ACTA IMEKO
ISSN: 2221-870X
March 2024, Volume 13, Number 1, 1 - 7

Validity of iPhone health application step count in semi free-living conditions

Pedro B. Júdice[1,2], Jorge Fouto[1], Miguel Horta[1], Sara Pereira[1,2]

[1] *CIDEFES - Centro de Investigação em Desporto, Educação Física e Exercício e Saúde; Universidade Lusó-fona, Lisbon, Portugal*
[2] *CIFI2D, Faculty of Sport, University of Porto, Porto, Portugal*

ABSTRACT
The iPhone's validity for measuring steps has been mainly investigated under laboratory conditions, while studies that include real-world conditions are still scarce. We examined the validity of iPhones in measuring steps in real-world walking conditions, while using direct observation with video as reference. A sample of 100 adults who owned an iPhone 5S or higher was included and participants were randomly allocated to one of two protocols. Limits of Agreement (LoA), Mean Absolute Percentage Error (MAPE), linear-regression and Bland-Altman analyses were carried out. In Protocol-1, which includes straight-line and zigzag conditions, we observed a low MAPE lower than 4%. Bland-Altman analyses assessed a high accordance of approximately 6 steps under both conditions. Differences between the iPhone and direct observation were only noticeable in straight-line condition (p = 0.002). Likewise, protocol-2 (three 50-step conditions: an upward 5 % slope, a flat surface, and a downward -5% slope) showed low MAPE values (3.78 %) for the upward slope, 2.41 % for the downward slope, and 2.37 % in absence of slope), and differences between the iPhone and direct observation were only observed for the downward slope (p = 0.017), with the iPhone overestimating. Our findings revealed that iPhone might be a reliable tool for monitoring walking in real-life conditions, however, downward slope seems to generate overestimation, which deserves future investigation.

Section: RESEARCH PAPER

Keywords: Accuracy; free-living; precision; smartphones; validation

Citation: P. B. Júdice, J. Fouto, M. Horta, S. Pereira, Validity of iPhone health application step count in semi free-living conditions , Acta IMEKO, vol. 13 (2024) no. 1, pp. 1-7. DOI: 10.21014/actaimeko.v13i1.1673

Section Editor: Leonardo Iannucci, Politecnico di Torino, Italy

Received September 14, 2023; **In final form** March 11, 2024; **Published** March 2024

Copyright: This is an open-access article distributed under the terms of the Creative Commons Attribution 3.0 License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Corresponding author: Pedro B. Júdice, e-mail: pedro.judice@ulusofona.pt

1. INTRODUCTION

Physical activity (PA) promotes health, as people that attain the recommended levels of PA reduce their risk for the most common non-communicable diseases, such as cardiovascular diseases, diabetes, some forms of cancer [1], and even mortality [2]. The technological expansion entails several behavioural vicissitudes that ultimately lower the energy demands in a daily basis, by tumbling the obligation to carry out PA [3].

One of the technological advances mostly used, nowadays, by the people is the smartphone, a portable device that is devote not only to the primary functions of a phone, but that grants several other tasks as well (e.g., check the bank account, watch movies, pay for services, explore the internet, send emails). The multi-tasking characteristic of smartphones identified these devices as a constant "companion" of the users, while awake [4]. It is estimated that 83.72 % of the world's population owns a smartphone and, from those, the mobile manufacturer market share advances that approximately 59.12 % own an iPhone (i.e., smartphone from Apple).

Even though these devices can boost population' sedentary behaviours, there is evidence showing that they can also be a relevant tool for promoting PA [5], [6], as they incorporate movement sensors such as accelerometers that allow, for instance, the measurement of steps walked [6]. The constant feedback on this feature itself [7] can lead people to walk more [5], [8], [9], thus fostering higher PA levels [5], [6] and lower sedentary behaviour [6].

The fact that their owners persistently use these devices turns them into a convenient platform for community-based step count estimation to monitor and encourage PA. However, to increase the confidence in these devices as a valid tool to measure PA in large-scale studies, and according to the Lancet series Bergman, Spellman [10], it is essential to grasp the validity of these smartphones to estimate the number of steps. In this sense, many studies have evaluated the validity of smartphones (mostly iPhones) to measure steps, but these studies have been conducted in small samples and under laboratory conditions. These settings usually imply standardized walking speeds performed on a treadmill [10], [11], [12], which can hardly be translated to the real-world [11], [13]. The assessment of accuracy and its validation in real-world contexts are important steps for the metrological characterization of motion sensors responsible for counting steps on the smartphone [14], [15]. Thus, it is important to carry out controlled tests to assess the accuracy of smartphone's motion sensors. However, the findings from laboratory-based investigations are contradicting. Some studies suggested a good overall validity [12], [16], [17], while other ones showed low accuracy for smartphones [10], [12], [18]. It seems that walking speed is an important feature that can modify the validity of smartphones to measure steps. A recent model of iPhone was found to be accurate for customary and fast speeds, but weak for the slow speed [12], with a mean absolute percent error (MAPE) between iPhone and direct observation of 21 %, 8 %, and 4 % for the slow, customary, and fast speeds, respectively [12]. A further investigation confirmed these findings, as several models of iPhone differed from manually counted steps by a mean bias of less than 5 % when walking at 5 km/h, 7.5 km/h, and 10 km/h on a treadmill, which is acceptable, but the MAPE was higher than 5 % when a lower speed (i.e., 2.5 km/h) was adopted [17].

