

## Article

# Apple Watch 6 vs. Galaxy Watch 4: A Validity Study of Step-Count Estimation in Daily Activities

Kyu-Ri Hong <sup>1</sup>, In-Whi Hwang <sup>2</sup>, Ho-Jun Kim <sup>3</sup>, Seo-Hyung Yang <sup>4</sup> and Jung-Min Lee <sup>5,6,\*</sup><sup>1</sup> Department of Physical Education, Graduate School of Education, Kyung Hee University, Yongin-si 17014, Republic of Korea; rbf12526@khu.ac.kr<sup>2</sup> Department of Sports Medicine and Science, Graduate School of Physical Education, Kyung Hee University, Yongin-si 17014, Republic of Korea; inn95313@khu.ac.kr<sup>3</sup> Department of Physical Education, College of Physical Education, Kyung Hee University, Yongin-si 17014, Republic of Korea; wnsghrla1@khu.ac.kr<sup>4</sup> School of Global Sport Studies, Korea University, 2511 Sejong-ro, Sejong City 30019, Republic of Korea; trusting2000@naver.com<sup>5</sup> Sports Science Research Center, Global Campus, Kyung Hee University, Yongin-si 17014, Republic of Korea<sup>6</sup> Department of Physical Education, Kyung Hee University, Yongin-si 17014, Republic of Korea

\* Correspondence: jungminlee@khu.ac.kr

**Abstract:** The purpose of this study was to examine the validity of two wearable smartwatches (the Apple Watch 6 (AW) and the Galaxy Watch 4 (GW)) and smartphone applications (Apple Health for iPhone mobiles and Samsung Health for Android mobiles) for estimating step counts in daily life. A total of 104 healthy adults (36 AW, 25 GW, and 43 smartphone application users) were engaged in daily activities for 24 h while wearing an ActivPAL accelerometer on the thigh and a smartwatch on the wrist. The validities of the smartwatch and smartphone estimates of step counts were evaluated relative to criterion values obtained from an ActivPAL accelerometer. The strongest relationship between the ActivPAL accelerometer and the devices was found for the AW ( $r = 0.99$ ,  $p < 0.001$ ), followed by the GW ( $r = 0.82$ ,  $p < 0.001$ ), and the smartphone applications ( $r = 0.93$ ,  $p < 0.001$ ). For overall group comparisons, the MAPE (Mean Absolute Percentage Error) values (computed as the average absolute value of the group-level errors) were 6.4%, 10.5%, and 29.6% for the AW, GW, and smartphone applications, respectively. The results of the present study indicate that the AW and GW showed strong validity in measuring steps, while the smartphone applications did not provide reliable step counts in free-living conditions.

**Keywords:** ActivPAL; Apple Watch 6; Galaxy Watch 4; wearable devices; step counts

**Citation:** Hong, K.-R.; Hwang, I.-W.; Kim, H.-J.; Yang, S.-H.; Lee, J.-M. Apple Watch 6 vs. Galaxy Watch 4: A Validity Study of Step-Count Estimation in Daily Activities. *Sensors* **2024**, *24*, 4658. <https://doi.org/10.3390/s24144658>

Academic Editors: Angelo Maria Sabatini, Paolo Visconti and Eugenio Santoro

Received: 24 May 2024

Revised: 6 July 2024

Accepted: 16 July 2024

Published: 18 July 2024



**Copyright:** © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Engaging in regular physical activity (PA) is a widely recognized foundational determinant for enhancing overall health [1]. It plays a pivotal role in disease prevention and the promotion of well-being across all age groups [2]. Despite the clear benefits of PA, a 2020 report issued by the World Health Organization (WHO) highlights a significant deficiency in meeting recommended PA guidelines, with a substantial segment of the population falling short of the minimum advised levels of moderate to vigorous PA [3]. Specifically, this deficiency pertains to individuals who do not attain a minimum of 150 min of moderate intensity PA or 75 min of vigorous intensity PA per week. The insufficiency in PA levels may constitute a contributory element to the prevalence of obesity [4] and a spectrum of associated diseases. Consequently, the waning prevalence of PA among adults exerts a pervasive influence on individuals throughout their lifespan and imposes substantial financial burdens on the healthcare system and society at large [5].

Smartphones were once pivotal tools for boosting regular PA through their various applications that measured activity levels. However, their accuracy did face some constraints

based on whether they involved direct contact or were non-contact [6]. Both Samsung and iPhone, for instance, featured built-in applications that tracked daily step counts. As technology advanced, smart wearables subsequently emerged, setting a new standard in personal health tracking. This category encompasses a spectrum of devices, including fitness trackers, smartwatches, heart rate monitors, and GPS tracking devices, which have revolutionized our approach to PA and holistic health management [7,8]. Among these technological innovations, smartwatches have emerged as pioneering instruments for the enhancement of individual health. Going beyond their original functions of timekeeping and notification delivery, smartwatches have evolved into versatile health management tools [2,9]. Equipped with an array of sensors, smartwatches enable users to engage in self-monitoring of various health parameters. Self-monitoring of PA stands as an efficacious strategy for improving PA levels [10,11]. Basic fitness tracking capabilities encompass the recording of steps taken, distance covered, and calories expended. Furthermore, contemporary smartwatches extend their capabilities to encompass more sophisticated health metrics, including heart rate, sleep quality, stress levels, and even oxygen saturation.

In the realm of wearable technology, the landscape is dominated by two prominent providers, namely Apple and Samsung. As a titan of the technology industry, Apple's market value holds the top position, while Samsung is ranked eighth [12]. According to a 2024 study by Counterpoint Research, Apple is the leading company in the smartwatch industry in terms of shipment share, accounting for 26 percent of worldwide shipments as of the first quarter of 2023. This was followed by Samsung with nine percent of the market share [13]. Moreover, since its launch in 2015, the Apple Watch has become the best-selling wearable/smartwatch worldwide, with sales surpassing 12 million units. Samsung has also seen an increasingly positive consumer response, recording second place [14]. With the widespread use of smartwatches, the scale of interventions for PA is also increasing [15]. These two formidable smartwatch offerings have exerted a progressively substantial influence across various domains, with a particular emphasis on their impact in the arena of health monitoring and management [9,16]. However, despite the considerable prevalence of these devices, the accuracy of step-count measurement in free-living conditions remains an underexplored area in the existing body of research.

The quantification of step counts serves as a fundamental and indispensable metric in assessing an individual's daily PA level. These step counts, conventionally monitored through devices such as pedometers and contemporary smartwatches, provide a cumulative measure of daily energy expenditure. There is a close and intricate relationship between step counts, the volume of PA, and its associated intensity [9,17]. Therefore, this study investigated step counts, one of the fundamental metrics of evaluating an individual's PA level in daily life. Recent research conducted by Chaudhry (2020) has elucidated that the utilization of pedometers or similar step-count tracking devices specifically designed for health monitoring leads to a notable improvements in recorded step counts compared to conventional care groups [18,19]. This improvement is made possible by the sensors embedded within modern smartphones and smartwatches, which are capable of detecting and recording human activities [20]. These sensors, which include tri-axis acceleration and tri-axis angular velocity measurements, are critical for the accurate quantification of an individual's step count [21].

