terms of mean error, relative error, MAPE, MAD were comparable for all the trackers at 4 and 5 km/h with the maximum relative error equal to -6.69%. However, the ActiGraph is the only exception proven by a significant undercounting (MAPE: 20.44% and 23.39% at the two speeds). Walking at 3km/h is slightly different as it is well-known that commercial devices may not be suitable at low speeds [9]. Nevertheless, results were satisfying for most of the trackers, even though some of them presented larger errors, both undercounting (Fitbit: -10.74%, Withing: -16.37%, Xiaomi: -9.14%) and over-counting (ActiGraph: +12.57%).

Considering the household activities, results similar to the treadmill case were obtained in the walking with a box exercise, except for the Withings which presents a considerable difference between wrists (10.9% right and 4.75% left MAPE). Again, the ActiGraph was the least accurate device. Slightly larger errors were obtained for the climb up/down exercises. Fitbit and Xiaomi showed the best performance in both cases, whereas some of the trackers presented differences between the right and left wrist, in particular Garmin, Jawbone, Withings, and ActiGraph (all when climbing up). Good agreement is also obtained with the measurements collected by the LumoLift. However, when considering unstandardized and unregulated activities, such as dusting and vacuuming, errors in step counting can be remarkable. For example, when vacuuming, all the wrist-based trackers, and also the LumoLift, largely undercounted the steps from both wrists, with a relative error going from -17.03% to -97.45%. Similar results were also obtained when dusting but only for the left wrist. Indeed, given that 19 out of the 20 volunteers were right-handed, they all adopted the right hand for performing this task. As a consequence, the device on the left wrist, being not moved, was not able to measure steps effectively (error between -57.84 and -73.20% with the LumoLift reaching -89.73%); on the other side, the right wrist was continuously moved and, thus, the device attached presented a number of steps more similar to the one visually counted (error between -14.84 and 14.6%).

Surprisingly, even though most of the trackers did not count any steps during the performing of sedentary activities, some of them presented over-counting, especially on the right hand. In particular, when observing the right hand results, Philips counted steps also when writing, reading and computing, while all the trackers over-counted during the playing cards task. Much lower errors are observable on the left hand, again with the Philips watch presenting errors during reading and playing cards.

B. Heart Rate

Heart rate results are shown for the treadmill and household scenarios in Table III-IV. Interestingly, in both cases, all the trackers presented consistent results for all the activities and without significant discrepancies between left and right wrist. As an example, the range in which the relative error is included is between -12.41% and 9.25%, which indicates a good agreement with the gold-standard Polar H7. However, it is worth highlighting that these values are the worst-case scenario which occurred only during the carry a box activity. For the other activities the heart rate variation is more contained, with an average variation between -4.3% and 2.78%, which confirms that heart rate measurements are comparable for all the trackers and show good accuracy during every task. Maximum MAPE was instead 16.88%.

C. Distance

Distance results are shown for the treadmill and household scenarios in Table V-VI. It should be specified that the distance reported by all the trackers (except the Xiaomi) is expressed in tens of metres and not in metres. As a result, given the short distances considered in the present data capture, in order to simulate free-living environments and not fitness scenarios, error values are significant. In the treadmill test, the error obtained from all the trackers decreased when going from 3 to 5 km/h, presenting acceptable results at the higher speed but showing larger errors at lower speeds, which may be connected to the steps miscounting evident at these speeds. No specific difference was noted between the arms.

Climbing up/down-stairs also presented large errors over-estimating the travelled distance from 42.86% to 88.57%. The only significant exception was the Xiaomi with a limited over-estimation. However, as mentioned earlier, this may be due to the fact that the Xiaomi is the only tracker expressing distance in metres.

In the walking while carrying a box test, however, the results obtained present lower error, comparable to the treadmill at 5km/h test, with the exception of the Garmin and the Withings.

