scientific reports

OPEN

Validity of Hera Leto consumer earbuds for heart rate monitoring during stationary cycling

Junchao Yang[1,2], Zhihui Lu[3], Xiangxin Li[1], Kuan Tao[4]✉ & Junqiang Qiu[1,2]✉

Heart rate (HR) monitoring using earbuds presents a promising avenue for precise measurement, yet its validity remains under investigation during stationary cycling. In this study, we assessed the validity of the Hera Leto consumer earbuds by enrolling twenty-eight athletes as participants. Each participant wore a chest-strap Polar H10, considered the gold-standard reference, while simultaneously testing three HR monitors: the Hera Leto (worn on the ears), the Polar Verity Sense (worn on the upper arm), and the Polar Vantage V (worn on the wrist). The exercise protocol included four steady-state exercise sessions, six sessions of severe intensity exhaustive exercise, and three high-intensity interval training (HIIT) sessions. Our findings revealed a robust correlation (r=0.970, P<0.001) between the HR measurements obtained from Hera Leto and Polar H10 across all exercise intensities, with a minimal systematic bias of – 0.67 bpm. with a minimal systematic bias of -0.67 bpm. Notably, Hera Leto demonstrated superior performance, with a mean absolute error (MAE) of 2 bpm and concordance correlation coefficient (CCC) of 0.97, compared to the Polar Verity Sense (MAE = 4 bpm, CCC = 0.88) and Polar Vantage V (MAE = 4 bpm, CCC = 0.93). Although the precision of HR measurement decreased with increasing exercise intensity, Hera Leto exhibited a slower onset of this decline compared to the other devices. Overall, our study underscores the potential of Hera Leto earbuds as a valid tool for HR monitoring, particularly in scenarios involving steady-state exercise.

The measurement of heart rate (HR) in beats per minute (bpm) is a fundamental physiological parameter derived from assessing the temporal duration between successive heart cycles initiated by the sinoatrial node[1]. HR serves as a crucial indicator for evaluating both general health and athletic performance[2–4]. Traditionally, HR is assessed using electrocardiography (ECG)[5], utilizing either multiple-lead channels or basic chest straps equipped with two electrodes. However, this method is not ideal for prolonged monitoring during daily activities[6]. Photoplethysmography (PPG), especially when utilizing green light[7], has gained increasing popularity in consumer wearables such as watches, earbuds, armbands, and rings. This technology, known for its reduced susceptibility to motion artifacts compared to infrared, significantly enhances the accuracy of optical HR monitors and has established PPG as the standard method for continuous fitness monitoring[7].

Among various consumer devices incorporating optical HR monitoring, earbuds have emerged as an optimal location for integrating physiological signal sensors, offering many advantages over conventional wrist and finger recording sites[8,9]. For example, Since 2007, when research initially demonstrated the feasibility of collecting HR data from the ear[10], subsequent studies have revealed that the quality of in-ear PPG signals is comparable to those acquired from the finger[11,12]. Furthermore, in-ear PPG offers a higher beat detection rate and more accurate HR estimation compared to rings[13]. From a physiological perspective, the ear canal can resist the impact of sweat on PPG signal quality during exercise. Additionally, the ear probe is situated near the carotid artery that supplies blood to the brain, it can reduce the delay in measuring blood oxygen levels[14]. However, it is still unknown whether earbuds can accurately measure HR changes during stationary cycling without being affected by motion artifacts and low perfusion issues.

Apart from this, despite the significant expansion in the market share of commercial HR earbuds in recent years, the bulk of research has predominantly focused on validating the efficacy of non-commercial in-ear devices. For instance, studies have scrutinized disparities in validity between wakefulness and sleep states[13], as well as the precision during surgical procedures[15]. While a few investigations have delved into the performance of specific commercial earbuds during stationary cycling exercises, notable variances persist across different brands and devices[9]. Furthermore, studies reveal that the accuracy of HR measurements using wrist-worn

[1]School of Sport Science, Beijing Sport University, Beijing 100084, China. [2]Beijing Sports Nutrition Engineering Research Center, Beijing 100084, China. [3]China Football College, Beijing Sport University, Beijing 100084, China. [4]School of Sports Engineering, Beijing Sport University, Beijing 100084, China. ✉email: taokuan@bsu.edu.cn; qiujunqiang@bsu.edu.cn

PPG devices deteriorates as exercise intensity escalates[16]. Moreover, various types of activities can impact the accuracy of HR measurements[17,18]. Thus, it is imperative to assess the effectiveness of HR measurements across diverse activities and intensities.