In the current global context of increased smartwatch and smartphone usage and heightened interest in personalized health monitoring and management, it is essential to evaluate the validity of PA measurements by these devices. Previous research indicates that while smartwatches accurately monitor heart rate during activities like walking and running, their accuracy declines with increased intensity of activity [22,23]. Despite several studies addressing the accuracy of activity measurement by smartwatches, a comprehensive examination of step-count accuracy in free-living conditions remains lacking. Specifically, comparative studies on the accuracy of major smartwatches, such as the Apple Watch and Samsung Galaxy Watch, are scarce. Additionally, there is limited analysis of accuracy variations across different activity levels, which is essential for evaluating the practical

utility and validity of wearable devices. Therefore, this study aims to assess the accuracy of step counts measured by the Apple Watch and Samsung Galaxy Watch in daily life, comparing them with the ActivPAL accelerometer as a criterion measure. This research seeks to provide an in-depth evaluation of the practical application and contribution of smartwatches to health monitoring.

## 2. Materials and Methods

### 2.1. Participants

In this study, a total of 104 participants, comprising 36 Apple Watch 6 users (17 males and 19 females), 25 Galaxy Watch 4 users (10 males and 15 females), and 43 individuals (21 males and 22 females) who did not wear smartwatches (referred to as the “Smartphone Application” group), were included. The number of participants required to ensure sufficient statistical power was calculated using G\*Power 3.1.9.7 software and the effect size reported in previous studies [24]. G\*Power is a program used to calculate the necessary sample size before conducting a study. Based on an effect size (Cohen’s  $f$ ) of 0.25, a Type I error rate of 5% ( $\alpha$  error), and a power of 95% (power =  $1 - \beta$ ), the required total sample size was determined to be 45 participants. The participants, aged between 18 and 53 years, were recruited through word of mouth and direct approach. This study excluded variables that significantly affect PA in daily life, such as chronic diseases, physical disabilities, and pulmonary conditions, ensuring the selection of participants who could wear the devices without difficulty. In addition, if participants did not wear their smartwatches for more than 2 h a day, excluding sleep time, or if there was a discrepancy in the wearing periods between their ActivPAL devices and their smartwatches, the data were considered missing and excluded. To minimize any unreasonable discrepancies in measurements, participants were restricted from engaging in high-intensity activities and were instructed to record the time if they engaged in such activities. Each participant was instructed to wear both a smartwatch on their wrist and an ActivPAL device on their thigh continuously for 24 h. Furthermore, the smartphone application group consisted of participants who solely wore the ActivPAL device on their thighs. The number of steps of the participants in this group was measured by utilizing smartphone applications in conjunction with the ActivPAL device. To minimize error and ensure precise measurement, participants were instructed to continuously wear their smartwatches throughout each day, excluding sleep time, and to carry their mobile phones with them at all times. The Institutional Review Board of Kyung Hee University approved the study (KHGIRB-23-095).

### 2.2. Instruments

The study’s participants were instructed to utilize two distinct categories of activity monitoring devices. Each participant was required to wear a smartwatch on their wrist (Figure 1) in addition to wearing an ActivPAL device affixed to their thigh (Figure 2). In the case of the smartphone applications, they were held in the hand or carried in the participant’s clothes, while the smartphone-based accelerometer measured the step counts in the direction of the starting point of motion for each step [25]. Table 1 presents a comprehensive list of the devices used in this study and some extra information.

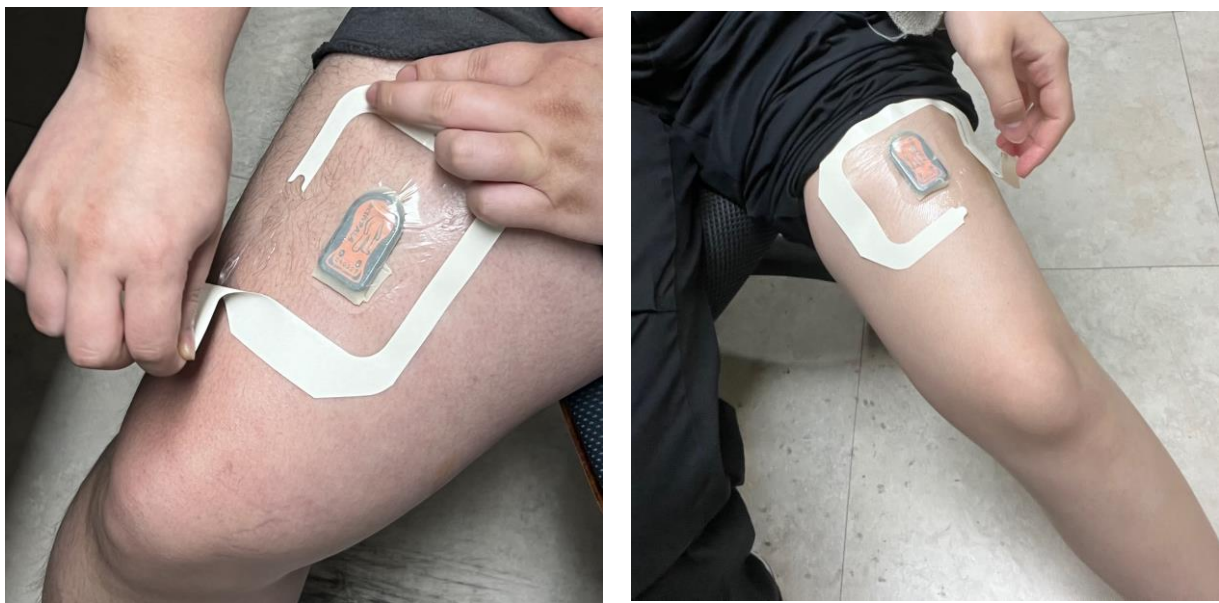
**Table 1.** A list of devices included in this study by versions used, locations worn, and software.

Device	Version	Location Worn	Software
ActivPAL (PAL Technologies Ltd., Glasgow, UK)	ActivPAL4	Thigh	ActivPAL4 v8.12.6
Apple Watch (Apple, Cupertino, CA, USA)	Apple Watch se2	Wrist	Apple Watch 6 v9.3.1
Galaxy Watch (Samsung, Suwon-si, Republic of Korea)	Galaxy Watch 4	Wrist	Galaxy Watch 4 v2.2.11.23082851
Apple iPhone (Apple, Cupertino, CA, USA)	iPhone	-	iPhone app; Health (iOS 16)
Samsung Galaxy Phone (Samsung, Suwon-si, Republic of Korea)	Galaxy Phone	-	Galaxy Phone app; Samsung Health (UI 5.0)

app: application.