TABLE I. STEPS. TREADMILL SCENARIO

TREADMILL		FC (R)	FC (L)	GVS (R)	GVS (L)	JU (R)	JU (L)	PHW (R)	PHW (L)	WP (R)	WP (L)	XMB (R)	XMB (L)	AG (R)	AG (L)	LL
3 Km/h	Mean Error (Step)	-29.71	-13.65	-2.15	0.15	13.26	-4.16	-10.11	-8.72	-32.12	-45.68	-25.94	-19.06	25.49	35.36	-2.63
	Relative Error (%)	-10.74	-4.94	-0.68	0.05	4.73	-1.48	-3.58	-3.09	-11.51	-16.37	-9.14	-6.71	9.06	12.57	-0.94
	MAPE (%)	13.34	14.11	2.87	2.56	14.09	13.30	7.89	7.21	13.68	17.48	11.08	7.06	26.17	25.83	2.66
	MAD (Step)	37.47	38.27	8.95	8.05	38.53	31.32	21.89	20.28	38.56	49.05	32.18	20.35	73.53	72.99	7.93
	Correlation	0.18	-0.13	0.72	0.74	0.24	-0.24	0.35	0.23	0.25	0.22	0.15	0.57	0.39	0.37	0.75
	p-value	**0.017**	0.349	0.716	0.979	0.198	0.663	0.190*	0.133*	**0.02***	**< 0.01***	**0.017***	**< 0.01**	**< 0.01***	**< 0.01***	0.472*
4 Km/h	Mean Error (Step)	-19.35	-18.18	-1.70	-11.70	-8.75	-8.95	-10.42	-13.00	-11.50	-14.37	-20.28	-6.11	-48.54	-55.54	2.28
	Relative Error (%)	-6.05	-5.68	-0.49	-3.35	-2.72	-2.78	-3.26	-4.07	-3.64	-4.55	-6.40	-1.93	-15.26	-17.46	0.72
	MAPE (%)	6.64	5.56	2.40	5.16	4.04	4.12	9.56	4.19	3.88	4.52	7.80	4.37	19.39	20.44	1.80
	MAD (Step)	50.11	17.83	8.70	19.50	13.25	13.55	30.21	13.53	12.33	14.68	24.94	13.67	63.03	66.23	5.44
	Correlation	-0.13	0.66	0.70	-0.30	0.62	0.70	0.48	0.83	0.73	0.30	0.09	0.54	-0.26	-0.15	0.75
	p-value	**0.012**	**< 0.01**	0.815	0.138	0.157	0.198	0.405	**0.049**	**< 0.01***	**0.017**	**0.012***	0.136*	**< 0.01***	**< 0.01***	0.666
5 Km/h	Mean Error (Step)	-15.86	-23.34	5.07	1.03	0.00	-9.96	-11.05	-10.60	-4.53	-9.37	-15.00	-22.99	-71.25	-80.11	-2.11
	Relative Error (%)	-4.55	-6.69	6.89	1.40	0.00	-2.87	-3.18	-3.05	-1.31	-2.70	-4.27	-6.55	-20.51	-23.06	-0.60
	MAPE (%)	8.49	7.98	14.21	18.94	2.31	6.12	3.68	3.83	1.90	2.76	5.37	6.33	20.75	23.39	1.33
	MAD (Step)	30.26	27.95	8.69	13.18	7.90	21.21	12.55	13.30	6.53	10.00	19.95	24.00	73.84	84.08	5.06
	Correlation	0.14	0.37	0.10	-0.05	0.77	0.48	0.79	0.74	0.86	0.75	0.37	-0.09	-0.18	-0.39	0.88
	p-value	**0.029**	**0.022**	0.120	0.204	1	0.332165	0.166	0.155	0.530	0.128	0.140	**< 0.01**	**< 0.01***	**< 0.01***	0.786