In summary, the rigorous scientific validation of commercial devices, particularly for research purposes, is paramount before their widespread adoption. Despite previous evaluations of in-ear devices for heart rate (HR) measurement, Hera Leto—a state-of-the-art in-ear device utilizing PPG technology—lacks specific validity assessments. Therefore, this study aims to assess the validity of Hera Leto in measuring HR across various activities and intensities by employing the Polar H10 as a standard reference. Additionally, the study seeks to compare the HR measurement accuracy of Hera Leto with that of the wrist-worn PPG device Polar Vantage V and the arm-worn PPG device Polar Verity Sense..

Methods
Participants
Twenty-eight national first-level athletes, including twelve females and sixteen males specializing in track and field and skiing disciplines, were successfully recruited for this study, as outlined in Table 1. The sample size was guided by similar studies, where a comparable number of participants was deemed adequate for evaluating the validity of HR monitoring technologies[19,20]. All participants were nonsmokers, had a normal BMI, and reported no use of medications that could influence heart rate or exercise performance, such as beta blockers or antiarrhythmic agents. Female participants confirmed they were not using hormonal contraceptives and reported a consistent menstrual cycle length of 28 ± 2 days. Prior to the study, participants were fully briefed on the procedures, along with the associated risks and benefits, and provided their informed consent. The study protocol was approved by the Conjoint Health Research Ethics Board at Beijing Sport University (NO. 2024016 H), adhering to the standards of the Declaration of Helsinki.

Study design
Participants in this study were required to visit 12 to 13 times to the laboratory over an 8-week period, ensuring a minimum of 48 h between each visit and tests scheduled consistently at specific times of the day. During the initial visit, participants first underwent assessments of height, weight, and body composition, followed by a ramp-incremental exercise test. Body composition was measured using bioelectrical impedance analysis (InBody 230, InBody, Seoul, Korea). Based on the results of the incremental test, appropriate intensities were determined for four steady-state exercises, six severe intensity exhaustive exercises, and three high-intensity interval training (HIIT) sessions. All experimental trials were conducted using a cycle ergometer (Monark 839E, Monark, Vansbro, Sweden) in a climate-controlled room, maintaining temperatures between 18 °C and 21 °C and humidity levels between 50% and 60%. Participants were instructed to avoid consuming any food or caffeine-containing beverages for at least 2 and 8 h, respectively, prior to each visit, and to refrain from intense physical activity for 24 h before each session. During all tests, metabolic gases were collected using a gas analyzer (MetaMax 3B, Cortex Biophysic, Leipzig, Germany). The analyzer was calibrated before each test according to the manufacturer's guidelines, including calibration for pressure, gas concentrations (standard gases: O_2 = 15.00%, CO_2 = 5.00%), and volume (3 L syringe).

Devices
Polar H10
The Polar H10 served as the criterion device in this study, showcasing its validity with a strong correlation of r =0.997 in comparison to ECG[21]. As a result, the Polar H10 has become a standard reference in validating the accuracy of numerous commercial HR monitoring devices[17]. The HR sensor was attached to a Polar Pro HR strap, which was positioned over the sternum. Real-time HR data from the Polar H10 were transmitted to the Polar Flow app, which captured HR readings at 1-second intervals.

Polar verity sense
The Polar Verity Sense was secured to the nondominant forearm using an armband, in accordance with the manufacturer's instructions. The sensor was placed on the upper forearm or upper arm, with the lens facing inward and positioned firmly against the skin on the underside of the armband. The strap notch used for securing

Characteristics	Mean ± SD (n = 28)
Age (*yr*)	20.39 ± 1.18
Height (*cm*)	174 ± 8
Body weight (*kg*)	65.06 ± 9.19
BMI (kg/m^2)	21.28 ± 1.71
Percentage of body fat (%)	13.79 ± 5.29
VO_2max (mL/kg/min)	43.98 ± 5.91
First ventilatory threshold (mL/kg/min)	23.51 ± 4.97
Second ventilatory threshold (mL/kg/min)	36.29 ± 5.35
Maximum power output (*W*)	256 ± 43

Table 1. Participants' characteristics and ramp-incremental test responses.

the band was recorded and replicated across all visits to ensure consistent placement and tightness. Care was taken to ensure a snug yet comfortable fit, avoiding excessive pressure while minimizing sensor movement during exercise. The device was set to “Indoor Cycling” mode using the Polar Flow app, activating the heart rate sensor mode for continuous real-time Bluetooth streaming. HR data were captured at 1-second intervals through six light-emitting diode sensors and transmitted live via Bluetooth to a smartphone equipped with the Polar Flow app (Polar Electro Oy). Upon completion of each visit, the data were uploaded to the Polar Flow web service (Polar Electro Oy).