**Figure 1.** Images showing how the Apple Watch (Left) and Galaxy Watch (Right) are worn on the wrist.



**Figure 2.** Images showing how the ActivPAL accelerometer is worn on the thigh.

#### 2.2.1. ActivPAL (ActivPAL 4 v8.12.6) (AP)

The ActivPAL accelerometer is a small capacitance-based accelerometer (5.3 cm × 3.5 cm × 0.7 cm) and light (15 g). During the study, the ActivPAL device was affixed to the thigh using hypoallergenic tape and was not detached except in special or hazardous situations. It could later be verified through the data analysis process. This procedure minimized errors in the criterion values. The ActivPAL device functions as an instrument for analyzing PA based on body positioning or morphology. Precisely, it possesses the capability to discern between three distinct activity states: sitting/lying, standing, and stepping, as expounded upon in reference [26]. The ActivPAL accelerometer was selected as the criterion measure for this study due to its proven accuracy and reliability in measuring PA and stepping activities [27,28]. The ActivPAL accelerometer is widely recognized in the field of PA research for its high validity in various settings, including free-living conditions. Additionally, it has been extensively validated against other gold-standard



methods, making it an ideal benchmark for evaluating the performance of consumer-grade wearable devices such as smartwatches.

### 2.2.2. Smartwatches (Apple Watch 6 and Galaxy Watch 4)

In this study, we evaluated the functionality of two distinct smart wearables devices: the Apple Watch 6 and the Galaxy Watch 4, occupying the highest ranks in the current technology market values [12]. This study selected the latest versions available at the time of the research to investigate the accuracy of their measurements. The Apple Watch Series 6 and the Galaxy Watch 4 are equipped with an array of sensors designed to monitor various health metrics. The Apple Watch Series 6 includes an accelerometer, gyroscope, heart rate monitor, and blood oxygen (SpO<sub>2</sub>) sensor. These sensors enable the device to track physical activity, detect irregular heart rhythms, and measure blood oxygen levels [29]. The Galaxy Watch 4 also features an accelerometer, gyroscope, heart rate monitor, and body composition sensor, which provides additional insights into the user's health by estimating metrics such as body fat percentage and skeletal muscle mass. The accelerometer and gyroscope in both devices allow for the detection of movement and orientation, which are crucial for accurately counting steps and assessing physical activity levels. The heart rate monitors use photoplethysmography (PPG) to measure the user's pulse by detecting blood flow changes in the wrist. The SpO<sub>2</sub> sensor in the Apple Watch and the body composition sensor in the Galaxy Watch enhance the devices' abilities to provide comprehensive health monitoring. By incorporating these advanced sensors, both smartwatches offer robust health tracking capabilities that are critical for the accurate measurement of step counts and other physical activity metrics. This could provide valuable information to a wide range of users. These devices are seamlessly integrated with their respective smartphones, the iPhone and Samsung Galaxy phone, facilitating the effortless visualization of data through mobile applications.

Both wearable devices are equipped with the capability to accurately quantify step counts, relying on tri-axis acceleration and tri-axis angular velocity measurements [30] to discern various physical activities, including walking and running. This capacity enables the smartwatches to effectively monitor and record step counts by analyzing the user's PA. Furthermore, wearable sensor technology advancements are closely linked to the development of low-power soft transistors and flexible electronic devices. Liu et al. (2024) demonstrated that low-power soft transistors could greatly enhance the efficiency of wearable devices by reducing energy consumption and maintaining stable performance over extended periods [31]. Additionally, Zhang et al. (2023) explored how flexible electronic devices could be utilized in cardiovascular healthcare monitoring, explaining that the application of flexible electronics plays a crucial role in improving the reliability and accuracy of wearable devices [32]. Thus, the accurate step-count measurement capabilities of the Apple Watch 6 and Galaxy Watch 4 are further enhanced by the application of the latest low-power soft transistor and flexible electronics technologies. The collected step data are automatically synchronized with the corresponding mobile application; namely, Apple Watch's Health application, and Galaxy Watch's Samsung Health application. All study participants were instructed to wear these smartwatches continuously for each 24 h period, with exceptions made solely in cases of health-related emergencies.

### 2.3. Study Protocol

At the beginning of the data acquisition, anthropometric measurements were ascertained. Participants' height and weight were garnered through self-report health history, subsequently permitting the calculation of their Body Mass Index (BMI) as the ratio of weight (kg) to height squared (m<sup>2</sup>). Subsequently, the ActivPAL device was affixed to the participants' thighs as per the manufacturer's recommendation. While both wearing the ActivPAL device and either wearing a smartwatch or carrying a smartphone featuring a health application, participants' step counts were recorded in free-living conditions for 24 h. During a follow-up meeting, the ActivPAL device and the amassed data were retrieved. To

enhance the precision and reliability of the study's outcomes, an assessment was conducted to identify instances where the smartwatch had been detached for more than one hour, and such instances were categorized as missing data values.

#### 2.4. Data Analysis

All data processing and statistical analyses in this study were conducted using SPSS version 28.0 (SPSS Inc., Chicago, IL, USA). Descriptive statistics were employed to succinctly summarize the demographic characteristics of the study participants, including gender, age, height, weight, and BMI. To evaluate the correlation between the smartwatch (or smartphone application) and the ActivPAL accelerometer in measuring step counts, Kendall's tau-b and Pearson's correlation coefficients were computed. The Mean Absolute Percent Error (MAPE) was calculated as the average of the absolute difference between the criterion measure value (as measured by the ActivPAL accelerometer) and the values recorded by each device, divided by the criterion measure value, then multiplied by 100. Furthermore, a paired-sample *t*-test was utilized to identify any significant mean differences between the ActivPAL accelerometer and each respective device. In analyzing the step-count data from all the devices, we applied equivalence testing, a novel statistical approach designed to assess the equivalence between diverse measures, as opposed to solely testing against a null hypothesis of zero difference [33]. To examine the proportional systematic biases, Bland–Altman plots were utilized. These plots incorporate 95% limits of agreement and fitted lines derived from regression analyses comparing the criterion measure to the differences in values obtained from the ActivPAL accelerometer and each other device. Parameters such as intercept and slope were included in the analyses to provide a comprehensive evaluation of the agreement between the devices.

### 3. Results

A total of 104 participants (48 male and 56 female) were involved in this study. All participants had smartwatches or smartphones. Table 2 presents comprehensive descriptive statistics summarizing the anthropometric characteristics of all the study participants.