TABLE II. STEPS. HOUSEHOLD ACTIVITIES SCENARIO

HOUSE ACTIVITIES		FC (R)	FC (L)	GVS (R)	GVS (L)	JU (R)	JU (L)	PHW (R)	PHW (L)	WP (R)	WP (L)	XMB (R)	XMB (L)	AG (R)	AG (L)	LL
Climb up the stairs	Mean Error (Step)	3.29	-0.15	5.07	1.03	5.59	0.95	2.95	-1.59	7.05	5.34	3.70	2.93	2.16	-4.11	2.19
	Relative Error (%)	4.37	-0.21	6.89	1.40	7.29	1.24	3.86	-2.08	9.88	7.48	4.96	3.92	2.87	-5.46	2.94
	MAPE (%)	6.14	5.78	14.21	18.94	9.26	5.33	7.40	8.52	26.37	8.23	6.94	8.26	8.97	12.01	11.47
	MAD (Step)	4.56	4.33	8.69	13.18	7.06	4.00	5.47	6.61	10.11	6.29	4.10	5.11	6.28	8.75	7.00
	Correlation	0.21	0.41	0.10	-0.05	0.32	0.78	0.27	0.25	-0.24	0.14	0.44	0.33	0.19	0.17	0.23
	p-value	**0.015**	0.923	0.120	0.204	**0.014***	0.464	0.137*	0.438*	0.082*	0.790	0.078*	0.111*	**< 0.01***	**< 0.01***	0.065*
Go down the stairs	Mean Error (Step)	-0.57	-3.32	3.56	1.39	-2.60	-2.31	-8.10	-9.72	-7.74	-5.02	-0.50	0.40	-0.29	-0.99	-1.59
	Relative Error (%)	-0.73	-4.25	4.53	1.78	-3.31	-2.93	-9.71	-11.65	-9.41	-6.10	-0.64	0.51	-0.37	-1.26	-1.94
	MAPE (%)	5.23	6.35	11.28	10.02	9.15	10.33	12.43	11.43	6.66	5.73	5.97	7.29	6.56	8.85	4.51
	MAD (Step)	4.06	5.17	8.53	7.94	7.06	8.29	11.60	8.95	5.22	4.56	5.00	5.80	5.04	6.97	3.50
	Correlation	0.69	0.38	0.68	0.18	0.55	0.43	0.36	0.39	0.79	0.82	0.29	0.46	0.57	0.38	0.81
	p-value	0.138*	0.092*	0.117*	0.078*	0.776*	0.173*	0.073*	0.066*	**0.011***	0.731*	0.534*	0.368*	0.626*	0.723*	0.737*
Carry a box	Mean Error (Step)	5.08	4.30	8.47	9.15	2.12	-3.30	-13.90	-15.15	-21.70	-13.20	-6.79	-5.16	-30.03	-25.94	-13.44
	Relative Error (%)	1.61	1.37	2.64	2.85	0.67	-1.04	-4.23	-4.61	-6.69	-4.07	-2.10	-1.60	-9.34	-8.07	-4.13
	MAPE (%)	4.72	5.83	3.59	3.46	5.60	4.94	5.78	5.32	10.90	4.75	3.60	3.78	12.41	12.15	5.32
	MAD (Step)	13.50	17.28	11.21	11.06	16.65	14.67	19.40	17.95	35.00	15.40	11.63	12.42	39.99	38.81	17.59
	Correlation	0.79	0.70	0.91	0.89	0.53	0.60	0.40	0.37	0.40	0.79	0.76	0.75	0.17	0.16	0.34
	p-value	0.612	0.675	0.227	0.314	0.805	0.686	0.067	**0.012**	0.072	0.104	0.150	0.400	**< 0.01***	**< 0.01***	0.234
Vacuuming	Mean Error (Step)	-45.30	-68.80	-35.90	-78.30	-45.50	-76.30	-17.30	-48.80	-83.80	-99.30	-44.90	-81.50	-57.52	-74.04	-104.40
	Relative Error (%)	-42.34	-64.30	-34.32	-74.86	-43.01	-72.12	-17.03	-48.03	-82.24	-97.45	-44.41	-80.61	-55.42	-71.34	-96.67
	MAPE (%)	64.23	77.42	66.93	80.38	69.15	72.90	56.64	49.59	82.68	97.01	75.02	82.62	68.59	76.00	96.63
	MAD (Step)	65.90	81.80	69.30	81.10	70.10	76.30	56.30	48.80	83.80	99.30	71.30	81.50	66.63	77.00	104.40
	Correlation	0.79	0.53	0.64	0.58	0.78	0.29	0.82	0.82	0.08	-0.39	0.82	0.48	0.75	0.46	0.05
	p-value	**0.049***	**0.012***	0.106*	**0.040***	0.084*	**< 0.01***	0.482	**< 0.01**	**< 0.01***	**< 0.01***	**0.047***	**< 0.01***	**< 0.01***	**< 0.01***	**< 0.01***
Dusting	Mean Error (Step)	-6.40	-58.30	6.60	-57.20	-10.40	-61.20	13.40	-67.20	10.00	-66.20	-13.80	-57.10	7.96	-57.30	-83.00
	Relative Error (%)	-6.69	-60.92	6.67	-57.84	-9.91	-58.34	14.60	-73.20	10.95	-72.51	-14.84	-61.40	8.25	-59.36	-89.73
	MAPE (%)	42.32	65.17	69.92	64.78	52.50	61.18	48.41	71.58	83.56	77.63	82.38	66.22	50.05	74.78	86.56
	MAD (Step)	42.60	63.10	63.40	61.60	49.80	61.20	38.20	67.20	73.20	73.00	71.60	60.50	42.26	70.11	83.00
	Correlation	0.36	0.21	0.38	0.49	0.30	0.86	0.25	0.08	-0.38	-0.20	-0.18	0.33	0.23	0.16	-0.38
	p-value	0.743	**< 0.01***	0.825	**< 0.01**	0.650	**< 0.01**	0.494	**< 0.01**	0.717	**< 0.01***	0.576	**< 0.01***	0.489*	**< 0.01***	**< 0.01***