Polar vantage V
The Polar Vantage V was securely attached to the nondominant wrist, positioned one finger width above the ulnar styloid process, according to the manufacturer’s instructions. The wristband was fastened firmly to ensure that the back sensor remained in constant contact with the skin and that the device did not shift during movement. The strap notch position was recorded and consistently used across all visits. Placement was verified by trained staff using a fit-check protocol recommended by the manufacturer: gently pressing both sides of the band to confirm that no green LED light was visible and that the sensor stayed in full contact with the skin. The device was set to “Indoor Cycling” mode using the device’s onboard interface to standardize LED intensity and HR recording settings. HR data were recorded at a frequency of 1-second intervals. The device utilizes Polar Precision Prime sensor fusion technology for HR measurement. After each visit, the collected data were uploaded to the Polar Flow web service (Polar Electro Oy).

Hera Leto
Hera Leto is developed and commercially sold by Actywell Digital Limited based in Hong Kong, China. The earbuds utilize multi-LED and multi-photodiode technology and employ multi-wavelength PPG (green light: 940 nm; red light: 625 nm) to measure pulse rate within the ear canal through reflection. In contrast, the specific wavelength used by the green PPG sensors in the Polar Verity Sense and Polar Vantage V has not been publicly disclosed by the manufacturer. Hera Leto is an in-ear earbud designed to fit snugly in the right ear, with the Ear-Gel nozzle pointing downward as recommended in the manual. The earbuds come with three sizes of sports Ear-Gel (S, M, L) for personal choice. During data collection, the Hera Leto HR app is used to acquire data, which is stored in files named according to the format: < *device id*> *__YYYY_MM_DD_hh_mm_ss_*< *file_type*>*.csv*, with a time resolution of 1-second intervals.

Study procedures

Ramp-incremental exercise
The ramp incremental test aims to assess maximal oxygen uptake ($VO_{2\,max}$), identify the first ventilatory threshold (VT_1), the second ventilatory threshold (VT_2) and determine the maximum power output (P_{max}) reached during the test[22]. Ventilatory thresholds were determined using the online platform https://www.exercisethresholds.com, which integrates multiple gas exchange-based detection methods. Two investigators independently identified the thresholds, and results were confirmed by consensus. During their initial visit to the laboratory, participants underwent a 4-minute warm-up cycling at 50 W, followed by an incremental ramp test with a 30 W per minute increase until exhaustion. At the start of the test, participants selected a cadence between 60 and 90 bpm and were instructed to maintain this cadence consistently throughout the test. This selected cadence was also required for all subsequent tests. The ramp test concluded when a participant voluntarily stopped due to exhaustion or when they were unable to sustain their chosen cadence for more than five seconds, despite receiving significant encouragement. This exhaustion criterion remained consistent across all tests designed to assess the limits of exhaustion.

Steady-state exercise
Participants in this study undertook four stages of constant load exercises at intensities corresponding to 45%, 55%, 65%, and 75% of the power output at VO_2max (Pmax), each stage lasting 10 min. The participants were divided into two groups for laboratory testing, with one group randomly assigned to the 45% and 75% Pmax conditions, and the other to the 55% and 65% Pmax conditions. Following completion of each intensity level, participants engaged in active rest periods lasting approximately 20–30 min before proceeding to the subsequent level. The average heart rate (HR) during the 8 th to 10 th minute of each load level was computed and then linearly correlated with the respective load intensity (power).

Severe intensity exhaustive exercise
Each participant was required to complete at least three constant load tests to exhaustion. These tests began with an initial 4-minute warm-up at 25 watts, immediately followed by an increase to a predetermined power output (PO). Based on the Pmax determined during the maximal oxygen uptake test, intensities are set at VT_2, 100%, 110%, 130%, 150%, and up to 170% of the Pmax. Participants were provided verbal encouragement to maintain maximum effort throughout the exercise.