In addition to the walking speed, there seems to be a difference in smartphones' validity when participants are tested in laboratory vs free-living conditions [17]. When iPhones were compared against accelerometer-derived steps in a free-living environment, a much higher MAPE was found (21.5 % or 1340 steps/day) [17], which is disconcerting. Evidence suggest that the lower validity found in free-living can be justified by the fact that smartphones may not be continuously carried by participants, thus justifying the typical underestimation [17]. It is possible that the wide variability of walking conditions when in free-living (e.g., inclination, interrupted walking, changes in direction and speed), gathered with some recognized limitations of the reference methods used in these studies (i.e., research-based accelerometers) may also explain the lower validity of smartphones while in free-living.

In real life, it is obvious that if a participant does not carry a smartphone, there will be an underestimation of step counts. In fact, an investigation found that the iPhone underestimated steps/day by approximately 12 % compared to a validated pedometer. The underestimation was higher for the participants who reported having some non-carrying time [19]. However, it is important to recognize that when examining the validity of iPhones in free-living conditions, the reference methods that have been used are pedometers or accelerometers, which can hardly be considered reference methods for measuring steps, as they likewise present some relevant limitations [11]. In this context, it is necessary to conduct real-world tests to assess the smartphone's performance in everyday situations such as walking, running, and climbing.

Investigations that aim to examine the validity of smartphones for measuring steps in real-world conditions while using definite gold-standard methods as the reference (i.e., direct observation with video) are still scarce [13], [18], [20]. A recent systematic review suggested that future studies should consider larger sample sizes and include semi free-living conditions combining gold-standard reference methods and real-life conditions [11]. Thus, the aim of this study was to examine the validity of iPhones in measuring steps in real-life walking conditions (i.e., distinct inclinations and inconstant speed and direction) while using direct observation with video as the reference method in a sample of adults.

2. MATERIAL AND METHODS

2.1. Participants

We recruited a convenience sample of 100 healthy adults, who owned an iPhone 5S, 6, 6S, 6plus, SE, 7, or higher (Apple Inc, California, United States), through direct outreach at a university in 2022. Each participant received an explanation of the purpose and duration of the protocol and gave prior consent to participation. Ethical approval was granted by the Ethical Committee of the Sport Faculty, Lusófona University (m1122).

2.2. Measures

Step count was measured during the protocols by using the health app preinstalled on the personal iPhones. The total number of steps shown in the health app immediately before and after each walking condition were recorded and based on those 2 values the number of steps from each condition were calculated. Participants were asked to carry their own iPhones as usual on the side pockets of their pants. A self-reported questionnaire evaluated socio-demographic factors.

The actual steps carried out during each walking condition (i.e., reference method) was assessed by direct observation performed by 3 evaluators plus video recording. If step counting differed between the 3 evaluators by 1 or more steps, then the video was used to assess the actual number of steps walked. The video only captured the legs and feet of the participants. A step was defined as a forward, sideward, or backward displacement of the foot together with a forward displacement of the trunk. Closing steps were not counted.

Participants' demographics such as age, weight, and height were self-reported and body mass index (BMI) was calculated as weight (kg)/(height (m)^2.

2.3. Design and procedures

The participants were randomly allocated to one of 2 protocols. The first structured walking protocol was completed in the following order: 1) overground walking in a straight-line for 50 meters with a comfortable walking speed; 2) overground walking in a Zigzag line for 50 meters, with an inconstant direction. In this second condition, the participants changed direction in a perpendicular manner (i.e., 90 degrees) every 5 meters, until completing 50 meters, thus performing 9 changes of direction and speed. Participants were asked to walk an arc at the turning points by reducing their walking speed at those moments. There were pins to signal the turning points, and participants had to walk toward the next pin until this condition was completed.

The second overground walking protocol consisted of 3 separate conditions. Participants were asked to complete 50 steps on an overground corridor in each condition. The first condition was carried out in a straight-line with an upward slope of 5 % at participants' comfortable walking speed. In the second condition, participants were asked to complete 50 steps on a straight-line with no slope at a comfortable walking speed. Finally, in the third condition, participants were asked to walk in a straight-line with a downward slope of -5 % at participants' comfortable walking speed.

In this second protocol, one assessor walked behind the participant and counted the number of steps silently while enunciating the last 3 steps so that the participant could anticipate the end of the condition. The second assessor also

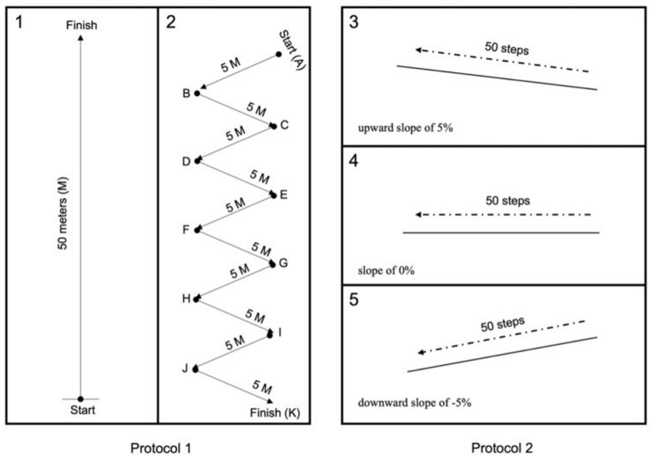

Figure 1. The two protocols and the conditions within each protocol (1 to 5).

counted the number of steps from a distance in silence. A third assessor video recorded and counted the number of steps in each condition, and if there were differences between the 3 assessors, the video recorded was used for counting the actual steps walked. For a more ecological approach, all conditions from both protocols were performed on an outdoor overground sidewalk at the University campus. Figure 1 illustrates the 2 protocols and the 5 distinct conditions within the 2 protocols.

