



Original Article

Ann Rehabil Med 2023;47(4):261-271

eISSN: 2234-0653

<https://doi.org/10.5535/arm.23019>

Validation of Wearable Digital Devices for Heart Rate Measurement During Exercise Test in Patients With Coronary Artery Disease

Chul Kim, MD, PhD, Jun Hyeong Song, MD, Seung Hyoun Kim, PT, MS

Department of Rehabilitation Medicine, Inje University Sanggye Paik Hospital, Inje University College of Medicine, Seoul, Korea

Objective: To assess the accuracy of recently commercialized wearable devices in heart rate (HR) measurement during cardiopulmonary exercise test (CPX) under gradual increase in exercise intensity, while wearable devices with HR monitors are reported to be less accurate in different exercise intensities.

Methods: CPX was performed for patients with coronary artery disease (CAD). Twelve lead electrocardiograph (ECG) was the gold standard and Apple watch 7 (AW7), Galaxy watch 4 (GW4) and Bio Patch Mobicare 200 (MC200) were applied for comparison. Paired absolute difference (PAD), mean absolute percentage error (MAPE) and intraclass correlation coefficient (ICC) were evaluated for each device.

Results: Forty-four participants with CAD were included. All the devices showed MAPE under 2% and ICC above 0.9 in rest, exercise and recovery phases (MC200=0.999, GW4=0.997, AW7=0.998). When comparing exercise and recovery phase, PAD of MC200 and AW7 in recovery phase were significantly bigger than PAD of exercise phase ($p<0.05$). Although not significant, PAD of GW4 tended to be bigger in recovery phase, too. Also, when stratified by HR 20, ICC of all the devices were highest under HR of 100, and ICC decreased as HR increased. However, except for ICC of GW4 at HR above 160 (=0.867), all ICCs exceeded 0.9 indicating excellent accuracy.

Conclusion: The HR measurement of the devices validated in this study shows a high concordance with the ECG device, so CAD patients may benefit from the devices during high-intensity exercise under conditions where HR is measured reliably.

Keywords: Wearable electronic devices, Cardiac rehabilitation, Heart rate, Exercise test

Received: February 9, 2023

Revised: June 2, 2023

Accepted: June 22, 2023

Correspondence:

Jun Hyeong Song
Department of Rehabilitation Medicine,
Inje University Sanggye Paik Hospital,
Inje University College of Medicine,
1342 Dongil-ro, Nowon-gu, Seoul,
Korea.

Tel: +82-2-950-1145

Fax: +82-2-950-1144

E-mail: sjh9233@hanmail.net

INTRODUCTION

Cardiac rehabilitation (CR) is an essential component in the continuum of care for patients with cardiovascular disease. The safety and efficacy of CR are well documented, but CR participation rates remain low and suboptimal worldwide. Despite

being eligible candidates, over 80% of patients in the United States and 50% in the United Kingdom do not participate in CR [1,2]. Factors such as reluctance to participate in group rehabilitation, inconvenient exercise schedule, career responsibilities, transportation, and related costs were prominent barriers to CR participation [3,4]. In addition, during the coronavirus disease

2019 pandemic, safe distancing measures have led to cessation of center based cardiac rehabilitation (CBCR) programs [4].

Home based cardiac rehabilitation (HBCR) has been introduced as an alternative strategy to expand access and participation over conventional CBCR [3,5]. In concept, HBCR could help overcome barriers for CBCR such as geographic, occupational and other access related barriers.

Heart rate (HR) is an important parameter for determining appropriate exercise intensity and establishing a safe zone when performing self-monitored aerobic exercise during HBCR [6,7]. Commonly, patients are instructed to take pulse rate (PR) at their radial or carotid artery and measure their arterial beats for 10 seconds (and multiply it with 6 to get HR per minute), but this method is difficult and inconvenient during aerobic exercise and also is less accurate [6].

With the recent development of technology, commercial electronic devices with HR monitor in forms of chest straps, smart watches and smart bands have been introduced and are being used worldwide. The HR measurement derived from these tools not only reflect patient's condition but also provides guidance for appropriate exercise [7].

Researches on the accuracy and validity of HR measuring wearable devices have been actively reported [6-13]. However, the reports on HR accuracy so far are somewhat inconsistent. Depending on type of the devices, such as chest strap, patch, smart band or smart watch, the HR accuracy was variable. Chest strap type tended to be more accurate than other wrist worn devices [8]. Smart watches were more accurate in terms of HR (actually PR) than relatively inexpensive smart bands [8-10]. Also, the accuracy of these devices fluctuated in accordance to the intensity or type of exercise [6,10]. HR measured by smart band seemed to be inaccurate with high intensity aerobic exercises [6].

To the best of our knowledge, although previous reports were made recently, these studies were conducted with relatively old and inexpensive series of available wearable devices [7,8-13]. In this study, we used the latest products from the two electronics companies with the largest market share in the world, which are Apple watch 7 (AW7; Apple Inc.) and Galaxy watch 4 (GW4; Samsung). We also included a recently introduced patch type HR monitoring device, Bio Patch Mobicare 200 (MC200; Seers Technology), to compare with the smart watches.

We hypothesized that the newer wearable devices compared to the older models would be more accurate with HR moni-

toring and show higher correlation with the 12 lead electrocardiograph (ECG), which is the gold standard. This study aimed to confirm the validity and accuracy of HR measurement in recently released commercial smart watches and chest patch device, during cardiopulmonary exercise test (CPX).

METHODS

The study protocol was approved by the Institutional Review Board of Inje University (no. 2021-10-007). All the participants provided written informed consent. The researcher explained the purpose, methods, advantages, and potential risk of the study to every participants. Patient privacy and data confidentiality were maintained throughout the study period. The smart watches used in the study were purchased at researcher's own expense. The authors declare no conflict of interest.