High intensity intermittent training
We conducted three types of intermittent exercises, each lasting a total of 4 min with a fixed 10-second rest interval between activity bouts. The first type consisted of 10 s of activity followed by 10 s of rest, repeated 12 times ($HIIT_{10/10}$). The second type involved 20 s of activity followed by 10 s of rest, repeated 8 times ($HIIT_{20/10}$). The third type comprised 30 s of activity followed by 10 s of rest, repeated 6 times ($HIIT_{30/10}$). The exercise intensities were set at 100% of VO_2max for $HIIT_{10/10}$, 90%VO_2max for $HIIT_{20/10}$, and 85%VO_2max for $HIIT_{30/10}$, respectively.

Heart rate kinetics

This study focused on examining the differences in HR kinetics during various phases of exercise, specifically the ON phase (during exercise) and the OFF phase (during recovery). The analysis was limited to steady-state exercise and post-exercise periods due to the unsuitability of current kinetic equations for extreme exercise conditions[23]. The HR kinetics during exercise followed a pattern similar to that of VO_2, albeit with noticeable differences[24]. The HR kinetics during the exercise (ON phase) was modeled using two exponential terms: $\mathrm{HR}\ (\mathrm{t}) = \mathrm{HRbaseline} + \mathrm{A} \times\ (1 - e^{-\mathrm{t}/\tau\,1})$. Consequently, the HR during recovery (OFF phase) was described by a model incorporating two exponential terms as well: $\mathrm{HR}\ (\mathrm{t}) = \mathrm{HREnd} - \mathrm{Ex} + \mathrm{A} \times\ (1 - e^{-(\mathrm{t}-\mathrm{TD})/\tau\,1})$. HRbaseline and HREnd-Ex represent the starting heart rates at the beginning of exercise, while TD accounts for the time delay before recovery begins. A represents the amplitude of heart rate change, and τ1 is the time constant that determines the rate of recovery—the smaller the value, the faster the recovery.

Data analysis

Data processing and analysis in this study were conducted in accordance with the HR validity verification guidelines published by the Towards Intelligent Health and Well-Being: Network of Physical Activity Assessment (INTERLIVE)[6]. Validity comparisons were made between second-by-second data from Hera Leto, Polar Verity Sense, and Polar Vantage V against data from Polar H10.

Statistical analyses were performed using SPSS 21.0, with results presented as mean ± standard deviation ($x \pm s$). The normality of the data was tested using the Shapiro-Wilk test. Differences in HR measurements for $HRVT_1$, $HRVT_2$, and the maximum HR (HRmax) across devices were assessed using one-way repeated measures ANOVA, with pairwise comparisons conducted using the Bonferroni post hoc test. A significance level of 0.05 was set for statistical tests. Prior to analysis, all data underwent missing value testing; datasets with less than 10% missing or anomalous data points were retained[16].

The relevant libraries were imported in PyCharm to facilitate data processing, mathematical operations, and statistical analyses, including Bland-Altman and repeated measures correlation analyses. Using data from four devices, we fitted a linear equation between HR and power. The accuracy and quality of these fitted equations were assessed using the Standard Error of Estimate (SEE), where a lower SEE value indicates a better fit, reflecting a closer alignment between the observed HR and the model's predictions. Furthermore, these libraries were used to compute evaluation metrics such as Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), and Concordance Correlation Coefficient (CCC) for assessing the effectiveness of Hera Leto, Polar Verity Sense, and Polar Vantage V[6]. Finally, Excel was employed for calculating dropout rate. For validity thresholds, we used a MAPE value of less than 5% and a CCC greater than 0.90[25].

Results

Validity of HR measurement across all activities and intensities

Hera Leto demonstrated high validity in HR measurement across different exercise intensities, as shown in Table 2. While Verity Sense and Vantage V performed comparably under steady-state and HIIT conditions, their HR readings showed greater deviations during severe-intensity constant load exercise (≥ 130% Pmax), particularly during the ON phases.

Table 3 presents the aggregated validity metrics across all exercise intensities. Hera Leto exhibited the lowest overall MAE and MAPE values and the highest CCC, compared to Polar Verity Sense and Polar Vantage V. However, as shown in Table 2, the differences in device performance varied across specific exercise conditions, with comparable accuracy observed among the devices during steady-state and HIIT exercises.