**Table 2.** Physical characteristics of participants ( $n = 104$ ).

Variables		Apple Watch		Galaxy Watch		Smartphone Applications	
		No. (%)	Mean (SD)	No. (%)	Mean (SD)	No. (%)	Mean (SD)
Gender	Male	17 (47.2%)	-	10 (40%)	-	21 (48.8%)	-
	Female	19 (52.8%)	-	15 (60%)	-	22 (51.2%)	-
Anthropometrics	Male	Age (year)	25.7 (5.6)		31.1 (11.8)		27.9 (8.6)
		Height (cm)	175.2 (0.1)		177.0 (0.0)		176.0 (0.1)
		Weight (kg)	73.9 (6.9)	10	82.7 (13.1)	21	77.4 (8.7)
		BMI (kg/m <sup>2</sup> )	24.1 (1.6)		26.4 (4.3)		25.0 (2.5)
	Female	Age (year)	23.3 (3.6)		27.3 (10.2)		25.0 (5.1)
		Height (cm)	163.8 (0.0)		165.9 (0.0)		163.8 (0.0)
		Weight (kg)	50.3 (3.6)	15	52.3 (1.8)	22	52.6 (3.7)
		BMI (kg/m <sup>2</sup> )	18.8 (1.4)		19.0 (1.3)		19.6 (1.3)
	Total	Age (year)	24.4 (4.7)		28.8 (10.8)		26.4 (7.1)
		Height (cm)	169.2 (0.1)		170.4 (0.1)		169.7 (0.1)
		Weight (kg)	61.5 (13.1)	25	64.5 (17.2)	43	64.7 (14.1)
		BMI (kg/m <sup>2</sup> )	21.3 (3.1)		22.0 (4.6)		22.2 (3.4)

BMI: Body Mass Index; SD: standard deviation.

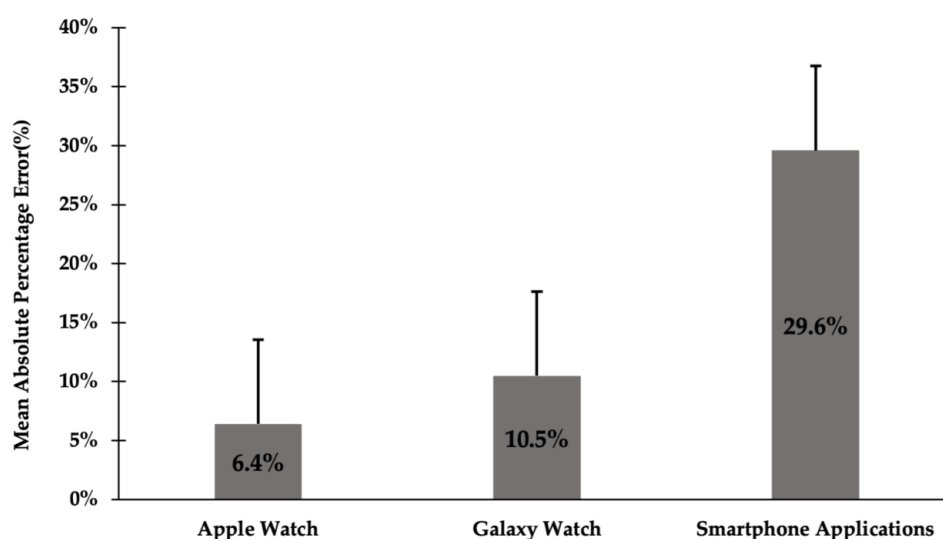
Meanwhile, Table 3 displays the correlation coefficients (*r*) representing the relationships between the ActivPAL accelerometer and three distinct devices; the Apple Watch 6, Galaxy Watch 4, and smartphone applications. These devices exhibited substantial correlations with the ActivPAL device (Apple Watch 6:  $r = 0.986$ ,  $p < 0.01$ ; Galaxy Watch 4:  $r = 0.824$ ,  $p < 0.01$ ; and smartphone applications:  $r = 0.926$ ,  $p < 0.01$ ), with the strongest association observed between the ActivPAL device and the Apple Watch 6.

**Table 3.** Correlations between the ActivPAL accelerometer and the Apple Watch, Galaxy Watch, and non-wearable smartphone applications ( $n = 104$ ).

	ActivPAL	Apple Watch	Correlation Galaxy Watch	Smartphone Applications
ActivPAL	1	0.986 **	0.824 **	0.926 **

\*\* Correlation is significant at the 0.01 level (2-tailed).

Figure 3 shows the MAPE for each device, computed as the average absolute error relative to the ActivPAL criterion. The smallest errors were associated with the Apple Watch 6 (6.4%), followed by the Galaxy Watch 4 (10.5%), and the smartphone applications (29.6%). These findings underscore the superior accuracy of smartwatches in comparison to smartphone applications, with the Apple Watch 6 demonstrating the lowest MAPE and, consequently, the highest accuracy.

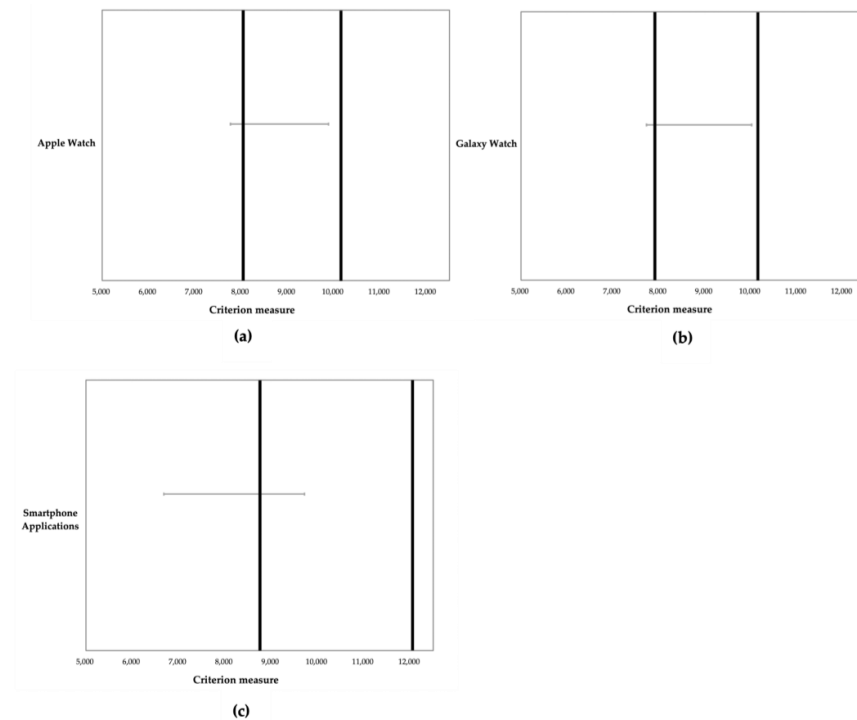


**Figure 3.** Mean absolute percentage error ( $\pm$ SD) for all devices; Apple Watch, Galaxy Watch, and smartphone applications.