Sample size calculation

Since free software calculating sample sizes for reliability coefficients are relatively scarce, the calculation was performed on a web-based calculator in accordance to the article by Arifin [14]. The test-retest reliability of a measurement tool by intraclass correlation coefficient (ICC) is expected=0.9. The measurement was taken at two occasions (12 lead ECG vs. each wearable device). The lowest acceptable ICC is 0.75 [14,15]. A significance level of $\alpha=0.05$ and a power of 90% were set. In conclusion, 44 subjects were required for this study.

Participants

This was a comparative observational study recruiting outpatients of Cardiac Rehabilitation Clinic at Inje University Sanggye Paik Hospital, Inje University College of Medicine, from June 2022 to September 2022. We included 44 patients aged 20 to under 75, diagnosed with coronary artery disease, including acute coronary syndrome and stable angina, and subsequently referred for CR. These patients were scheduled for outpatient CPX as a regular follow-up exam. For this study, the participants had to put on extra wearable devices while going through CPX. Patients with contraindications to CPX were excluded according to the guideline by the American Heart Association (AHA) [16]. Also, patients referred for CR with other cardiac etiologies, such as valvular or aortic disease, were excluded from the study.

CPX

CPX was performed on a treadmill and stress test system (T-2100 & CASE; GE Healthcare) according to the modified Bruce protocol. In this study, respiratory gas analyzer (Quark CPET; COSMED), automatic blood pressure and pulse monitor (TANGO M2; SunTech) were used to measure physiologic variables. The HR recorded from the CPX machine was set as the gold standard.

All three wearable devices used in this study were recently released products within a year. MC200 is a device approved for long term ECG recording by the Korean Ministry of Health and Welfare in February 2022. MC200 was attached along the axis of lead II, from lateral aspect of left upper sternum to apex of heart. GW4 and AW7, which are wrist strap type, were worn around left and right wrists respectively. We looked up in the product instructions and fitted the watches exactly where it should be (2–3 cm above styloid process of radius). It was fixed to a hole of a strap that could fit the device as tight as we could so it doesn't move sideways, but also not as much as it would choke patient's wrist. To reduce the bias caused while fastening the strap, one researcher helped put on the device in person so that the straps were adjusted to participants' wrist with less inconsistency.

In the rest phase, which lasted for 6 minutes, patients were sat on a chair while blood pressure and ECG were being monitored. HR from 12 lead ECG and wearable devices were recorded every two minutes during the rest phase. After 6 minutes, exercise phase started according to the modified Bruce protocol with gradual increase in intensity and HR was recorded every minute during all exercise stages until termination of the test. The participants held handrails in front of them to prevent seri-

ous harm during the CPX test since the most participants were middle-aged. Termination of the test was decided according to the "Indications for termination of exercise testing" of AHA guidelines [16]. Lastly, in the recovery phase, patients kept walking slowly for additional 5 minutes, and HR was measured every minute (Fig. 1).

Statistical analysis

All data were analyzed using IBM SPSS ver. 25 (IBM Corp.) and the values were presented as means \pm standard deviation (SD) or numbers and percentages. The HR data from gold standard 12 lead ECG and each wearable devices were analyzed with following parameters.

First, paired relative difference and paired absolute difference (PAD) were calculated to assess accuracy, by subtracting the HR recorded by wearable devices (HR_{DEV}) from the HR recorded by the 12 lead ECG from CPX machine (HR_{ECG}). Also, mean absolute percentage error (MAPE) values were calculated as the average absolute value of the errors of each wearable device relative to 12 lead ECG, the gold standard measure, expressed in percentage. Fokkema et al. [17] suggest a MAPE threshold of 5%, whereas Nelson and Allen [18] used a MAPE threshold of 10% to classify a wearable device as valid.

The degree of agreement between the values of two devices was examined using the ICC with 95% confidence interval (CI).

The thresholds suggested by Fokkema et al. [17] were used; excellent, 0.90 or above; good, 0.75–0.90; moderate, 0.60–0.75; and low, 0.60 or below.

In addition, Bland–Altman plot with 95% limits of agreement (LoA) was used to measure each outcome relative to the gold standard. The differences of each sample are scattered on the

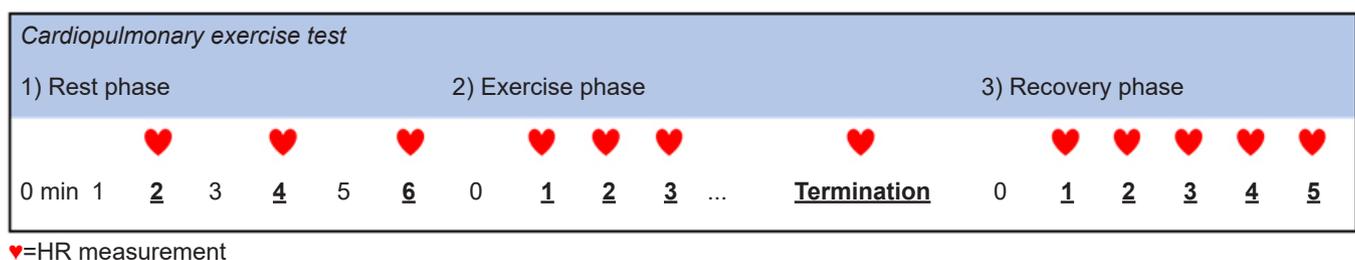


Fig. 1. Heart rate (HR) measurement during cardiopulmonary exercise test. HR was recorded at 2, 4, and 6 minutes in rest phase. HR was recorded every minute beginning from 1 minute during exercise phase until termination of the test. HR at termination, which was the beginning of recovery phase, was also recorded and once every minute subsequently in recovery phase. The time points for HR measurement are in bold and underlined. Also, these time points are marked with heart shape (♥), respectively.

vertical axis and the average of two measurements are scattered on the horizontal axis. The plot contains three horizontal lines indicating the mean difference and upper and lower LoA which could be calculated as the mean difference ± 1.96 SD of differences [19].