In the Bland-Altman plots, the Hera Leto had lowest bias (ON=−2.24 bpm, OFF=−0.15 bpm, ALL=−0.67 bpm) in estimating HR compared to the other device (Polar Verity Sense: ON=−5.54 bpm, OFF=−0.22 bpm, ALL=−1.56 bpm; Polar Vantage V: ON=−4.34 bpm, OFF = 0.52 bpm, ALL=−0.70). Hera Leto (ON: **r** = 0.970, *P* < 0.001, OFF: **r** = 0.957, *P* < 0.001, ALL: **r** = 0.970, *P* < 0.001) had the strongest correlation with Polar H10 in estimating HR compared to other devices (Polar Verity Sense, ON: **r** = 0.781, *P* < 0.001, OFF: = 0.887, *P* < 0.001, ALL: **r** = 0.878, *P* < 0.001; Polar Vantage V:, ON: **r** = 0.865, *P* < 0.001, OFF: **r** = 0.933, *P* < 0.001, ALL: **r** = 0.927, *P* < 0.001) (Fig. 1).

Validity of HR measurement across various intensities

As exercise intensity increases, the VT_2 serves as a distinct demarcation. Below this intensity, the accuracy of measurements from Hera Leto, Polar Verity Sense, and Polar Vantage V is very high, with very similar MAE, MAPE, and CCC. However, above this intensity, the effectiveness of HR measurements from all devices decreases, although Hera Leto shows the slowest rate of decline, but it has the highest dropout rate (Fig. 2).

HR response for specific exercise

Ramp-incremental exercise

Table 4 shows that during incremental load exercise, Hera Leto, Polar Verity Sense, and Polar Vantage V exhibit consistent performance in terms of MAE, MAPE, and CCC; however, Hera Leto generally has a higher dropout rate.

The Polar Verity Sense, Polar Vantage V, and Hera Leto show significant differences in measurements of HR_{peak} and HR-VT_2 when compared to the Polar H10 (*P* < 0.05). The differences in HR_{peak} are respectively 2 (95% CI: 1 to 3) bpm, 4 (95% CI: 2 to 6) bpm, and 2 (95% CI: 0 to 4) bpm. For $HRVT_2$, the differences are also 2 (95% CI: 1 to 3) bpm, 4 (95% CI: 2 to 6) bpm, and 2 (95% CI: 0 to 4) bpm. However, there are no significant differences in the measurements of HR-VT_1 between these devices and the H10. Additionally, there are no significant differences in the OFF-phase kinetic parameters measured during Ramp-Incremental Exercise between the Polar Verity Sense, Polar Vantage V, Hera Leto, and the Polar H10 (*P* > 0.05) (Table 5).

Type	Stage (*N* of paired observations)	Polar H10	Polar Verity Sense	Polar Vantage V	Hera Leto
45%VO_2max	ON (24)	120 ± 12	119 ± 12	119 ± 12	119 ± 12
	OFF (23)	103 ± 11	104 ± 11	104 ± 11	103 ± 11
55%VO_2max	ON (15)	135 ± 13	134 ± 15	134 ± 14	134 ± 14
	OFF (18)	111 ± 14	112 ± 14	112 ± 14	111 ± 14
65%VO_2max	ON (19)	142 ± 15	141 ± 16	141 ± 16	141 ± 16
	OFF (22)	112 ± 12	113 ± 13	113 ± 13	112 ± 13
75%VO_2max	ON (19)	149 ± 14	149 ± 14	150 ± 12	149 ± 13
	OFF (19)	113 ± 15	113 ± 16	113 ± 16	112 ± 15
VT_2	ON (15)	165 ± 19	155 ± 28	157 ± 25	162 ± 20
	OFF (19)	124 ± 16	125 ± 17	125 ± 17	124 ± 17
100%VO_2max	ON (17)	157 ± 23	127 ± 32	135 ± 28	150 ± 25
	OFF (21)	125 ± 19	125 ± 19	125 ± 19	125 ± 19
110%VO_2max	ON (22)	153 ± 22	118 ± 24	129 ± 23	148 ± 23
	OFF (25)	123 ± 18	121 ± 16	123 ± 17	122 ± 18
130%VO_2max	ON (26)	150 ± 22	109 ± 20	122 ± 25	140 ± 25
	OFF (25)	121 ± 17	119 ± 15	121 ± 17	120 ± 17
150%VO2 max	ON (20)	146 ± 23	105 ± 15	114 ± 15	133 ± 23
	OFF (24)	115 ± 21	116 ± 16	119 ± 18	118 ± 19
170%VO_2max	ON (19)	148 ± 23	107 ± 12	119 ± 19	132 ± 23
	OFF (24)	119 ± 17	116 ± 15	118 ± 16	119 ± 17
$HIIT_{10/10}$	ON (14)	142 ± 15	139 ± 18	138 ± 17	138 ± 19
	OFF (15)	97 ± 19	98 ± 21	97 ± 20	96 ± 20
$HIIT_{20/10}$	ON (20)	152 ± 18	145 ± 23	148 ± 21	149 ± 21
	OFF (19)	103 ± 19	103 ± 20	103 ± 20	102 ± 20
$HIIT_{30/10}$	ON (19)	153 ± 18	144 ± 27	148 ± 22	149 ± 22
	OFF (20)	104 ± 20	104 ± 21	104 ± 21	103 ± 21