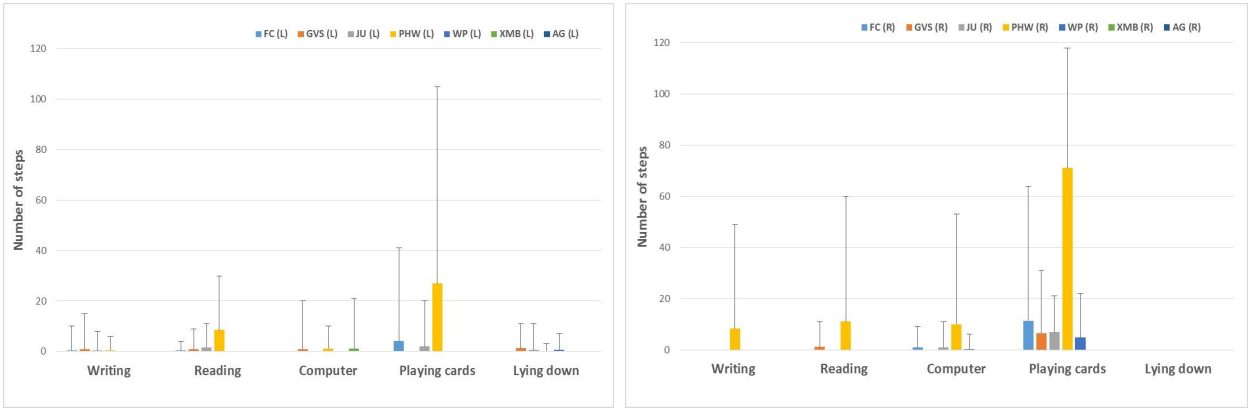


Fig. 1. Steps measured in the sedentary scenario. Left wrist on the left and right wrist on the right.