RESULTS

Total 44 of 36 male and 8 female patients were included for the study. In CPX results, average attained stage was 6.0, lasting average 16 minutes 26 seconds. Average rate of perceived exertion was 16.4 with respiratory exchange ratio of 1.2. The average of peak oxygen uptake and peak metabolic equivalents were 25.8 and 7.4, respectively. The baseline characteristics and summary of CPX results are presented in Table 1. Incidence of arrhythmias

during CPX test and total number of HR records for each patient are described in Table 2. There were participants with frequent premature beats but there were no sustained arrhythmias that could fluctuate HR during CPX.

Average paired relative difference and relative percent difference of all the devices were within absolute value of 1 at all phases (Table 3). Average PAD were also around 1 at all phases, with AW7 showing maximal 1.29 at recovery phase. Absolute percent difference of all three devices at all phases did not exceed value of 2%. ICC of all the devices exceeded 0.99 at all the phases, except for AW7 resulting in 0.984 at rest phase but still showing excellent correlation.

On the other hand, influence of HR itself on accuracy of the devices were evaluated. HR recordings from exercise phase and recovery phase were grouped into two subgroups of “HR below 100” and tachycardia, which were “HR at or above 100” (Table 4). PAD, MAPE, and ICC were obtained. Although not always statistically significant, there were some tendency of HR accuracy variability according to HR and CPX phases. PAD values tended to be larger when HR was above 100, and MAPE values tended to be larger when HR was below 100. However, PAD and MAPE did not exceed value of 2 and ICC values were above 0.9 in every circumstances regardless of HR or phases.

Fig. 2 presents Bland–Altman analysis, where solid horizontal lines indicate the average HR differences in each device. Two dashed lines indicate upper and lower 95% confidence LoA for each device respectively. MC200 showed average difference of 0.24, with upper and lower LoA of 2.80 and -2.32 , respectively. Average difference of GW4 was -0.18 with 95% LoA of 3.58 and -3.95 . Lastly, AW7 had mean value of 0.11 and 95% LoA of 3.33 and -3.22 , respectively.

In Table 5, ICC values of the devices in accordance to HR divided into interval of 20, ranging from below 100, 100–119, 120–139, 140–159 to at or above 160 were analyzed. The ICC value was highest at HR below 100 in all three devices. The accuracy decreased when HR was over 100 in all groups. GW4 showed the lowest ICC of 0.867 at HR at or above 160 and AW7 showed the lowest value of 0.925 at HR between 140–159. MC200, on the other hand, showed minimum ICC of 0.980 at HR between 100–119 and 120–139.

The overall ICC of three devices, analyzed with the overall HR measurements, were 0.999, 0.998, and 0.997 for MC200, AW7 and GW4 respectively with MC200 being the most accurate (Fig. 3).

Table 1. Baseline characteristics of participants

Characteristic	Value (n=44)
Age (yr)	60.9 \pm 7.9
Sex, male:female	36:8
Height (cm)	165.5 \pm 8.0
Weight (kg)	70.0 \pm 10.9
Body mass index (kg/m ²)	25.5 \pm 2.9
Diagnosis	
STEMI	12 (27.3)
NSTEMI	9 (20.5)
Unstable angina	13 (29.5)
Stable angina	10 (22.7)
Electrocardiograph at rest	
Normal sinus rhythm	20 (45.5)
Sinus bradycardia	16 (36.4)
1°AV block	3 (6.8)
Premature atrial beat	7 (15.9)
Premature ventricular beat	4 (9.1)
Intervention	
Percutaneous coronary intervention	42 (95.5)
Coronary artery bypass graft	2 (4.5)
Cardiopulmonary exercise test results	
Attained stage	6.0 \pm 0.59
Duration of test (min:s)	16:26 \pm 1:53
VO _{2AT} (mL/kg/min)	18.7 \pm 4.6
VO _{2peak} (mL/kg/min)	25.8 \pm 5.1
METS _{peak}	7.4 \pm 1.5
RPE _{max}	16.4 \pm 1.4
Respiratory exchange ratio	1.2 \pm 0.1

Values are presented as mean \pm standard deviation, number only, or number (%). STEMI, ST segment elevation myocardial infarction; NSTEMI, non STEMI; 1°AV block, 1st degree atrioventricular block; VO_{2AT}, oxygen uptake at anaerobic threshold; VO_{2peak}, peak oxygen uptake; METS_{peak}, peak metabolic equivalents; RPE_{max}, maximal rate of perceived exertion.

Table 2. Numbers of arrhythmia and total heart rate measurements in each patient

Patient no.	Atrial premature complex	Ventricular premature complex	Paroxysmal supraventricular tachycardia	Other arrhythmias	No. of heart rate records
1	16	21	1	0	21
2	0	0	0	0	24
3	196	1	1	0	23
4	0	2	4	0	24
5	0	0	3	0	20
6	0	16	8	0	23
7	0	21	0	0	25
8	0	0	1	0	27
9	0	0	0	0	25
10	6	1	4	0	28
11	54	8	6	0	25
12	0	0	9	0	24
13	5	0	1	0	23
14	2	1	1	0	26
15	0	15	17	0	23
16	0	0	5	0	26
17	18	9	5	0	21
18	0	3	4	0	23
19	18	0	1	0	26
20	0	6	1	0	26
21	0	1	12	0	24
22	0	2	2	0	24
23	4	1	0	0	24
24	0	2	0	0	24
25	0	0	0	0	22
26	0	2	0	0	24
27	0	109	4	0	23
28	6	0	4	0	21
29	0	4	6	0	24
30	6	0	0	0	25
31	0	8	4	0	25
32	0	0	0	0	24
33	0	0	0	0	24
34	0	0	1	0	21
35	0	0	0	0	24
36	0	1	9	0	23
37	0	10	2	0	28
38	0	0	0	0	24
39	0	0	1	0	24
40	0	0	1	0	20
41	0	0	0	0	25
42	0	2	4	0	23
43	4	4	4	0	25
44	0	0	5	0	26

Values are presented as number for each variable.