Table 2. Mean and standard deviation in each intensity level.

Phase	Metrics	Polar Verity Sense	Polar Vantage V	Hera Leto
ON	MAE	7	6	3
	MAPE	4.86%	4.16%	2.28%
	CCC	0.74	0.84	0.96
OFF	MAE	3	3	2
	MAPE	2.79%	3.06%	2.12%
	CCC	0.88	0.93	0.96
ALL	MAE	4	4	2
	MAPE	3.13%	3.24%	2.14%
	CCC	0.88	0.93	0.97

Table 3. Validity of HR measurement with Polar H10 and other devices.

Steady-state exercise

Figure 3 demonstrates a high level of consistency in the linear equations for steady-state HR and power fitted using four devices. Table 6 displays the results of HR kinetics during steady-state exercise; as intensity increases, there are no significant differences in HR kinetics parameters between Hera Leto and Polar H10 ($P > 0.05$). However, there are significant differences between the Polar H10 and both Polar Verity Sense and Polar Vantage V in the parameters A or $\tau 1$ ($P < 0.05$). Additionally, there are no significant differences in the OFF-phase kinetic parameters of steady-state exercise measured at different intensities between Polar Verity Sense, Polar Vantage V, Hera Leto, and Polar H10 ($P > 0.05$).

Severe intensity exhaustive exercise

Figure 4 illustrates the OFF-phase kinetics for severe intensity exhaustive exercise. As the intensity increases, there are no significant differences in the kinetic parameters between Hera Leto and Polar H10. However, Polar Verity Sense and Polar Vantage V struggle to accurately capture the full HR kinetics, particularly the rapid recovery phase of HR.