2.4. Statistical analysis

The descriptive statistics for the number of steps in each walking condition was obtained. The Mean Absolute Percent Error (MAPE) between the number of steps measured by the iPhones and the direct observation was calculated. The difference in the step count between iPhone measurement and direct observation was calculated by subtracting the step count of the iPhone from that of the reference in protocol 1. As differences between two measurements did not follow a normal distribution, a non-parametric approach was chosen in all analyses [21]. Limits of accordance were also calculated as previously stated [22]. Moreover, differences in measurements were plotted in a histogram. Wilcoxon sign-rank tests were carried out to determine whether the differences between step counts were statistically significant in the several conditions of both protocols (1 and 2) to determine if the step counts were statistically different from the reference. Linear regression analysis and a Bland-Altman analysis were performed to assess the accordance between the 2 measurements, and sample quantile estimation was used to determine the limits of accordance. Analyses were conducted using STATA v.14.2, and a 5 % significance was adopted.

3. RESULTS

From the 100 participants that were included in this investigation, 60 were randomly allocated to protocol 1 and 40 to the second protocol. No data were lost during the first protocol. However, there was a problem related to data collection in five participants in the second protocol. Thus, a total of 95 participants (43 males) were considered.

Descriptive statistics of the protocol 1 participants are presented in Table 1. Overall, participants were mostly young adults (22 ± 3 years) and had normal weight (23 ± 4 kg/m²). Table 1 shows that, on average, the iPhone step counts in both conditions (straight-line and zigzag) were very close to direct observation step counts.

Table 1. Descriptive statistics for protocol 1.

Protocol 1 (*N* = 60)	Mean ± SD	Min-max
Age (years)	22 ± 3	18.0 - 34.0
BMI (kg/m²)	23 ± 4	17.1 - 31.3
iPhone step counts in straight-line	70 ± 7	53.0 - 87.0
Direct step counts in straight-line	71 ± 6	55.0 - 85.0
iPhone step counts in zigzag line	67 ± 7	48.0 - 84.0
Direct step counts in zigzag line	67 ± 6	50.0 - 81.0

BMI = body mass index

Comparisons between iPhone step count and direct observation are shown in Table 2.

The MAPE for the iPhone against direct observation was small in both conditions (3.0 % in straight-line and 3.2 % in zigzag line). Wilcoxon signed-rank test for differences between the iPhone and direct observation was significant in the straight-line condition (z = -3.165, p = 0.002) but no differences were found for the zigzag line condition (z = -0.04, p = 0.964). This result means that there are significant differences between the iPhone and direct observation step counting in the straight-line condition, but no differences were found between methods in the zigzag line condition ($p > 0.05$).

However, Bland-Altman plots (Figure 2) demonstrated that the iPhone had a high level of measurement concordance with direct observation for both conditions. More specifically, during 50 meters, the iPhone estimated lower limits of agreement were -6.0 and -9.4 steps, and upper limits of agreement were 9.9 and 6.0 for straight-line and zigzag line, respectively. In a more individualized analysis, for the straight-line condition, the iPhone overestimated the number of steps in 16.7 % and underestimated 66.7 % of the participants, with 16.7 % presenting exactly equal values compared to the reference. For the zigzag condition, there was 48.3 % of underestimation, 33.3 % of overestimation, and 18.3 % of equal values. Moreover, 98.3 % of the observations are within the 2.5th and 97.5th limits in both conditions.

As in protocol 1, participants in protocol 2 were mostly young adults (21.0(4.5) and had normal weight (24.1(5.8)). Table 3 presents the distribution of estimated step counts by iPhone in 3 different conditions:

1) straight-line with a downward slope of -5 % (51.0(4.0));
2) straight-line with no slope (50.0(3.5));
3) straight-line with an upward slope of 5 % (51.0(3.0)).

Moreover, MAPE results showed a higher value for the upward slope condition (3.78 %), followed by a downward slope (2.41 %) and no slope (2.37 %). However, the Wilcoxon sign-rank test was only significant for the downward slope condition ($p < 0.001$), with the iPhone overestimating the number of steps, as shown in Figure 3.

Table 2. Limits of agreement (LoA), mean absolute percentage error (MAPE), and Wilcoxon signed-rank test for each condition in protocol 1.