DISCUSSION

In this study, we aimed to verify the accuracy of HR monitoring function of wearable devices by performing the CPX according

to the modified Bruce protocol. The devices used in the study, all commercialized recently, were AW7 and GW4, both of which were released by the top 2 largest electronics companies in the global smart watch market. For comparison, MC200, a chest

Table 3. Average HR differences between each wearable device and 12 lead ECG

Device	HR differences from 12 lead ECG				ICC	
	Paired relative difference ^{a)}	Paired absolute difference ^{b)}	Relative percent difference ^{c)}	Absolute percent difference ^{d)}		
REST	MC200	0.17±1.47	0.85±1.21	0.27± 2.33	1.33±1.93	0.990
	GW4	-0.30±1.46	0.86±1.21	-0.49±2.37	1.37±1.99	0.990
	AW7	-0.34±1.89	1.01±1.63	-0.49±2.81	1.55±2.39	0.984
EXE	MC200	0.18±1.23	0.67±1.05	0.16±1.38	0.70±1.20	0.999
	GW4	0.03±1.95	0.99±1.57	-0.08±1.88	0.99±1.60	0.997
	AW7	0.33±1.41	0.78±1.22	0.25±1.50	0.78±1.31	0.998
REC	MC200	0.57±1.37	1.07±1.03	0.51±1.29	1.04±0.89	0.998
	GW4	-0.98±2.14	1.26±1.99	-1.00±2.33	1.25±2.49	0.994
	AW7	-0.46±2.24	1.29±1.88	-0.44±2.39	1.31±2.31	0.994

Values are presented as mean±standard deviation.

HR, heart rate; ECG, electrocardiograph; ICC, intraclass correlation coefficient; REST, rest phase; MC200, Mobicare 200; GW4, Galaxy watch 4; AW7, Apple watch 7; EXE, exercise phase; REC, recovery phase; DEV, test device.

^{a)}HR_{ECG}-HR_{DEV}, ^{b)}the absolute value of (HR_{ECG}-HR_{DEV}), ^{c)}(HR_{ECG}-HR_{DEV})/HR_{ECG}×100, and ^{d)}the absolute value of (HR_{ECG}-HR_{DEV})/HR_{ECG}×100.

Table 4. PAD, MAPE, and ICC of each device during exercise phase and recovery phase

	Mobicare 200			Galaxy watch 4			Apple watch 7		
	PAD	MAPE	ICC	PAD	MAPE	ICC	PAD	MAPE	ICC
<exe100	0.66±1.27 ^{a)}	0.81±1.52 ^{b)}	0.989	0.89±1.57 ^{b)}	1.07±1.85 ^{b)}	0.983	0.69±1.35 ^{a,b)}	0.84±1.63	0.987
≥exe100	0.68±0.73 ^{a)}	0.57±0.63 ^{b)}	0.999	1.10±1.56 ^{b)}	0.89±1.24 ^{b)}	0.994	0.88±1.03 ^{a,b)}	0.70±0.78	0.998
<rec100	0.86±0.74 ^{a,b)}	1.01±0.84	0.991	1.24±2.52 ^{b)}	1.45±2.96	0.952	1.41±2.38 ^{a)}	1.64±2.68	0.947
≥rec100	1.27±1.20 ^{a,b)}	1.07±0.99	0.993	1.27±1.30 ^{b)}	1.06±1.05	0.993	1.16±1.25 ^{a)}	0.98±1.08	0.993

Values are presented as mean±standard deviation.

PAD, paired absolute difference; MAPE, mean absolute percentage error; ICC, intraclass correlation coefficient; <exe100, exercise phase with heart rate under 100 beats/min; ≥exe100, exercise phase with heart rate 100 or over 100 beats/min; <rec100, recovery phase with heart rate under 100 beats/min; ≥rec100, recovery phase with heart rate 100 or over 100 beats/min.

^{a)}Significant difference (p<0.05) of parameters within each device in accordance to phases with same heart rate interval.

^{b)}Significant difference (p<0.05) of parameters within each device, in accordance to heart rate interval within same phase.

patch type device for a long-term ECG recording approved by the Korean Ministry of Health and Welfare on February 2022, was adopted.

A few earlier studies have assessed the HR accuracy of various smart devices. Boudreaux et al. [9] reported the average MAPE of Apple watch 2 as 4.14 and the average ICC as 0.9. But, the exercise intensity was reported to affect the HR accuracy with the MAPE as 7.16 at maximum at high intensity exercise and the ICC as 0.8 at minimum [9]. In another study, the values of the MAPE were all affected by the exercise intensity. The MAPE of Apple watch (series unclarified) varied from 1.14% to 6.70%—the highest MAPE value in the moderate intensity exercise; the MAPE of Fitbit from 2.38% to 16.99%; and the MAPE of Garmin Forerunner 225 from 7.87% to 24.38% [10]. Likewise, the overall HR accuracy depended on the type of devices, series, and models, and the type of exercise and intensity.

Many other studies have showed somewhat high MAPE and low correlation coefficient value with less accurate HR values of

the devices during exercise [7,8-13]. Though chest-patch devices are known to be more accurate than other types of devices, it was reported that even the chest-patch devices (Mobicare 100) showed a correlation coefficient of 0.69 during high intensity exercise with above HR 160 [6,8].

In contrast, the devices used in this study showed the MAPE value of less than 2% and ICC value of over 0.9 in every circumstance, which shows a high correlation with the gold standard 12 lead ECG—except the HR measurement from GW4 in high intensity exercise phase with HR over 160, showing ICC of 0.867. Still, the minimum value of ICC still showed a good correlation that is similar or higher than that of previous reports conducted with other older devices.