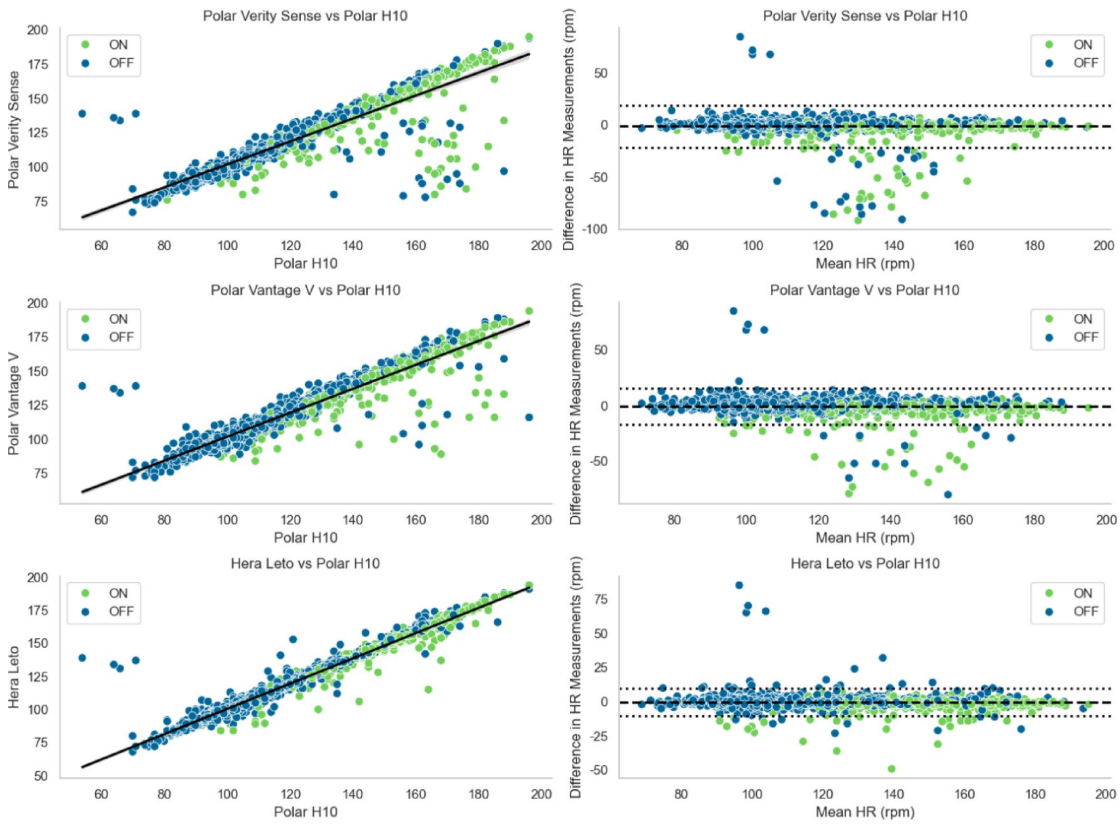

Fig. 1. Using Bland-Altman analysis with limits of agreement and repeated-measures correlation to compare HR from the Polar H10 with other devices during ON and OFF phases. For visualization purposes, 1% of the total second-by-second HR data was randomly sampled for the Bland-Altman plots, while all paired observations were used for statistical analysis.

High intensity intermittent exercise

Table 7 shows that the validity of HR measurements by Hera Leto, Polar Verity Sense, and Polar Vantage V—assessed using MAPE and CCC—remained within acceptable limits across HIIT intervals.

Figure 5 illustrates the MAE, MAPE, CCC, and Mean ± SD of HR measurements during the intervals and rest periods of HIIT, taken by different devices. As the number of HIIT sets increases, both the HRs measured by the devices and their validity improve, with Hera Leto showing validity closest to that of the Polar H10. Additionally, there are no significant differences in the OFF-phase kinetic parameters of HIIT measured at different intensities between Polar Verity Sense, Polar Vantage V, Hera Leto, and Polar H10 ($P > 0.05$).

Discussion

This study is the first to evaluate the accuracy of a commercial earbud device in monitoring heart rate (HR) during stationary cycling across different intensities and exercise types. Hera Leto showed relatively stable agreement with the criterion device (Polar H10) under varying conditions. In the recovery (OFF) phases of steady-state and HIIT sessions, its validity metrics (e.g., MAE, MAPE) remained within acceptable limits, comparable to those of Polar Verity Sense and Polar Vantage V. During the exercise (ON) phases, performance varied among devices, with larger errors emerging at higher intensities. In incremental load tests, all three devices remained within acceptable ranges overall, though accuracy declined at VO_2max and VT2. In constant load exercises, increasing intensity led to greater HR measurement errors across all devices, but Hera Leto maintained acceptable accuracy over a wider range. During low-volume HIIT, the validity of all three devices improved across intervals and generally remained acceptable. Notably, all devices exhibited consistent HR kinetics with Polar H10 during steady-state ON phases. However, in recovery phases, only Hera Leto showed HR kinetics closely matching those of Polar H10, while the other two devices underperformed, particularly during severe intensity exhaustive exercise.

Equipping with wearable devices to measure and further analyze physiological signals are gaining popularities in exercise[26–28]. Our research demonstrated that the validity of the Hera Leto, Polar Verity Sense, and Polar Vantage V devices deteriorates as exercise intensity increases. This pattern was confirmed across both incremental load intensity and different constant load intensities. Similar findings have been reported in several studies examining the validity of commercial wearable devices, noting an increase in HR measurement errors with rising exercise intensities[20,29–31]. For example, we observed that the HR measurements at VT_1 from all three devices were consistent with the Polar H10, while discrepancies were noted at VT_2 and VO_2max.