	LoA 2.5th to 97.5th	MAPE (%)	One sample t-test		
			Mean ± SD	z	p-value
Straight line	-6.00 to 9.90	3.04	-1.00(2.00)	-3.165	0.002
Zigzag line	-9.38 to 6.00	3.22	0.00 (3.50)	-0.045	0.964

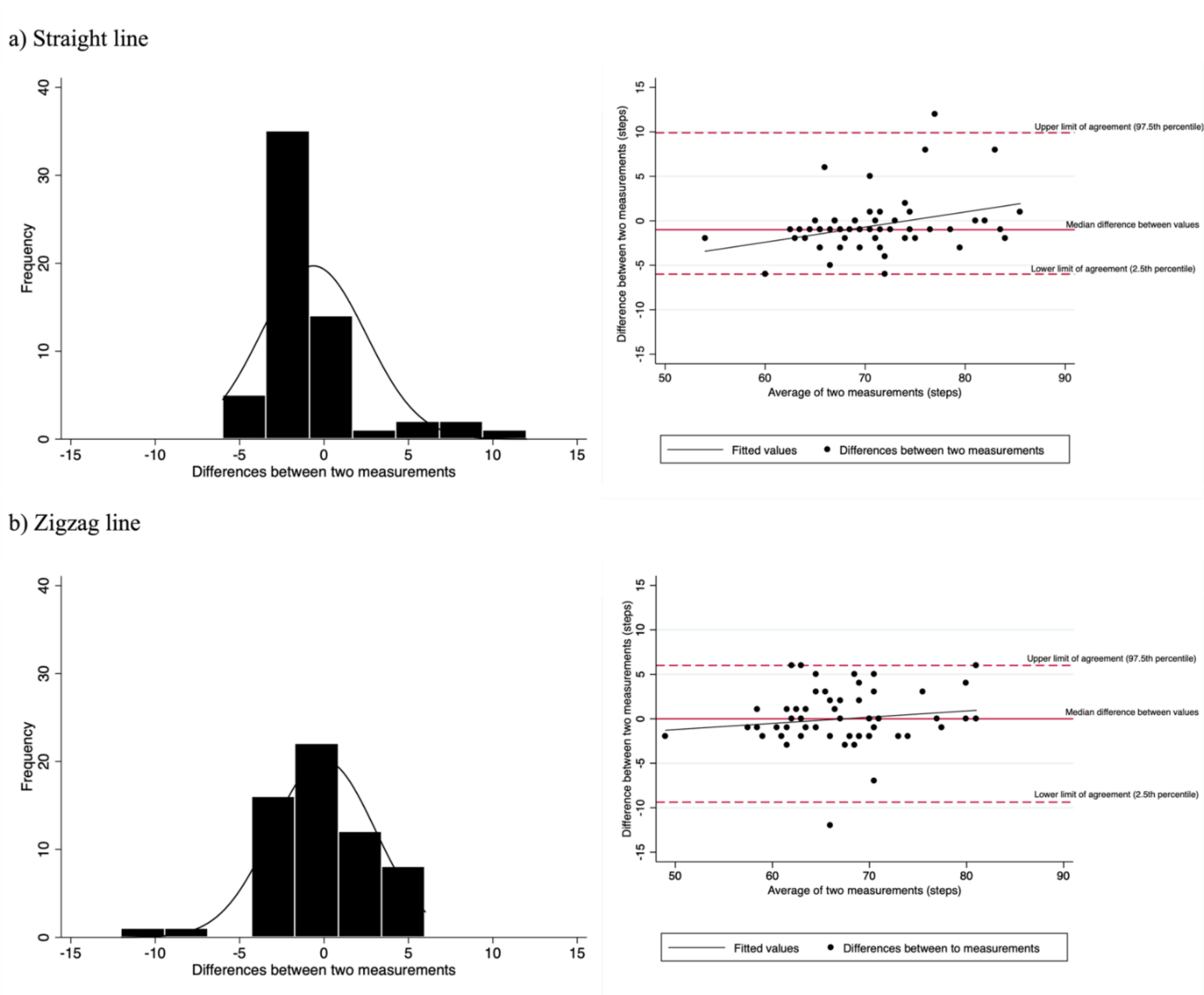

Figure 2. Linear regression, Bland-Altman and histogram plots for straight line (a) and zigzag line (b).

4. DISCUSSION

We aimed to investigate the validity of the iPhone's step count in 5 overground semi-free-living conditions (straight-line, zig-zag pattern, no slope, positive, and downward slopes). Regardless of the condition, our findings showed that iPhones tend, on average, to be highly precise in measuring step counts. The MAPE for different semi-free-living conditions was low, ranging from 2.37 % to 3.78 %. Previously published reports showed different results. For example, Duncan et al. 2018 [17] found a mean bias of ± 5 % in laboratory conditions (iPhone versus video record). Still, in the free-living condition, the mean bias increased to 21.5 % (iPhone versus Actigraph GTX3+). Additionally, another investigation [19] reported a bias of 12 % between the iPhone and the criterion (validated pedometer). Finally, measurement inaccuracies during intermittent walking have been found with MAPE ranging from 11.2-47.3 % during intermittent walking [20], which is not aligned with our results, suggesting a similar MAPE for the straight-line and the zigzag walking conditions. However, the discrepancies between studies may be due to different methodological strategies, such as applying other conditions, but mostly explained by the gold standard used (e.g., accelerometers and pedometers), which may not be considered a true reference method given their limitations [11].