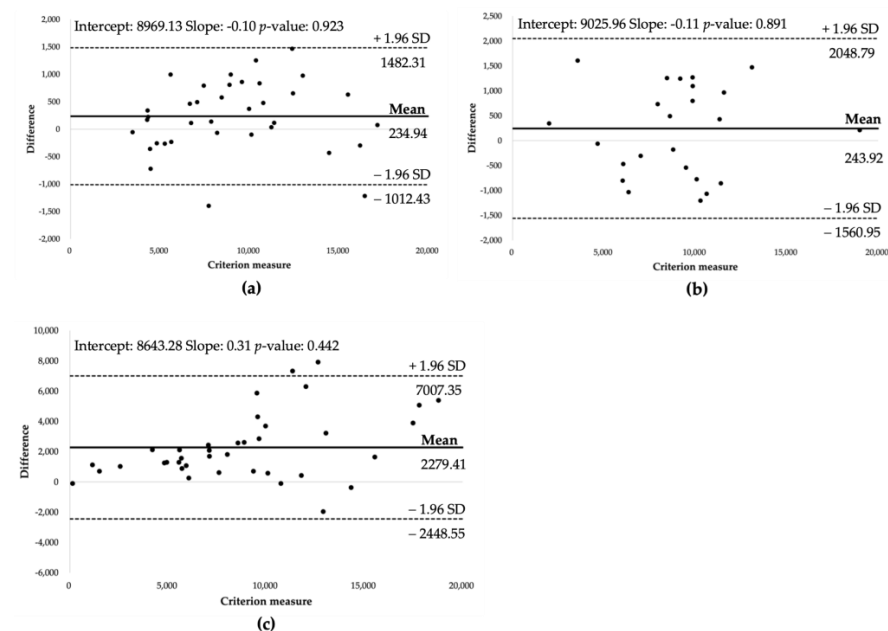
The results from equivalence testing are delineated in Figure 4, demonstrating the substantial equivalence of the smartwatches with the criterion measure. Specifically, the step counts recorded by the Apple Watch 6 and the Galaxy Watch 4 largely fell within the equivalence zone (90% confidence interval within the equivalence zone: Apple Watch 6 = 7770.0–9886.1 and Galaxy Watch 4 = 7732.6–10022.9). Conversely, the step counts from the smartphone applications were outside the equivalence zone (smartphone applications' 90% confidence interval outside of the equivalence zone: 6681.0–9718.2).

Bland–Altman plot analyses, as displayed in Figure 5, were carried out to assess the distribution of errors and identify the potential proportional systematic biases in the estimates. These plots show the differences between the criterion measure (the ActivPAL device) and each device on the Y-axis, plotted against the average step counts of the criterion and each device on the X-axis. Notably, the narrowest 95% limits of agreement were observed for the Apple Watch 6 (difference = 2494.7), with slightly wider limits for the Galaxy Watch 4 (difference = 3609.7). In contrast, the smartphone applications exhibited the widest limits (difference = 9455.9). The distribution of points around the mean was more concentrated for the Apple Watch 6, followed by the Galaxy Watch 4, while the smartphone applications displayed more dispersed data points. Furthermore, the fitted line slopes for all the devices were found to be statistically nonsignificant (Apple Watch 6: slope =  $-0.10$ ,  $p = 0.923$ ; Galaxy Watch 4: slope =  $-0.11$ ,  $p = 0.891$ ; and smartphone

applications: slope = 0.31,  $p = 0.442$ ), indicating the absence of substantial patterns of proportional systematic bias with these devices.



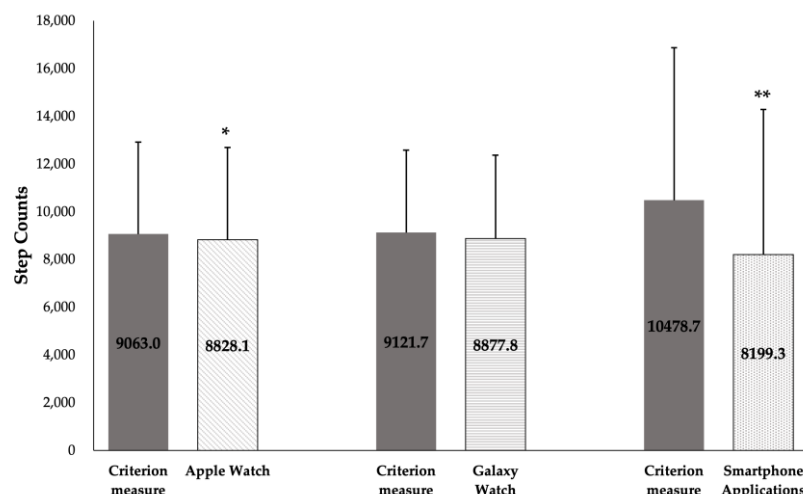
**Figure 4.** Results from 90% equivalence testing for agreement between the ActivPAL accelerometer and the devices: (a) Apple Watch; (b) Galaxy Watch; and (c) smartphone applications. Dark lines indicate the proposed equivalence zone ( $\pm 10\%$  of the mean); Grey bars indicate the 90% confidence interval for the means of the devices.



**Figure 5.** Bland–Altman plots showing the agreement of step counts between ActivPAL accelerometer and each device: (a) Apple Watch; (b) Galaxy Watch; and (c) smartphone applications. Dots represent the difference between the step count measured by each device and the criterion value measured by ActivPAL.



Figure 6 illustrates the outcomes of a paired-sample *t*-test comparing the criterion measure (the ActivPAL accelerometer) against the various devices. This analysis indicates a lack of statistically significant difference between the Galaxy Watch 4 and the criterion measure ( $p = 0.198$ ), with a small effect size ( $d = 0.265$ ). However, notable disparities were observed between the criterion measure and the Apple Watch 6 ( $p < 0.05$ ,  $p = 0.033$ ), reflecting a small effect size ( $d = 0.369$ ), as well as with the smartphone applications ( $p < 0.01$ ), with a large effect size ( $d = 0.945$ ).



**Figure 6.** Paired-sample *t*-test ( $\pm$ SD) between criterion measure (the ActivPAL accelerometer) and each device; Apple Watch, Galaxy Watch, and smartphone applications. \*  $p$ -value  $< 0.05$ , \*\*  $p$ -value  $< 0.01$ .

#### 4. Discussion

This is a valuable study to assess the validity of the Apple Watch, the Galaxy Watch, and smartphone applications for measuring step counts compared to the criterion measure of the ActivPAL accelerometer under free-living conditions. Previous studies have also documented the measurement accuracy of the smartwatches, particularly in the reliable quantification of heart rate and energy expenditure [34,35], and it is evident that the measurement accuracy of the smartwatches varies with the intensity of the exercise [22,23]. Additionally, according to prior studies that measured the Apple Watch's step counts on a treadmill at varying speeds, the most accurate readings occurred at moderate pace, with the tendency to undercount or overcount at faster or slower speeds, respectively [36]. However, questions remain regarding the details of how the Apple Watch and the Galaxy Watch measure PA [15]. Furthermore, research is required related to measuring PA in free-living conditions, not fixed to any specific intensity. Therefore, the present study assists smartwatch users in accurately gauging their PA levels in daily life, and the significance lies in the ability to increase exercise and enhance overall health based on such precise data.