TABLE III. HEART RATE. TREADMILL SCENARIO

TREADMILL		FC (R)	FC (L)	GVS (R)	GVS (L)	PHW (R)	PHW (L)	XMB (R)	XMB (L)
3 Km/h	Mean Error (bpm)	-3.21	-2.93	-1.21	0.37	1.26	3.05	-2.21	-5.07
	Relative Error (%)	-3.89	-3.54	-1.53	0.47	1.55	3.75	-2.73	-6.26
	MAPE (%)	7.05	6.59	9.37	6.39	4.93	5.58	5.69	6.66
	MAD (bpm)	5.74	5.59	0.63	5.00	3.89	4.42	4.74	5.80
	Correlation	0.89	0.84	0.36	0.60	0.85	0.87	0.85	0.71
	p-value	**0.029***	0.061*	0.714	0.777	0.846	0.660	0.520	0.163
4 Km/h	Mean Error (bpm)	-4.83	-0.83	-2.77	-2.94	0.81	-0.45	-1.16	-3.51
	Relative Error (%)	-5.58	-0.96	-3.35	-3.56	0.96	-0.53	-1.42	-4.31
	MAPE (%)	6.61	7.06	9.15	8.67	5.28	5.51	3.79	4.65
	MAD (bpm)	5.83	6.17	6.50	7.18	4.21	4.85	3.06	3.71
	Correlation	0.89	0.84	0.26	0.43	0.91	0.87	0.93	0.93
	p-value	**0.012***	0.906	0.159	0.138	0.811	0.688	0.989	0.634
5 Km/h	Mean Error (bpm)	-1.20	0.64	-2.85	-1.60	0.81	-0.10	-0.21	-2.65
	Relative Error (%)	-1.44	0.77	-3.53	-1.98	0.96	-0.12	-0.26	-3.31
	MAPE (%)	6.53	6.56	7.77	9.52	4.74	7.82	2.69	6.61
	MAD (bpm)	5.40	5.11	6.15	7.30	3.84	6.40	2.26	5.27
	Correlation	0.89	0.90	0.69	0.42	0.86	0.71	0.95	0.86
	p-value	0.444*	0.732*	0.109	0.164	0.915	0.986	0.538	0.189*

TABLE IV. HEART RATE. HOUSEHOLD ACTIVITIES SCENARIO

HOUSE ACTIVITIES		FC (R)	FC (L)	GVS (R)	GVS (L)	PHW (R)	PHW (L)	XMB (R)	XMB (L)
Climb up the stairs	Mean Error (bpm)	-0.40	1.10	-2.37	-2.20	0.55	-0.05	-0.05	-2.06
	Relative Error (%)	-0.49	1.35	-3.04	-2.82	0.67	-0.06	-0.06	-2.59
	MAPE (%)	4.39	12.16	4.54	12.18	5.60	7.64	6.50	5.11
	MAD (bpm)	3.60	8.50	3.89	8.50	4.65	6.15	5.00	4.00
	Correlation	0.95	0.75	0.73	0.63	0.87	0.80	0.77	0.87
	p-value	0.758*	0.177*	0.140	0.888	0.947	0.654	0.844	0.379*
Go down the stairs	Mean Error (bpm)	-5.32	-2.19	-4.63	-2.95	5.41	0.80	-1.58	-5.90
	Relative Error (%)	-6.43	-2.65	-5.67	-3.61	6.61	0.98	-1.95	-7.30
	MAPE (%)	6.79	5.57	8.01	9.09	8.52	6.49	6.36	9.16
	MAD (bpm)	5.53	4.78	6.74	7.26	6.26	5.20	4.95	7.44
	Correlation	0.95	0.88	0.56	0.58	0.57	0.78	0.91	0.78
	p-value	**< 0.01***	0.071*	0.099	0.260	0.218	0.680	0.785	0.478
Carry a box	Mean Error (bpm)	-2.45	0.55	-0.40	5.70	2.42	6.79	-8.90	-0.41
	Relative Error (%)	-3.23	0.73	-0.55	7.80	3.30	9.25	-12.41	-0.57
	MAPE (%)	9.48	5.35	11.27	12.70	6.51	5.83	16.88	7.50
	MAD (bpm)	4.25	4.25	7.00	9.40	4.11	4.47	10.70	5.94
	Correlation	0.97	0.93	0.83	0.18	0.94	0.90	0.67	0.60
	p-value	**0.019**	0.447	0.813*	0.831	0.168*	0.794	0.093*	0.614
Vacuuming	Mean Error (bpm)	-0.67	-0.56	4.40	2.07	4.11	2.49	2.00	-0.88
	Relative Error (%)	-0.82	-0.69	5.41	2.54	5.15	3.11	2.46	-1.08
	MAPE (%)	4.56	6.73	4.91	4.16	1.62	1.78	2.97	2.46
	MAD (bpm)	3.33	5.22	3.88	3.13	1.43	1.43	2.25	1.88
	Correlation	0.93	0.88	0.91	0.95	0.98	0.99	0.99	0.99
	p-value	0.893	0.983	0.720	0.970	0.968	0.887	0.762	0.900
Dusting	Mean Error (bpm)	-2.58	0.08	0.07	0.38	-0.44	-0.22	-1.00	-0.50
	Relative Error (%)	-3.10	0.10	0.08	0.45	-0.53	-0.26	-1.19	-0.60
	MAPE (%)	6.38	4.11	1.19	2.50	1.79	1.32	2.78	5.79
	MAD (bpm)	5.25	3.38	1.00	1.75	1.56	1.11	2.25	4.75
	Correlation	0.92	0.90	0.99	0.99	0.97	0.99	0.97	0.90
	p-value	0.404	0.809	0.969	0.956	0.934	0.967	0.870	0.944