In line with our findings, a recent investigation found that Android smartphones can be a promising alternative to measure steps, and they advocated that the step-counting algorithms require robust validation that accounts for temporal sensor body location, individual gait characteristics, and heterogeneous health states [23]. This study found that an open-source, step-counting method for smartphone data provided reliable step counts across sensor locations, measurement scenarios, and populations, including healthy adults and patients with cancer, with a mean bias of 0.1 % versus direct observation and a higher mean bias (3.4 %), when compared to a commercial wearable validation data set (i.e., Fitbit) [23]. The fact that smartphones operate different step-counting algorithms according to their brand and

Table 3. Descriptive statistics, Mean absolute percentage error (MAPE) percentages, and Wilcoxon sign-rank test for each condition in protocol 2.

N = 35	Median (IQR)	Min-max	MAPE (%)	Wilcoxon sign-rank test	
				z	p-value
Age (years)	21.0 (4.5)	18.0-40.0			
BMI (kg/m²)	24.1 (5.8)	19.6-32.1			
C1	51.0 (4.0)	35.0-73.0	2.41	3.551	**<0.001**
C2	50.0 (3.5)	42.0-72.0	2.37	1.660	0.097
C3	51.0 (3.0)	19.0-59.0	3.78	-0.087	0.931

BMI = body mass index; IQR = interquartile range; C1 = Step counts in straight-line with a downward slope of 5 %; C2 = Step counts in straight-line with no slope; C3 = Step counts in straight-line with an upward slope of 5 %. The bold means a significant difference from 50.

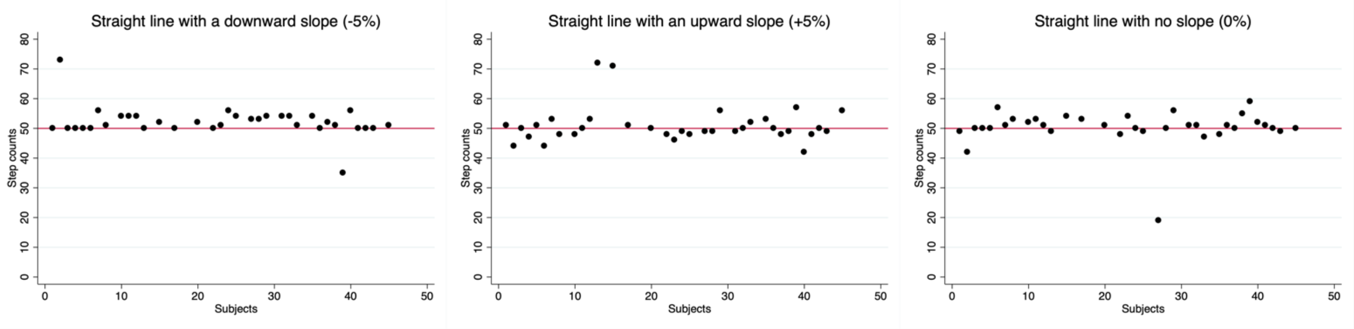

Figure 3. Illustrates the deviance between each estimated step count by the iPhone and the reference (50 steps). It is possible to observe that in the downward slope condition iPhone overestimate their counting, i.e., a large majority of individuals that deviated.

models and the non-open-source nature of these algorithms makes difficult comparisons between studies. Furthermore, data on the metrological characterization of smartphones are often missing or obtained with non-standardized methods, resulting in barely comparable results [14].

Comparisons between studies are, therefore, difficult. Moreover, available reports rarely exploited, within the same study, distinct semi-free-living conditions to examine the iPhone's validity in different conditions. For example, Hochsmann et al. (2018) investigated the validity of different devices in a walking course of 620.6 m with a total ascent of 12 m and a total descent of 9.4 m. They reported a MAPE of 3 % for the iPhone [13]. Yet, the authors did not count the steps separately in the different inclinations.

Nevertheless, the reported MAPE value is similar to our results. In any case, further studies are needed to expand this topic to other samples and brands of smartphones. The only significant difference between the iPhone's estimation of steps and the reference was found in the straight-line with a downward slope condition, with the iPhone consistently overestimating the number of steps. Current step-count estimation techniques use an accelerometer or gyroscope sensors to calculate the number of steps. However, due to smartphones' unfixed placement and direction, their accuracy can be affected [24]. We need to consider the impact of the carrying position on the accuracy of the pedometer algorithm. Even though all participants placed the iPhone in the pants side pocket, depending on the pocket size (i.e., which was not controlled), the smartphone could have been carried in various positions with a potential impact on the accuracy of measuring the number of steps [24]. Given the non-open-source nature of the algorithm within the smartphones, we do not know if the iPhone uses a carrying-position independent ensemble step-counting algorithm suitable for unconstrained smartphones in different carrying positions or if the algorithm does not comprise a classification algorithm that identifies the carrying position of the smartphone before the regression algorithm that considers the identified carrying position and calculates the number of steps [24]. If the second alternative may be the case, then the downward slope may slightly change the smartphone position in a way that can potentially explain the overestimation of step counting.