In the exercise phase of modified Bruce protocol, the exercise intensity increases every three minutes. Also, when the CPX terminates, the exercise intensity and the HR begins to decrease radically in recovery phase. The initial increase of HR in the exercise phase are caused by the parasympathetic withdrawal,

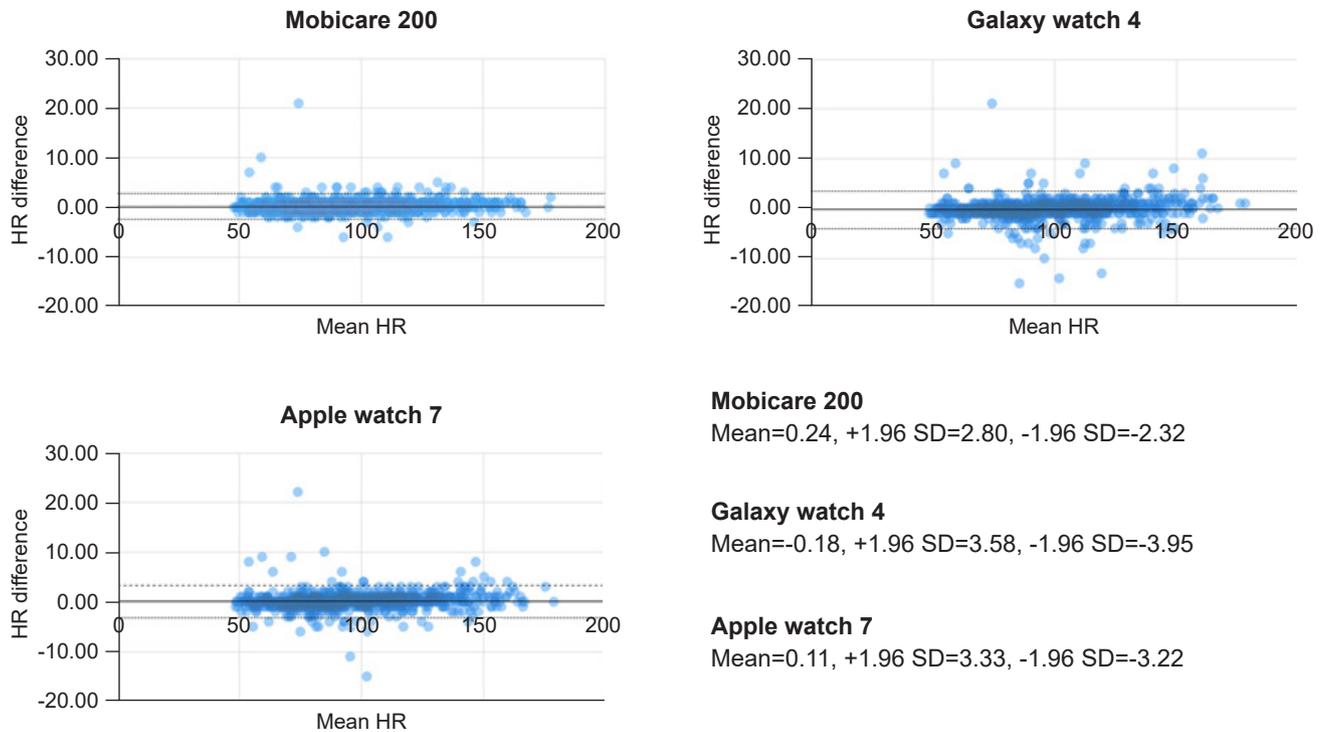


Fig. 2. Bland–Altman plots show agreement between 12 lead electrocardiograph, the gold standard and each wearable device. Solid horizontal lines indicate average heart rate (HR) differences in each device. Dashed lines indicate 95% confidence limits of agreement for each device. SD, standard deviation.

Table 5. Intraclass correlation coefficient in accordance to heart rate interval

Devices	Heart rate				
	<100	100–119	120–139	140–159	≥160
Mobicare 200	0.994	0.980	0.980	0.984	0.990
Galaxy watch 4	0.989	0.955	0.970	0.912	0.867
Apple watch 7	0.989	0.979	0.977	0.925	0.970

Values are presented as intraclass correlation coefficient of each device in accordance to heart rate interval.

while sympathetic activation is responsible for HRs greater than 100 beats/minute. Also, the rapid drop in HR for the first minute after the cessation of exercise is mostly determined by parasympathetic reactivation [20]. Due to the impaired autonomic regulation, CAD patients show delayed HR recovery which may take up to five minutes [21,22]. We postulated that such changes in HR according to the exercise intensity would affect HR accuracy of the devices.

Accordingly, the HR measurements in the exercise and recovery phase were divided into four subgroups based on “HR below 100” and “HR above 100.” Since the MAPE is an indicator expressed as percentage, the MAPE values tended to be small

when the denominator was greater than 100, and be big when the denominator was less than 100 (Table 4). Thus, the PAD was analyzed along with the MAPE for precise comparison.

The PAD values of MC200 and AW7 were significantly larger in recovery phase than in exercise phase. The PAD values of GW4 and MAPE values of all the devices did not show any significant difference between phases, but the values tended to be larger in recovery phase than those in exercise phase as postulated. However, since the HR difference between phases were about 1 bpm on average, the difference would not indicate any clinical significance in the real world. Also, ICC values exceeded 0.9 in all conditions, which showed a high reliability.