TABLE V. DISTANCE. TREADMILL SCENARIO

TREADMILL		FC (R)	FC (L)	GVS (R)	GVS (L)	JU (R)	JU (L)	WP (R)	WP (L)	XMB (R)	XMB (L)
3 Km/h	Mean Error (Km)	0.02	0.03	0.06	0.05	0.07	0.03	0.01	0.02	0.00	0.02
	Relative Error (%)	15.00	21.67	37.04	36.49	44.91	16.67	7.78	10.37	3.00	13.33
	MAPE (%)	19.67	22.33	37.78	37.89	44.91	36.67	30.00	24.44	17.67	16.67
	MAD (Km)	0.03	0.03	0.06	0.06	0.07	0.06	0.05	0.04	0.03	0.03
	p-value	**< 0.01**	0.06	**< 0.01***	**< 0.01***	**< 0.01***	0.394*	0.136	**< 0.01**	0.167*	**< 0.01***
4 Km/h	Mean Error (Km)	0.01	0.00	0.04	0.05	0.03	0.04	0.03	0.04	0.01	0.02
	Relative Error (%)	6.00	1.75	19.75	23.00	13.50	17.50	15.28	20.79	3.50	9.75
	MAPE (%)	8.50	7.75	19.75	23.00	14.50	17.50	15.83	20.79	14.50	13.75
	MAD (Km)	0.02	0.02	0.04	0.05	0.03	0.04	0.03	0.04	0.03	0.03
	p-value	**0.012**	0.105*	**< 0.01***	**< 0.01***	**< 0.01**	**0.016***	**< 0.01**	**< 0.01**	0.358*	**0.034**
5 Km/h	Mean Error (Km)	-0.01	-0.02	0.01	0.01	0.02	0.01	0.03	0.04	-0.01	-0.01
	Relative Error (%)	-5.89	-7.37	5.80	2.60	8.00	5.26	13.89	14.80	-2.60	-5.80
	MAPE (%)	9.26	13.26	5.80	7.80	10.53	22.11	14.74	14.80	9.80	15.00
	MAD (Km)	0.02	0.03	0.01	0.02	0.03	0.06	0.04	0.04	0.02	0.04
	p-value	**< 0.01***	**0.026***	**< 0.01***	0.170*	**< 0.01**	0.393*	**< 0.01**	**< 0.01***	0.876*	0.686

TABLE VI. DISTANCE. HOUSEHOLD ACTIVITIES SCENARIO

HOUSE ACTIVITIES		FC (R)	FC (L)	GVS (R)	GVS (L)	JU (R)	JU (L)	WP (R)	WP (L)	XMB (R)	XMB (L)
Climb up the stairs	Mean Error (Km)	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.01	0.01
	Relative Error (%)	55.56	62.41	63.91	63.91	64.29	42.86	65.08	70.68	17.14	36.84
	MAPE (%)	54.99	58.12	67.84	73.84	69.07	142.33	75.59	71.59	46.86	49.94
	MAD (Km)	0.02	0.02	0.02	0.03	0.02	0.04	0.03	0.03	0.02	0.02
	p-value	**< 0.01**	**< 0.01***	**< 0.01**	**< 0.01**	**< 0.01***	0.121*	**< 0.01***	**< 0.01***	0.291	**< 0.01**
Go down the stairs	Mean Error (Km)	0.02	0.02	0.03	0.03	0.03	0.03	0.03	0.03	0.02	0.01
	Relative Error (%)	69.75	49.21	72.93	88.57	78.15	71.43	79.83	82.14	62.41	35.34
	MAPE (%)	69.13	48.66	79.10	89.34	79.22	150.80	79.17	81.48	67.88	63.50
	MAD (Km)	0.02	0.02	0.03	0.03	0.03	0.05	0.03	0.03	0.02	0.02
	p-value	**< 0.01**	**< 0.01***	**< 0.01**	**< 0.01***	**< 0.01**	**0.016***	**< 0.01***	**< 0.01***	**< 0.01**	0.612
Carry a box	Mean Error (Km)	0.01	0.02	0.04	0.03	0.01	0.00	0.01	0.03	0.01	-0.01
	Relative Error (%)	7.02	8.27	19.55	14.29	4.83	1.01	2.53	11.72	2.39	-6.20
	MAPE (%)	35.68	38.79	47.03	10.79	10.72	21.75	10.13	11.56	18.94	19.95
	MAD (Km)	0.02	0.02	0.04	0.02	0.02	0.15	0.02	0.03	0.04	0.14
	p-value	0.199	0.102	**< 0.01***	**< 0.01***	0.358	0.554	0.725*	**0.014**	0.687	0.058