It is also possible that differences in speed from walking while descending or the higher impact of the feet on the floor can explain this overestimation of the number of steps by the iPhone. In contrast, a higher MAPE was found for the straight-line with an upward slope, which may be justified by the potentially lower speed in this condition [12]. There is evidence showing that the MAPE for iPhone and direct observation varied from 21 % for slow walking speed to 4 % for fast speeds [12]. Thus, a potential lower speed in the upward slope condition may be responsible for a slightly higher MAPE (3.78) in this condition, in comparison to the no slope (2.37) and the downward slope (2.41) conditions. Again, as previously explained, the unfixed placement and direction of the smartphone in the pocket may also impact the accuracy of measurement, which can be the case in the upward slope condition. A recent investigation trained neural network models on publicly available data and tested on an independent cohort using two approaches: generalization and personalization [25]. This study suggests that applying generalized and personalized deep learning on accelerometer signals may increase the accuracy to 96-99 %. Also, another investigation proposed a model that relies on four parameters (i.e., minimal peak distance, minimal peak prominence, dynamic thresholding, and vibration elimination) that seemed to solve the false walking problem [26]. These are two approaches that, together or alone, may potentially counteract or even eliminate the problem associated with the overestimation of step counting in the downward slope.

Besides these slight differences according to the specific conditions, our results confirmed previous evidence based on smartphones; that is, they seem to have a very good accelerometer system that can precisely estimate step counts [27]. Smartphone accelerometery provides better estimates of mobility and disability than a wrist-worn standard accelerometer in a free-living context [27]. Furthermore, by guaranteeing the accuracy of the iPhone to estimate step counts, we can use it for two aims: *i)* to monitor PA levels and *ii)* to implement intervention programs focused on increasing PA. Some studies attempted to understand the impact of smartphone Apps [8], [28], [29]. For example, in a systematic review, Romeo et al. (2019) [8] investigated the effectiveness of smartphone Apps for increasing objectively measured PA in adults and showed that smartphone Apps increased on average 476.75 steps/day. Also, Zhang et al. (2022) [29] in their recent systematic review and meta-analysis based on physically inactive individuals, reported that mobile health intervention improved PA and reduced sedentary behaviour among inactive individuals.

Taken together, our findings highlight the importance of including these devices in future research projects focused on PA epidemiology to help researchers use methods that are "closer to real life" (everyone uses smartphones in their daily routines), as well as in intervention programs designed to improve PA levels.

Notwithstanding the relevance of these results, this study has some limitations. First, the fact that participants were somehow homogeneous in terms of age. Ideally, these protocols must be replicated while considering more variability in terms of age, specifically by including older individuals, in which the walking pattern may impact the results for the validity of iPhones [27],

[30]. Although evidence suggests that age may not play such an important role in the validity of iPhones [13], we believe this issue must be further explored while considering older ages. Also, all participants were apparently healthy, without any walking constraints, and with an average healthy BMI. Thus, future studies must consider including people with distinct body composition profiles and disabilities to extensively generalize these findings. Furthermore, there is an inherent physiological variability when studying humans, and this variance is important to consider when assessing PA in the form of other outcomes coming from other wearables such as the smartwatches (e.g., combined information from heart rate and skin temperature). In the case of smartphones, they simply use data from the accelerometer sensors. Thus, this physiological variability may slightly impact the validity of smartphones. Regardless, other variables such as cardiorespiratory fitness levels of the participants (i.e., not measured in our study), can modify the walking pattern of someone in response to these conditions, which indirectly may impact the results for the validity of smartphones to measure steps. Moreover, using different samples in two different protocols requires caution in interpreting the results. More specifically, protocol 2 has fewer subjects than protocol 1 (60 vs 35). However, considering previous studies, we have a good sample size in both protocols. Previous studies, tend to have no more than 30 participants [10], [12], [13], [17]. Finally, we only considered one position for carrying the iPhone (i.e., pants side pocket). In real-life settings, people tend to carry their iPhones in other places (e.g., bags). However, a recent study found that the smartphone's position does not impact the accuracy of step detection, which rallies the versatility for PA assessment in research and clinical settings [13].

This study congregated several real-life conditions, such as overground walking with different inclinations or in a zigzag pattern, which are typical from free-living, while still using a valid method as the reference (i.e., direct observation with video recording), only customary in laboratorial studies [16], thus representing a strength. This semi-free-living approach is a strength, and experts in the field have pointed it out as the way to go [11]. Another strength was the sample size, which was significantly larger compared to previous laboratory and free-living studies [20], [31], and that further allowed us to examine the heterogeneity among the participants in addition to the overall sample (i.e., Bland-Altman analyses). In this sense, it is interesting to point out that our results suggest that walking with changes in direction and speed (i.e., closer to what happens in the real world) did not lower the validity of the iPhones to measure steps in comparison to the straight-line constant speed walking condition The MAPE for both conditions were similar (3.22 for the zigzag condition and 3.04 for the straight-line condition), as well as for Bland-Altman analysis.

A final message is crucial to this issue of being precise while monitoring PA (i.e., steps). Our results were promising and warranted a high validity for the iPhones to measure steps in real-life situations. Still, one must ponder that the participants carried their iPhones for the entire protocol (100 % of the time), which may not be the case when in free-living conditions. One investigation found that the largest underestimation of steps by the iPhone was observed among those who reported to have seldom carried their iPhones [19]. The participants seldom carrying their iPhones had a mean of -3036, SD 2990, steps/day; the ones sometimes carrying their iPhones had a mean of -1424, SD 2619, steps/day; and the ones almost always carrying their iPhones with a mean of -929, SD 1443, steps/day [19]. Thus, iPhones can only be valid in measuring steps if the person carries them. Data on step counts should be interpreted cautiously because of the possibility of underestimation due to non-carrying time [18], [19]. Other solutions in future free-living studies may be, for example, the smart socks used for detecting step counts, which have shown good validity even at slower walking speeds [32], or the smartwatch/smart band combination.