Numerous studies have been conducted to assess variations in measurement accuracy associated with different types of research-oriented activity monitors [37–39]. Hergenroeder et al. conducted a study wherein participants underwent two 100-step walking tests while simultaneously wearing three research-grade activity monitors. The step-count accuracy of these devices was evaluated by comparison with a manual step-count hand tally [38]. Their findings indicated that the ActivPAL accelerometer, when positioned on the thigh, consistently exhibited a high level of accuracy, accounting for  $97.3 \pm 11.1\%$  of manually tallied steps; in contrast to the ActiGraph GT9X, which, when worn at the waist, exhibited a considerably wider range of measurement variance ( $51.4 \pm 35.7\%$ ) in relation to the criterion measure. Additionally, the study highlighted that the ActivPAL accelerometer provided a reasonable degree of accuracy in measuring gait speed, with values ranging from  $86.8 \pm 14.0\%$  to  $95.1 \pm 9.2\%$  [38]. This study's findings are consistent with the body of literature supporting the accuracy of the ActivPAL accelerometer. In a study by An et al.

titled “Accuracy of Inclinometer Functions of the ActivPAL and ActiGraph GT3X+: A Focus on Physical Activity”, the ActivPAL device was shown to exhibit high accuracy for measuring stepping and it proved to be significantly accurate in detecting various forms of PA, including sitting, standing, and stepping [27]. In the current study, the ActivPAL device served as the criterion measure, reinforcing the study’s robust validity.

Comparative analyses involving smartwatches, namely the Apple Watch and Galaxy Watch, exhibit a notable degree of concordance with the criterion measure, accompanied by low MAPE values. Specifically, step-count data originating from the Apple Watch demonstrate a statistically significant equivalence with the criterion measure and an MAPE below 10%, signifying a high level of accuracy, which is consistent with prior investigations suggesting that the Apple Watch boasts superior accuracy compared to the established criteria [40,41]. Notably, MAPE values below the 10% threshold, as recorded by the Apple Watch, are indicative of a robust level of accuracy [42]. Furthermore, the Apple Watch exhibits a statistically significant correlation with the criterion measure ( $p < 0.01$ ). While the Apple Watch demonstrates the most favorable overall performance, it is noteworthy that the Galaxy Watch yields commendable results within this study. The Galaxy Watch displays a statistically significant correlation with the criterion measure ( $p < 0.01$ ). It records an MAPE value of 10.5%, a figure closely approaching the 10% threshold, indicative of a high degree of accuracy [42]. It is worth mentioning that only the Galaxy Watch exhibited no significant difference compared to the criterion measure.

Ultimately, the smartphone applications also showed a pronounced correlation with the ActivPAL device. Existing research indicates that, regardless of the smartphone’s position, whether it is in a bag or a pants pocket, high accuracy and generally low variability are demonstrated. Additionally, all accelerometer apps tend to systematically underestimate steps during free-living walking activities [43]. However, in this study they yielded notably lower accuracy in measuring the number of steps. Specifically, the MAPE of the smartphone applications showed the highest value at 29.6%, deviating considerably from the comparable range within the equivalence zone. In comparison to the smartwatches this discrepancy was substantial. Previous research corroborates these findings, as exemplified by Presset et al. [44], who reported diminished step-count accuracy of smartphone applications when smartphones were positioned in the most comfortable locations, such as within a “jacket”. Notably, for the Runtastic Pedometer smartphone application, the ‘jacket’ placement ranked as the least accurate among the three positions assessed (i.e., arm, belt, and jacket) [44]. This suggests that smartphones, which must be carried in clothing or bags, could show significant discrepancies in accuracy compared to smartwatches, which are attached to the body. Compared to the previous study, which was conducted under controlled conditions, such as treadmill settings, the current study proceeded within unconstrained free-living conditions, which evidently underscored the diminished accuracy inherent in smartphone applications for step-count quantification.

This is one of the first studies to compare the Apple and Galaxy Watches in free-living conditions. Notably, these smartwatches are globally recognized and extensively utilized for fitness and health tracking due to their acknowledged precision and widespread user acceptance. Furthermore, the current research differs significantly from previous studies, which predominantly compared smartwatches to wrist-worn accelerometers, such as the ActiGraph GT9X and Fitbit, as the criterion measures [2,45]. Instead of these devices, we employed the ActivPAL accelerometer, known for its minimal margin of error in measuring PA. Lastly, our study was conducted in the context of uncontrolled daily life, which was a distinguishing feature. While numerous studies have revealed the accuracy of smartwatches in step counting within controlled environments like laboratories and treadmills, the literature is less populated with studies addressing step counting during daily activities. Given the ubiquity of smartwatch usage in individuals’ everyday lives, our findings assume heightened relevance and applicability. They stand poised to offer valuable insights for self-monitoring of PA and health through smartwatch technology, potentially highlighting meaningful implications for health-conscious individuals seeking

to leverage these devices for wellness management. The authors acknowledge several limitations of this study. Firstly, the small sample sizes employed in the research must be recognized. Data collection in an uncontrolled environment resulted in the exclusion of numerous data points, particularly affecting the sample size for the Galaxy Watch, which was less than 30. Despite conducting a normality test, the constrained sample size inevitably restricted the study's generalizability. The small sample size unavoidably constrained the study. Future research should aim to gather larger and more representative samples to yield more comprehensive outcomes and measure a broader spectrum of physical activities, including factors such as sleep patterns and sedentary behaviors, to enhance the study's overall robustness and applicability.

## 5. Conclusions

In conclusion, we examined the accuracy of step-counting functionality across the Apple Watch 6, Galaxy Watch 4, and smartphone applications. The results unequivocally establish the Apple Watch 6 as the most accurate among the devices, a distinction supported by statistically significant findings. Furthermore, the Galaxy Watch 4 also demonstrated a level of step-counting accuracy akin to that of the Apple Watch 6. Conversely, smartphone applications lagged in terms of accuracy. Consequently, both the Apple Watch 6 and Galaxy Watch 4 are expected to provide data with sufficient accuracy to allow users to plan, measure, and enhance their PA in everyday life. On the other hand, data provided by smartphone applications demonstrated lower validity and reliability for contributing to PA. These findings underscore the value and validity of smartwatches as effective tools for routine self-monitoring of daily steps and the management of PA.

**Author Contributions:** K.-R.H., conceptualization, data curation, methodology, writing—original draft preparation, and supervision; I.-W.H., validation, visualization, writing—original draft preparation, and co-supervision; H.-J.K., formal analysis, data curation, and writing—original draft preparation; S.-H.Y., validation, and writing—review and editing; J.-M.L., validation, data curation, and writing—review and editing. All authors have read and agreed to the published version of the manuscript.

**Funding:** This study was supported by the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea (NRF-2021S1A5A806440412).

**Institutional Review Board Statement:** All subjects gave their informed consent for inclusion before they participated in the study. The study was conducted in accordance with the Declaration of Helsinki, and the Institutional Review Board of Kyung Hee University approved the study (KHGIRB-23-095).

**Informed Consent Statement:** Informed consent has been obtained from the healthy participants to publish this paper.

**Data Availability Statement:** The data presented in this study are available on request from the corresponding author.

**Acknowledgments:** We express our sincere appreciation to the anonymous referees for their valuable and thoughtful feedback, which greatly enhanced the quality of this paper. Additionally, we extend our gratitude to the editorial board and the staff for their support during the publication process.