IV. CONCLUSION AND FUTURE WORK

This study investigated the performance and accuracy of different commercial activity trackers as regards the estimation of steps taken, distance covered, and heart rate in a number of walking/household/sedentary activities recreated in a lab-environment in a young adult cohort. Results show that heart rate measurements are comparable for all the trackers and show good accuracy during every task. However, measurements were obtained only once at the end of the task. Further studies are needed to investigate if those measurements are also accurate when obtained continuously. Steps are mostly accurate for all trackers, especially when considering standard walking scenarios. However, unstandardized activities, such as domestic or leisure tasks, may cause large errors in some devices. Finally, travelled distance can also represent a challenge when climbing up/down-stairs. The results obtained from evaluating daily activities further motivates the need to identify the activity using other means/classification approaches (e.g., typing detection), to compensate for the over/under-estimations. This preliminary work will support the next phase of the project whose goal is to evaluate elderly subjects in lab- and free-living environments. Data collection is currently ongoing. Additional variables, such as calories consumption, will be also considered. The comparison of the results obtained from the young and older adults cohorts will be extremely useful to understand where the gaps in commercial devices are and provide indications on how to fill those gaps for the cohort of interest of senior citizens.

ACKNOWLEDGMENTS

This publication has emanated 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 is co-funded under the European Regional Development Fund under Grant Number 13/RC/2077.

REFERENCES

[1] http://proact2020.eu (accessed 7/11/17)

[2] S.C. Mukhopadhyay, "Wearable sensors for human activity monitoring: A review," IEEE Sensors J., 15, 3, 1321-1330, 2015.

[3] B. Mortazavi, et al. "Can smartwatch replace smartphone for posture tracking?," Sensors, 15, 10, 26783-26800, 2015.

[4] S. Tedesco, J. Barton, B. O'Flynn, "A review of activity trackers for senior citizens: Research perspectives, commercial landscape and the role of the insurance industry," Sensors, 17, 6, E1277, 2017.

[5] E. Nemati, Y. Soo Suh, B. Motamed, "Gait velocity estimation for a smartwatch platform using Kalman filter peak recovery", IEEE Int. Conf. Wearable Implantable Body Sensor Networks (BSN) 2016, S. Francisco, USA, 14-17 June 2016.

[6] Y. Soo Suh, E. Nemati, M. Sarrafzadeh, "Kalman filter-based walking distance estimation for a smart-watch," IEEE Int. Conf. Connected Health: Applications, Systems, Engineering Technologies (CHASE) 2016, Washington, DC, USA, 27-29 June 2016.

[7] M.B. Nelson, L.A. Kaminsky, D.C. Dickin, A.H. Montoye, "Validity of consumer-based physical activity monitors for specific activity types," Med Sci Sports Exerc, 48, 8, 1619-28, 2016.

[8] Y. Huang, J. Xu, B Yu, P. B. Shull, "Validity of FitBit, Jawbone UP, Nike+ and other wearable devices for level and stair walking," Gait Posture, 48, 36-41, 2016.

[9] R. McCullagh, et al. "A review of the accuracy and utility of motion sensors to measure physical activity of frail older hospitalised patients," J Aging Phys Act, 24, 3, 465-475, 2016.