5. CONCLUSIONS

Our results reveal a high validity for the step counting of iPhones regardless of the inclination or the inconstant direction of the track, which is an important finding, once in real-life this will happen very often. Increased PA assisted by these devices may lead to clinical benefits, and these results may assist individuals' trust in using iPhones' applications to monitor steps, which could have important health implications. Future investigations should include other non-healthy groups of the population and older individuals to confirm these findings. Finally, iPhones seemed to slightly overestimate the number of steps while descending, which deserves further investigation.

REFERENCES

[1] A. Wahid, N. Manek, M. Nichols, P. Kelly, C. Foster, (+ 7 more authors), Quantifying the Association Between Physical Activity and Cardiovascular Disease and Diabetes: A Systematic Review and Meta-Analysis, J Am Heart Assoc. 2016;5(9). DOI: 10.1161/jaha.115.002495

[2] W. Geidl, S. Schlesinger, E. Mino, L. Miranda, K. Pfeifer, Dose-response relationship between physical activity and mortality in adults with noncommunicable diseases: a systematic review and meta-analysis of prospective observational studies, Int J Behav Nutr Phys Act. 2020;17(1):109. DOI: 10.1186/s12966-020-01007-5

[3] T. S. Church, D. M. Thomas, C. Tudor-Locke, P. T. Katzmarzyk, C. P. Earnest, (+ 4 more authors), Trends over 5 decades in U.S. occupation-related physical activity and their associations with obesity. PLoS One. 2011;6(5):e19657. DOI: 10.1371/journal.pone.0019657

[4] C. Fennell, A. Lepp, J. Barkley, Smartphone Use Predicts Being an "Active Couch Potato" in Sufficiently Active Adults, Am J Lifestyle Med 2021, 15, pp. 673-681. DOI: 10.1177/1559827619861383

[5] J. Bort-Roig, N. D. Gilson, A. Puig-Ribera, R. S. Contreras, S. G. Trost, Measuring and Influencing Physical Activity with Smartphone Technology: A Systematic Review, Sports Medicine. 2014;44(5):671-86. DOI: 10.1007/s40279-014-0142-5

[6] R. Daryabeygi-Khotbehsara, S. M. Shariful Islam, D. Dunstan, J. McVicar, M. Abdelrazek, R. Maddison, Smartphone-Based Interventions to Reduce Sedentary Behavior and Promote Physical Activity Using Integrated Dynamic Models: Systematic Review, J Med Internet Res. 2021;23(9):e26315. DOI: 10.2196/26315

[7] S. S. Intille, J. Lester, J. F. Sallis, G. Duncan, New Horizons in Sensor Development, Medicine & Science in Sports & Exercise 2012; 44, S24-S31. DOI 10.1249/MSS.0b013e3182399c7d

[8] A. Romeo, S. Edney, R. Plotnikoff, R. Curtis, J. Ryan, I. Sanders, A. Crozier, C. Maher, Can Smartphone Apps Increase Physical Activity? Systematic Review and Meta-Analysis, J Med Internet