**Conflicts of Interest:** The authors declare no conflicts of interest.

## References

1. Rhodes, R.E.; Janssen, I.; Bredin, S.S.D.; Warburton, D.E.R.; Bauman, A. Physical activity: Health impact, prevalence, correlates and interventions. *Psychol. Health* **2017**, *32*, 942–975. [CrossRef] [PubMed]
2. Degroote, L.; De Bourdeaudhuij, I.; Verloigne, M.; Poppe, L.; Crombez, G. The Accuracy of Smart Devices for Measuring Physical Activity in Daily Life: Validation Study. *JMIR mHealth uHealth* **2018**, *6*, e10972. [CrossRef] [PubMed]
3. World Health Organization. *Global Status Report on Physical Activity 2022: Country Profiles*; World Health Organization: Geneva, Switzerland, 2022. Available online: <https://www.who.int/publications/i/item/9789240064119> (accessed on 10 July 2022).
4. Bassett, D.R., Jr.; Wyatt, H.R.; Thompson, H.; Peters, J.C.; Hill, J.O. Pedometer-measured physical activity and health behaviors in United States adults. *Med. Sci. Sports Exerc.* **2010**, *42*, 1819–1825. [CrossRef] [PubMed]

5. World Health Organization. Global Status Report on Physical Activity 2022. Available online: <https://www.who.int/teams/health-promotion/physical-activity/global-status-report-on-physical-activity-2022> (accessed on 6 July 2022).
6. Coppetti, T.; Brauchlin, A.; Müggler, S.; Attinger-Toller, A.; Templin, C.; Schönrrath, F.; Hellermann, J.; Lüscher, T.F.; Biaggi, P.; Wyss, C.A. Accuracy of smartphone apps for heart rate measurement. *Eur. J. Prev. Cardiol.* **2017**, *24*, 1287–1293. [[CrossRef](#)] [[PubMed](#)]
7. Chan, M.; Estève, D.; Fourniols, J.-Y.; Escriba, C.; Campo, E. Smart wearable systems: Current status and future challenges. *Artif. Intell. Med.* **2012**, *56*, 137–156. [[CrossRef](#)] [[PubMed](#)]
8. Appelboom, G.; Camacho, E.; Abraham, M.E.; Bruce, S.S.; Dumont, E.L.; Zacharia, B.E.; D’Amico, R.; Slomian, J.; Reginster, J.Y.; Bruyère, O. Smart wearable body sensors for patient self-assessment and monitoring. *Arch. Public Health* **2014**, *72*, 28. [[CrossRef](#)] [[PubMed](#)]
9. Li, J.; Ma, Q.; Chan, A.H.S.; Man, S.S. Health monitoring through wearable technologies for older adults: Smart wearables acceptance model. *Appl. Ergon.* **2019**, *75*, 162–169. [[CrossRef](#)] [[PubMed](#)]
10. Michie, S.; Abraham, C.; Whittington, C.; McAteer, J.; Gupta, S. Effective techniques in healthy eating and physical activity interventions: A meta-regression. *Health Psychol* **2009**, *28*, 690–701. [[CrossRef](#)]
11. Williams, S.L.; French, D.P. What are the most effective intervention techniques for changing physical activity self-efficacy and physical activity behaviour—And are they the same? *Health Educ. Res.* **2011**, *26*, 308–322. [[CrossRef](#)] [[PubMed](#)]
12. Forbes. The World’s Most Valuable Brands. Available online: <https://www.forbes.com/powerful-brands/list/> (accessed on 17 April 2023).
13. Research, C. Quarterly Smartwatch Unit Shipment Share Worldwide from 2018 to 2023, by Vendor. Available online: <https://www.statista.com/statistics/910862/worldwide-smartwatch-shipment-market-share/#statisticContainer> (accessed on 17 April 2023).
14. Canalys. Apple Shipped Two-Thirds of All Smart Watches in 2015. Available online: [https://www.canalys.com/static/press\\_release/2016/media-alert-05022016-apple-shipped-two-thirds-all-smart-watches-2015.pdf](https://www.canalys.com/static/press_release/2016/media-alert-05022016-apple-shipped-two-thirds-all-smart-watches-2015.pdf) (accessed on 17 April 2023).
15. Abt, G.; Bray, J.; Benson, A.C. Measuring Moderate-Intensity Exercise with the Apple Watch: Validation Study. *JMIR Cardio* **2018**, *2*, e6. [[CrossRef](#)]
16. Raja, J.M.; Elsakar, C.; Roman, S.; Cave, B.; Pour-Ghaz, I.; Nanda, A.; Maturana, M.; Khouzam, R.N. Apple watch, wearables, and heart rhythm: Where do we stand? *Ann. Transl. Med.* **2019**, *7*, 417–422. [[CrossRef](#)] [[PubMed](#)]
17. Kraus, W.E.; Janz, K.F.; Powell, K.E.; Campbell, W.W.; Jakicic, J.M.; Troiano, R.P.; Sprow, K.; Torres, A.; Piercy, K.L. Daily Step Counts for Measuring Physical Activity Exposure and Its Relation to Health. *Med. Sci. Sports Exerc.* **2019**, *51*, 1206–1212. [[CrossRef](#)] [[PubMed](#)]
18. Chaudhry, U.A.R.; Wahlich, C.; Fortescue, R.; Cook, D.G.; Knightly, R.; Harris, T. The effects of step-count monitoring interventions on physical activity: Systematic review and meta-analysis of community-based randomised controlled trials in adults. *Int. J. Behav. Nutr. Phys. Act.* **2020**, *17*, 129–144. [[CrossRef](#)] [[PubMed](#)]
19. Masi, E.; Peterman, J.E.; Kaminsky, L.A. The health benefits of a pedometer-based 100,000 steps/Week Physical Activity Program. *J. Sci. Sport Exerc.* **2019**, *1*, 176–183. [[CrossRef](#)]
20. Ahmed Bhuiyan, R.; Ahmed, N.; Amiruzzaman, M.; Islam, M.R. A robust feature extraction model for human activity characterization using 3-axis accelerometer and gyroscope data. *Sensors* **2020**, *20*, 6990. [[CrossRef](#)] [[PubMed](#)]
21. Wakaizumi, T.; Togawa, N. An indoor positioning method using smartphone and smartwatch independent of carrying modes. In Proceedings of the 2021 IEEE International Conference on Consumer Electronics (ICCE), Las Vegas, NV, USA, 10–12 January 2021; pp. 1–6.
22. Sarhaddi, F.; Kazemi, K.; Azimi, I.; Cao, R.; Niela-Vilén, H.; Axelin, A.; Liljeberg, P.; Rahmani, A.M. A comprehensive accuracy assessment of Samsung smartwatch heart rate and heart rate variability. *PLoS ONE* **2022**, *17*, e0268361. [[CrossRef](#)] [[PubMed](#)]
23. Støve, M.P.; Hansen, E.C.K. Accuracy of the Apple Watch Series 6 and the Whoop Band 3.0 for assessing heart rate during resistance exercises. *J. Sports Sci.* **2022**, *40*, 2639–2644. [[CrossRef](#)] [[PubMed](#)]
24. Castro, J.M.C.; Martini, H.A. Potencia estadística y cálculo del tamaño del efecto en G\* Power: Complementos a las pruebas de significación estadística y su aplicación en psicología. *Salud Soc.* **2014**, *5*, 210–244. [[CrossRef](#)]
25. Naqvib, N.Z.; Kumar, A.; Chauhan, A.; Sahni, K. Step counting using smartphone-based accelerometer. *Int. J. Comput. Sci. Eng.* **2012**, *4*, 675–681.
26. Kozey-Keadle, S.; Libertine, A.; Lyden, K.; Staudenmayer, J.; Freedson, P.S. Validation of wearable monitors for assessing sedentary behavior. *Med. Sci. Sports Exerc.* **2011**, *43*, 1561–1567. [[CrossRef](#)]
27. An, H.-S.; Kim, Y.; Lee, J.-M. Accuracy of inclinometer functions of the activPAL and ActiGraph GT3X+: A focus on physical activity. *Gait Posture* **2017**, *51*, 174–180. [[CrossRef](#)]
28. Wu, Y.; Petterson, J.L.; Bray, N.W.; Kimmerly, D.S.; O’Brien, M.W. Validity of the activPAL monitor to measure stepping activity and activity intensity: A systematic review. *Gait Posture* **2022**, *97*, 165–173. [[CrossRef](#)]
29. Spaccarotella, C.; Polimeni, A.; Mancuso, C.; Pelaia, G.; Esposito, G.; Indolfi, C. Assessment of non-invasive measurements of oxygen saturation and heart rate with an Apple smartwatch: Comparison with a standard pulse oximeter. *J. Clin. Med.* **2022**, *11*, 1467. [[CrossRef](#)]