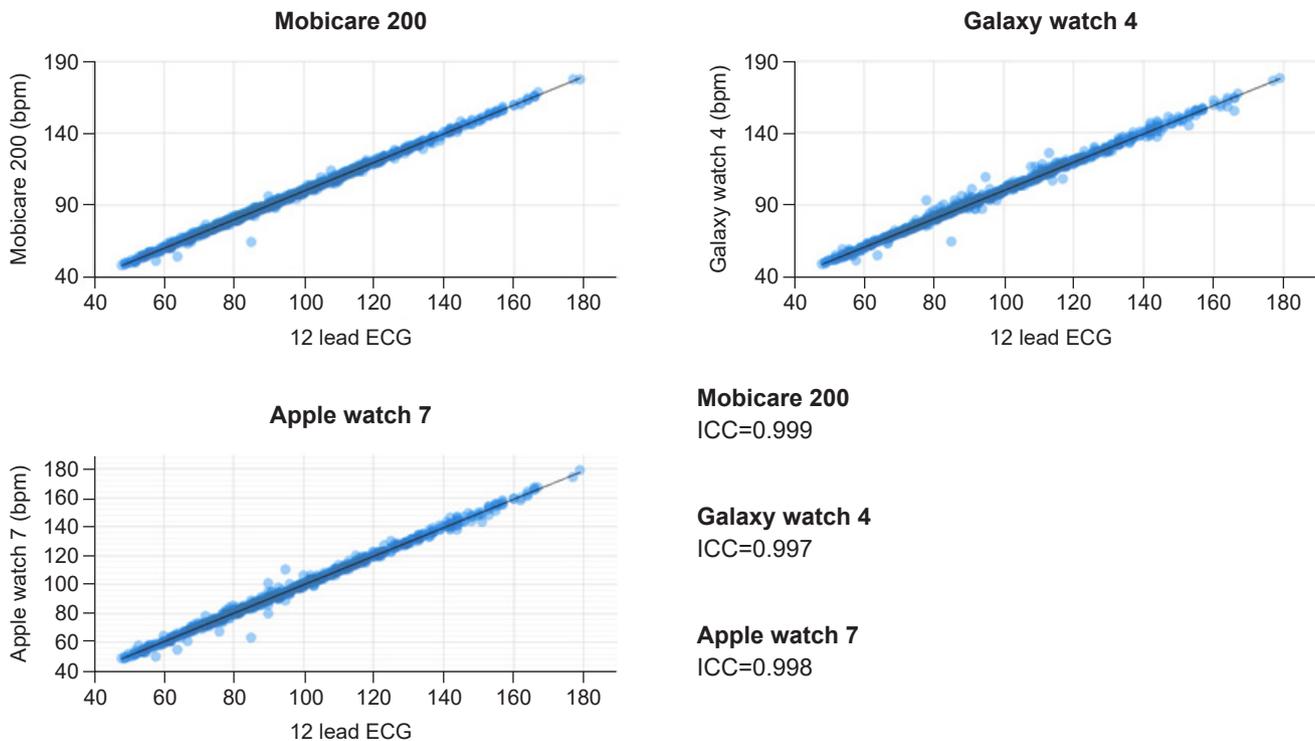


Fig. 3. Scatter plots show correlation between 12 lead electrocardiograph (ECG) vs. each wearable device. Diagonal lines are lines of identity. Each dot is a separate measurement of heart rate. ICC, intraclass correlation coefficient.

In addition, the comparison between the HR subgroups within the same phases showed some significant differences, but the results were inconsistent.

To examine HR accuracy in accordance with the exercise intensity, we stratified the HR measurements by the interval of 20. As in Table 5, the ICC of all three devices were highest when HR was below 100. Though the ICC tended to decrease as HR increased, the relationship was not linear, in that it showed the lowest ICC at the moderate intensity of HR between 140 and 159 in MC200 and AW7. It is consistent with the previous studies that suggest a higher accuracy with vigorous exercise than with slow walking [23-25]. However, as previously mentioned, the lowest ICC values of all three devices were still higher than those of other studies, indicating an overall high level of accuracy regardless of the exercise intensity.

As chest-strap-based HR monitors were first introduced in the 1980s, several studies have already confirmed the accuracy of these monitors as a prototype of HR measuring devices [26-28]. However, these devices were only used by elite athletes due to its inconvenience. Nowadays, new, convenient wrist-worn HR monitors have garnered public attention. Unlike the chest-

strap type directly sensing one's cardiac electric activity, these devices use photoplethysmogram (PPG) sensors that detect variations in blood volume in peripheral vessels produced by each cardiac contraction by projecting photons into body tissues and analyzing the reflected light. Such a PPG technique is one of the most widely adopted HR measuring technologies in smart devices produced these days [23].

Sampling frequency, defined by "the number of samples per second (or per other unit) taken from a continuous signal to make a discrete or digital signal," is an important concept to better understand mechanism and pitfalls of PPG sensors. A higher sampling frequency indicates a higher accuracy of analysis because more data samples are collected in the same time interval [29]. Gold standard ECG used in this study collects data at the frequency of 512 Hz, and MC200 collects at the frequency of 256 Hz. The exact sampling frequency of AW7 and SW4 are not made public. While the sampling frequency of the previous Samsung Watch is estimated to be between 20-25 Hz [30,31], Apple describe that their HR monitor blink hundreds of times per second to measure HR, without mentioning the precise sampling frequency [32]. Likewise, the discrepancy in

each sampling frequency may be an important reason for different accuracy between devices.

On the other hand, PPG sensors have some drawbacks. According to previous research, PPG sensors show low accuracy of HR measurements with darker skin likely because darker skin with more melanin absorbs more green light than lighter skin [29]. Moreover, the accuracy of PPG sensors are susceptible to motion artifact especially when the motion was cyclic or repetitive (e.g., walking and jogging). An accelerometer, which senses changes in velocity over time, is one of the methods used to detect motion artifact and filter this output as noise reference [29,33].

In this study, we established that the new devices demonstrate the high accuracy of HR measurements even in the high-intensity aerobic exercise and during a rapid change in HR. It could be attributable to advanced technologies such as algorithms for signal processing in the models released recently.