Res. 2019;21(3):e12053.
DOI: 10.2196/12053
[9] L. Laranjo, D. Ding, B. Heleno, B. Kocaballi, J. C. Quiroz, (+ 10 more authors), Do smartphone applications and activity trackers increase physical activity in adults? Systematic review, meta-analysis and metaregression, Br J Sports Med. 2021;55(8):422-32. DOI: 10.1136/bjsports-2020-102892
[10] R. J. Bergman, J. W. Spellman, M. E. Hall, S. M. Bergman, Is there a valid app for that? Validity of a free pedometer iPhone application, J Phys Act Health. 2012;9(5):670-6.
DOI: 10.1123/jpah.9.5.670
[11] W. Johnston, P. B. Judice, P. Molina García, J. M. Mühlen, E. Lykke Skovgaard, (+ 10 more authors), Recommendations for determining the validity of consumer wearable and smartphone step count: expert statement and checklist of the INTERLIVE network, British Journal of Sports Medicine. 2021;55(14):780-93. DOI: 10.1136/bjsports-2020-103147
[12] M. J. Major, M. Alford, Validity of the iPhone M7 motion co-processor as a pedometer for able-bodied ambulation, Journal of Sports Sciences. 2016;34(23):2160-4.
DOI: 10.1080/02640414.2016.1189086
[13] C. Höchsmann, R. Knaier, J. Eymann, J. Hintermann, D. Infanger, A. Schmidt-Trucksäss, Validity of activity trackers, smartphones, and phone applications to measure steps in various walking conditions, Scand J Med Sci Sports. 2018;28(7):1818-27. DOI: 10.1111/sms.13074
[14] G. Cosoli, L. Antognoli, L. Scalise, Methods for the metrological characterization of wearable devices for the measurement of physiological signals: state of the art and future challenges, MethodsX. 2023;10:102038.
DOI: 10.1016/j.mex.2023.102038
[15] D. J. Macfarlane, C. C. Lee, E. Y. Ho, K. L. Chan, D. Chan, Convergent validity of six methods to assess physical activity in daily life, J Appl Physiol (1985). 2006;101(5):1328-34.
DOI: 10.1152/japplphysiol.00336.2006
[16] M. A. Case, H. A. Burwick, K. G. Volpp, M. S. Patel, Accuracy of smartphone applications and wearable devices for tracking physical activity data, Jama. 2015;313(6):625-6.
DOI: 10.1001/jama.2014.17841
[17] M. J. Duncan, K. Wunderlich, Y. Zhao, G. Faulkner, Walk this way: validity evidence of iphone health application step count in laboratory and free-living conditions, J Sports Sci. 2018;36(15):1695-704.
DOI: 10.1080/02640414.2017.1409855
[18] K. Orr, H. S. Howe, J. Omran, K. A. Smith, T. M. Palmateer, A. E. Ma, G. Faulkner, Validity of smartphone pedometer applications, BMC Res Notes. 2015;8:733.
DOI: 10.1186/s13104-015-1705-8
[19] S Amagasa, M Kamada, H Sasai, N Fukushima, H Kikuchi, I-M. Lee IM, Sh. Inoue, How Well iPhones Measure Steps in Free-Living Conditions: Cross-Sectional Validation Study, JMIR Mhealth Uhealth. 2019;7(1):e10418.
DOI: 10.2196/10418
[20] V. Hartung, M. Sarshar, V. Karle, L. Shammas, A. Rashid, (+ 6 more authors), Validity of Consumer Activity Monitors and an Algorithm Using Smartphone Data for Measuring Steps during Different Activity Types, Int J Environ Res Public Health. 2020;17(24).
DOI: 10.3390/ijerph17249314
[21] J. M. Bland, D. G. Altman, Statistical methods for assessing agreement between two methods of clinical measurement, Lancet. 1986;1(8476):307-10.
[22] J. M. Bland, D. G. Altman, Measuring agreement in method comparison studies, Stat Methods Med Res. 1999;8(2):135-60. DOI: 10.1177/096228029900800204
[23] M. Straczkiewicz, N. L. Keating, E. Thompson, U. A. Matulonis, S. M. Campos, A. A. Wright, J.-P. Onnela, Open-Source, Step-Counting Algorithm for Smartphone Data Collected in Clinical and Nonclinical Settings: Algorithm Development and Validation Study, JMIR Cancer. 2023;9:e47646.
DOI: 10.2196/47646
[24] Z. Song, H.-Y. Park, N. Thapa, J.-G. Yang, K. Harada, S. Lee, H. Shimada, H. Park, B.-K. Park, Carrying Position-Independent Ensemble Machine Learning Step-Counting Algorithm for Smartphones, Sensors (Basel). 2022;22(10).
DOI: 10.3390/s22103736
[25] L. Luu, A. Pillai, H. Lea, R. Buendia, F. M. Khan, G. Dennis, Accurate Step Count with Generalized and Personalized Deep Learning on Accelerometer Data, Sensors (Basel). 2022;22(11). DOI: 10.3390/s22113989
[26] V. T. Pham, D. A. Nguyen, N. D. Dang, H. H. Pham, V. A. Tran, K Sandrasegaran, D.-C. Tran, Highly Accurate Step Counting at Various Walking States Using Low-Cost Inertial Measurement Unit Support Indoor Positioning System. Sensors (Basel). 2018;18(10).
DOI: 10.3390/s18103186
[27] Y. Zhai, N. Nasseri, J. Pöttgen, E. Gezhelbash, C. Heesen, J. P. Stellmann, Smartphone Accelerometry: A Smart and Reliable Measurement of Real-Life Physical Activity in Multiple Sclerosis and Healthy Individuals, Front Neurol. 2020;11:688.
DOI: 10.3389/fneur.2020.00688
[28] P. Lunde, B. B. Nilsson, A. Bergland, K. J. Kværner, A. Bye, The Effectiveness of Smartphone Apps for Lifestyle Improvement in Noncommunicable Diseases: Systematic Review and Meta-Analyses, J Med Internet Res. 2018;20(5):e162.
DOI: 10.2196/jmir.9751
[29] M. Zhang, W. Wang, M. Li, H. Sheng, Y. Zhai, Efficacy of Mobile Health Applications to Improve Physical Activity and Sedentary Behavior: A Systematic Review and Meta-Analysis for Physically Inactive Individuals, Int J Environ Res Public Health. 2022;19(8). DOI: 10.3390/ijerph19084905
[30] 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;5(6):e88.
DOI: 10.2196/mhealth.7870
[31] B. Presset, B. Laurenczy, D. Malatesta, J. Barral, Accuracy of a smartphone pedometer application according to different speeds and mobile phone locations in a laboratory context, J Exerc Sci Fit. 2018;16(2):43-8.
DOI: 10.1016/j.jesf.2018.05.001
[32] B. N. Balmain, N. Tuttle, J. Bailey, K. Cheng, M. Duryea, J Houlihan, J. Wotherspoon, N. Morris, Using Smart Socks to Detect Step-count at Slow Walking Speeds in Healthy Adults, Int J Sports Med. 2019;40(2):133-8.
DOI: 10.1055/a-0732-5621