30. Lee, B.-G.; Lee, B.-L.; Chung, W.-Y. Smartwatch-based driver alertness monitoring with wearable motion and physiological sensor. In Proceedings of the 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Milan, Italy, 25–29 August 2015; pp. 6126–6129.
31. Liu, Z.; Zhao, Y.; Yin, Z. Low-power soft transistors triggering revolutionary electronics. *Innovation* **2024**, *5*, 100616. [[CrossRef](#)] [[PubMed](#)]
32. Zhang, T.; Liu, N.; Xu, J.; Liu, Z.; Zhou, Y.; Yang, Y.; Li, S.; Huang, Y.; Jiang, S. Flexible electronics for cardiovascular healthcare monitoring. *Innovation* **2023**, *4*, 100485. [[CrossRef](#)] [[PubMed](#)]
33. Hauck, W.W.; Anderson, S. A new statistical procedure for testing equivalence in two-group comparative bioavailability trials. *J. Pharmacokinet. Biopharm.* **1984**, *12*, 83–91. [[CrossRef](#)]
34. Khushhal, A.; Nichols, S.; Evans, W.; Gleadall-Siddall, D.O.; Page, R.; O'Doherty, A.F.; Carroll, S.; Ingle, L.; Abt, G. Validity and reliability of the Apple Watch for measuring heart rate during exercise. *Sports Med. Int. Open* **2017**, *1*, E206–E211. [[CrossRef](#)] [[PubMed](#)]
35. Hajj-Boutros, G.; Landry-Duval, M.-A.; Comtois, A.S.; Gouspillou, G.; Karelis, A.D. Wrist-worn devices for the measurement of heart rate and energy expenditure: A validation study for the Apple Watch 6, Polar Vantage V and Fitbit Sense. *Eur. J. Sport Sci.* **2023**, *23*, 165–177. [[CrossRef](#)]
36. Veerabhadrapa, P.; Moran, M.D.; Renninger, M.D.; Rhudy, M.B.; Dreisbach, S.B.; Gift, K.M. Tracking steps on apple watch at different walking speeds. *J. Gen. Intern. Med.* **2018**, *33*, 795–796. [[CrossRef](#)]
37. Hartung, V.; Sarshar, M.; Karle, V.; Shammas, L.; Rashid, A.; Roullier, P.; Eilers, C.; Mäurer, M.; Flachenecker, P.; Pfeifer, K. 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*, 9314. [[CrossRef](#)]
38. Hergenroeder, A.L.; Barone Gibbs, B.; Kotlarczyk, M.P.; Kowalsky, R.J.; Perera, S.; Brach, J.S. Accuracy of objective physical activity monitors in measuring steps in older adults. *Gerontol. Geriatr. Med.* **2018**, *4*, 2333721418781126. [[CrossRef](#)] [[PubMed](#)]
39. Bai, Y.; Welk, G.J.; Nam, Y.H.; Lee, J.A.; Lee, J.-M.; Kim, Y.; Meier, N.F.; Dixon, P.M. Comparison of consumer and research monitors under semistructured settings. *Med. Sci. Sports Exerc.* **2016**, *48*, 151–158. [[CrossRef](#)] [[PubMed](#)]
40. Fuller, D.; Colwell, E.; Low, J.; Orychok, K.; Tobin, M.A.; Simango, B.; Buote, R.; Van Heerden, D.; Luan, H.; Cullen, K. Reliability and validity of commercially available wearable devices for measuring steps, energy expenditure, and heart rate: Systematic review. *JMIR mHealth uHealth* **2020**, *8*, e18694. [[CrossRef](#)]
41. Bai, Y.; Hibbing, P.; Mantis, C.; Welk, G.J. Comparative evaluation of heart rate-based monitors: Apple Watch vs Fitbit Charge HR. *J. Sports Sci.* **2018**, *36*, 1734–1741. [[CrossRef](#)] [[PubMed](#)]
42. Moreno, J.J.M.; Pol, A.P.; Abad, A.S.; Blasco, B.C. Using the R-MAPE index as a resistant measure of forecast accuracy. *Psicothema* **2013**, *25*, 500–506.
43. Höchsmann, C.; Knaier, R.; Eymann, J.; Hintermann, J.; Infanger, D.; Schmidt-Trucksäss, A. Validity of activity trackers, smartphones, and phone applications to measure steps in various walking conditions. *Scand. J. Med. Sci. Sports* **2018**, *28*, 1818–1827. [[CrossRef](#)]
44. Presset, B.; Laurenczy, B.; Malatesta, D.; Barral, J. Accuracy of a smartphone pedometer application according to different speeds and mobile phone locations in a laboratory context. *J. Exerc. Sci. Fit.* **2018**, *16*, 43–48. [[CrossRef](#)]
45. Henriksen, A.; Haugen Mikalsen, M.; Woldaregay, A.Z.; Muzny, M.; Hartvigsen, G.; Hopstock, L.A.; Grimsgaard, S. Using fitness trackers and smartwatches to measure physical activity in research: Analysis of consumer wrist-worn wearables. *J. Med. Internet Res.* **2018**, *20*, e110. [[CrossRef](#)]

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.