Still, there are several limitations in this study. First, there is an imbalance of the gender distribution of the participants with more male ones, probably due to male-dominant prevalence of CAD [34,35]. However, previous studies concluded that the effect of gender on HR accuracy is none or unlikely clinically relevant [35,36]. Second, participants with severe or sustained arrhythmia (such as atrial fibrillation or ventricular tachycardia) were not recruited for the study. The HR accuracy might decline when participants have such arrhythmias while measuring the HR. Third, participants held handrails in front of them for safety measures during the CPX. Since the hand position was fixed during the CPX, compared to regular running or jogging, the HR accuracy might have been overestimated with less motion artifact. Also, other aerobic exercises such as tennis, swimming, and skiing, or other forms of exercise such as resistance training may yield a different result.

Fourth, the results from this study may not be able to be generalized because only one device from each brand was used due to the cost. Some variations might be observed between devices when customers actually purchase these products. In addition, the effect of whether the wrist to wear a watch is left or right was not considered. One previous research concluded that neither left nor right hand wearing the device affected the accuracy of HR measurements [6]. Yet, it could have been more accurate if the hand wearing the device was randomized as part of an attempt to determine the factors affecting the accuracy of HR measurement. Also, even though one researcher helped the participants put on the devices as tight as possible, since the strap had holes in a straight line with regular interval for the users to

fasten the strap, the devices could not be worn with fine adjustment. The tightness of the strap was decided without objective measurement of tension or pressure, but rather subjectively. Hence, the tightness of the devices might not have been consistent among participants. Lastly, although the HR values from all the devices were displayed on the monitor in real-time, we were not able to check all the values simultaneously. However, the time lag between HR records were less than a few seconds. As the electronics technology advances with newer products, additional studies that supplements such limitations will be necessary in the future.

In the present study, we evaluated the HR accuracy of wearable devices recently released during the CPX test performed by participants holding handrails. We concluded that new devices demonstrate the superior accuracy of HR measurement, compared to previous studies conducted with old devices even during the high intensity exercise. It is likely that cardiac patients participating in the high-intensity exercise under condition where HR is measured reliably, may benefit from utilizing these wearable devices. Also, as many people worldwide already own these watches, it will be helpful for both cardiac patients and healthy people when they do self-exercise.

CONFLICTS OF INTEREST

No potential conflict of interest relevant to this article was reported.

FUNDING INFORMATION

None.

AUTHOR CONTRIBUTION

Conceptualization: Kim C. Methodology: Kim C, Song JH, Kim SH. Formal analysis: Song JH, Kim SH. Project administration: Kim C. Visualization: Kim C, Song JH, Kim SH. Writing – original draft: Song JH. Writing – review and editing: Kim C, Song JH, Kim SH. Approval of final manuscript: all authors.

ORCID

Chul Kim, <https://orcid.org/0000-0001-8223-2945>

Jun Hyeong Song, <https://orcid.org/0000-0001-9003-0032>

Seung Hyoun Kim, <https://orcid.org/0000-0003-3051-9287>

REFERENCES

1. Thomas RJ, Beatty AL, Beckie TM, Brewer LC, Brown TM, Forman DE, et al. Home-based cardiac rehabilitation: a scientific statement from the American Association of Cardiovascular and Pulmonary Rehabilitation, the American Heart Association, and the American College of Cardiology. *Circulation* 2019;140:e69-89.
2. Kim C, Sung J, Han JY, Jee S, Lee JW, Lee JH, et al. Current status of cardiac rehabilitation in the Regional Cardiocerebrovascular Centers in Korea. *J Clin Med* 2021;10:5079.
3. Buckingham SA, Taylor RS, Jolly K, Zawada A, Dean SG, Cowie A, et al. Home-based versus centre-based cardiac rehabilitation: abridged Cochrane systematic review and meta-analysis. *Open Heart* 2016;3:e000463.
4. Antoniou V, Davos CH, Kapreli E, Batalik L, Panagiotakos DB, Pepera G. Effectiveness of home-based cardiac rehabilitation, using wearable sensors, as a multicomponent, cutting-edge intervention: a systematic review and meta-analysis. *J Clin Med* 2022;11:3772.
5. Anderson L, Sharp GA, Norton RJ, Dalal H, Dean SG, Jolly K, et al. Home-based versus centre-based cardiac rehabilitation. *Cochrane Database Syst Rev* 2017;6:CD007130.
6. Kim C, Kim SH, Suh MR. Accuracy and validity of commercial smart bands for heart rate measurements during cardiopulmonary exercise test. *Ann Rehabil Med* 2022;46:209-18. Erratum in: *Ann Rehabil Med* 2023;47:78.
7. Bidargaddi NP, Sarela A. Activity and heart rate-based measures for outpatient cardiac rehabilitation. *Methods Inf Med* 2008;47:208-16.
8. Gillinov S, Etiwy M, Wang R, Blackburn G, Phelan D, Gillinov AM, et al. Variable accuracy of wearable heart rate monitors during aerobic exercise. *Med Sci Sports Exerc* 2017;49:1697-703.
9. Boudreaux BD, Hebert EP, Hollander DB, Williams BM, Cormier CL, Naquin MR, et al. Validity of wearable activity monitors during cycling and resistance exercise. *Med Sci Sports Exerc* 2018;50:624-33.
10. Dooley EE, Golaszewski NM, Bartholomew JB. Estimating accuracy at exercise intensities: a comparative study of self-monitoring heart rate and physical activity wearable devices. *JMIR Mhealth Uhealth* 2017;5:e34.
11. Etiwy M, Akhrass Z, Gillinov L, Alashi A, Wang R, Blackburn G, et al. Accuracy of wearable heart rate monitors in cardiac rehabilitation. *Cardiovasc Diagn Ther* 2019;9:262-71. Erratum in: *Cardiovasc Diagn Ther* 2020;10:644-5.
12. Martinato M, Lorenzoni G, Zanchi T, Bergamin A, Buratin A, Azzolina D, et al. Usability and accuracy of a smartwatch for the assessment of physical activity in the elderly population: observational study. *JMIR Mhealth Uhealth* 2021;9:e20966.
13. Prieto-Avalos G, Cruz-Ramos NA, Alor-Hernández G, Sánchez-Cervantes JL, Rodríguez-Mazahua L, Guarneros-Nolasco LR. Wearable devices for physical monitoring of heart: a review. *Biosensors (Basel)* 2022;12:292.
14. Arifin WN. A web-based sample size calculator for reliability studies. *Educ Med J* 2018;10:67-76.
15. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropr Med* 2016;15:155-63. Erratum in: *J Chiropr Med* 2017;16:346.
16. Fletcher GF, Ades PA, Kligfield P, Arena R, Balady GJ, Bittner VA, et al; American Heart Association Exercise, Cardiac Rehabilitation; Prevention Committee of the Council on Clinical Cardiology, Council on Nutrition, Physical Activity and Metabolism, Council on Cardiovascular; Stroke Nursing, and Council on Epidemiology and Prevention. Exercise standards for testing and training: a scientific statement from the American Heart Association. *Circulation* 2013;128:873-934.
17. Fokkema T, Kooiman TJ, Krijnen WP, VAN DER Schans CP, DE Groot M. Reliability and validity of ten consumer activity trackers depend on walking speed. *Med Sci Sports Exerc* 2017;49:793-800.
18. Nelson BW, Allen NB. Accuracy of consumer wearable heart rate measurement during an ecologically valid 24-hour period: intraindividual validation study. *JMIR Mhealth Uhealth* 2019;7:e10828.
19. Kwon S, Lee SR, Choi EK, Ahn HJ, Song HS, Lee YS, et al. Validation of adhesive single-lead ECG device compared with Holter monitoring among non-atrial fibrillation patients. *Sensors (Basel)* 2021;21:3122.
20. Ghaffari S, Kazemi B, Aliakbarzadeh P. Abnormal heart rate recovery after exercise predicts coronary artery disease severity. *Cardiol J* 2011;18:47-54.
21. Johnson NP, Goldberger JJ. Prognostic value of late heart rate recovery after treadmill exercise. *Am J Cardiol* 2012;110:45-9. Erratum in: *Am J Cardiol* 2012;110:763.
22. Gayda M, Bourassa MG, Tardif JC, Fortier A, Juneau M, Nigam A. Heart rate recovery after exercise and long-term prognosis in patients with coronary artery disease. *Can J Cardiol* 2012;28:201-7.
23. Navalta JW, Montes J, Bodell NG, Salatto RW, Manning JW, DeBello M. Concurrent heart rate validity of wearable technology devices during trail running. *PLoS One* 2020;15:e0238569.
24. Alzahrani A, Hu S, Azorin-Peris V, Barrett L, Eslinger D, Hayes M, et al. A multi-channel opto-electronic sensor to accurately monitor heart rate against motion artefact during exercise. *Sensors (Basel)* 2015;15:25681-702.
25. El-Amrawy F, Nounou MI. Are currently available wearable devices

- for activity tracking and heart rate monitoring accurate, precise, and medically beneficial? *Healthc Inform Res* 2015;21:315-20.
26. Laukkanen RM, Virtanen PK. Heart rate monitors: state of the art. *J Sports Sci* 1998;16 Suppl:S3-7.
 27. Léger L, Thivierge M. Heart rate monitors: validity, stability, and functionality. *Phys Sportsmed* 1988;16:143-51.
 28. Patel MS, Asch DA, Volpp KG. Wearable devices as facilitators, not drivers, of health behavior change. *JAMA* 2015;313:459-60.
 29. Bent B, Goldstein BA, Kibbe WA, Dunn JP. Investigating sources of inaccuracy in wearable optical heart rate sensors. *NPJ Digit Med* 2020;3:18.
 30. Sarhaddi F, Kazemi K, Azimi I, Cao R, Niela-Vilén H, Axelin A, et al. A comprehensive accuracy assessment of Samsung smartwatch heart rate and heart rate variability. *PLoS One* 2022;17:e0268361.
 31. Nissen M, Slim S, Jäger K, Flaucher M, Huebner H, Danzberger N, et al. Heart rate measurement accuracy of Fitbit Charge 4 and Samsung Galaxy Watch Active2: device evaluation study. *JMIR Form Res* 2022;6:e33635.
 32. Brophy E, Muehlhausen W, Smeaton AF, Ward TE. Optimised convolutional neural networks for heart rate estimation and human activity recognition in wrist worn sensing applications. *arXiv [Online]* 2020 [cited 2023 Apr 28]. Available from: <https://doi.org/10.48550/arXiv.2004.00505>.
 33. Kang JH, Cho BH, Lee JS, Chee YJ, Kim IY, Kim SI. A study on accelerometer based motion artifact reduction in photoplethysmography signal. *J Biomed Eng Res* 2007;28:369-76.
 34. Kappert K, Böhm M, Schmieder R, Schumacher H, Teo K, Yusuf S, et al; ONTARGET/TRANSCEND Investigators. Impact of sex on cardiovascular outcome in patients at high cardiovascular risk: analysis of the Telmisartan Randomized Assessment Study in ACE-Intolerant Subjects With Cardiovascular Disease (TRANSCEND) and the Ongoing Telmisartan Alone and in Combination With Ramipril Global End Point Trial (ONTARGET). *Circulation* 2012;126:934-41.
 35. Bots SH, Peters SAE, Woodward M. Sex differences in coronary heart disease and stroke mortality: a global assessment of the effect of ageing between 1980 and 2010. *BMJ Glob Health* 2017;2:e000298.
 36. Blok S, Piek MA, Tulevski II, Somsen GA, Winter MM. The accuracy of heartbeat detection using photoplethysmography technology in cardiac patients. *J Electrocardiol* 2021;67:148-57.