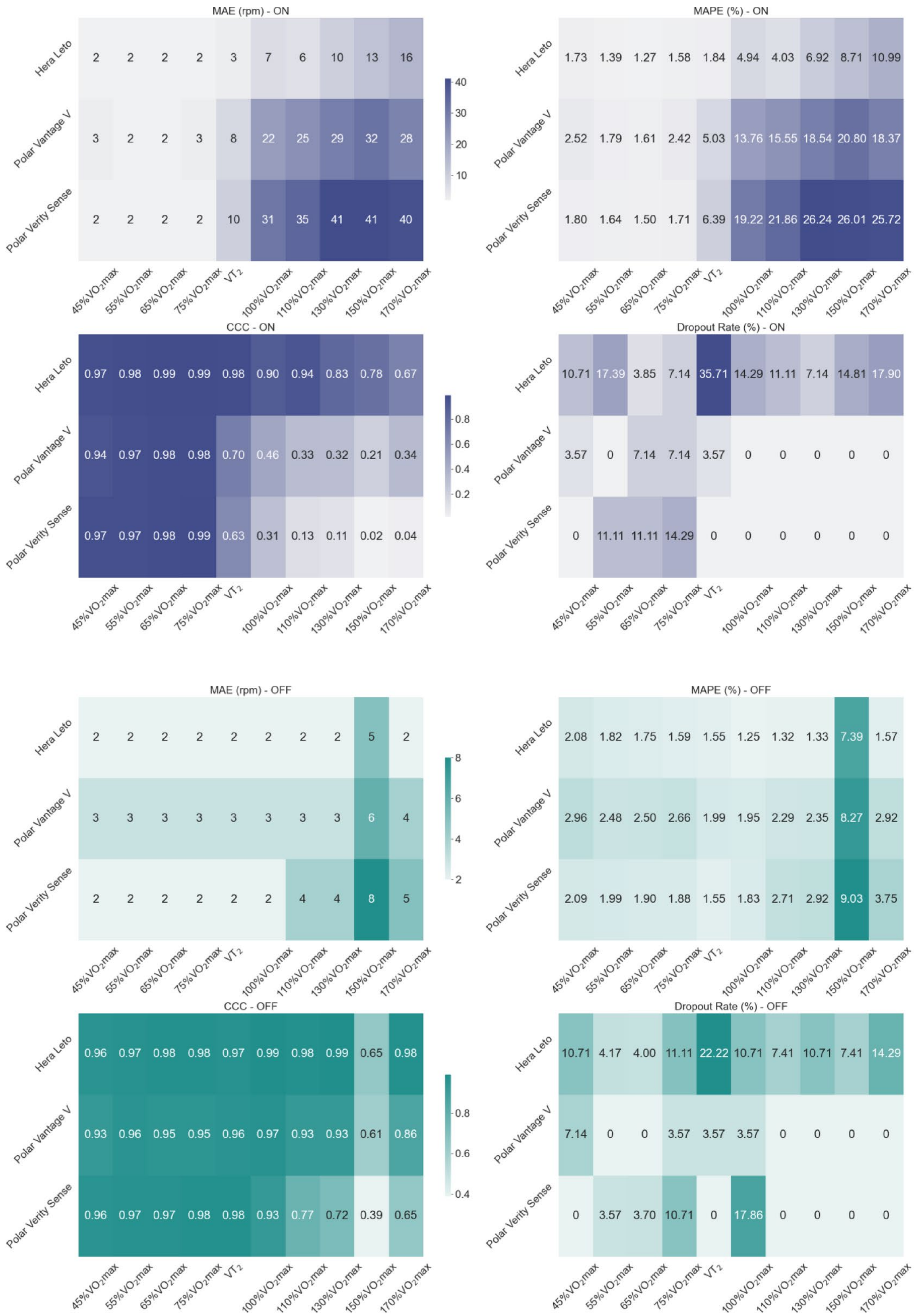

Fig. 2. The validity of an HR measuring device varies with intensity, with darker colors representing higher values and lighter colors indicating lower values for MAE, MAPE, CCC, and Dropout Rate.

Interestingly, studies have indicated that the Apple Watch possesses good to very good criterion validity for measuring HRmax. We hypothesize that the differences may arise from the modalities used in the incremental load tests: the Apple Watch study utilized a treadmill, whereas our tests were conducted on a cycle ergometer. This is supported by findings from Muggeridge et al.[16], who noted that compared to treadmill exercises, the validity of wrist-worn PPG devices was particularly poor during stationary cycling exercises. It's worth noting

Phase	Metrics	Polar Verity Sense	Polar Vantage V	Hera Leto
ON	MAE (bpm)	2	3	2
	MAPE (%)	1.93	2.65	1.44
	CCC	1	0.99	1
	Dropout Rate (%)	7.41	7.14	20.83
OFF	MAE (bpm)	3	2	2
	MAPE (%)	2.22	1.42	1.87
	CCC	0.99	1	1
	Dropout Rate (%)	3.57	16.67	3.7

Table 4. Validity of HR measurement in ramp-incremental exercise.

	HR-VT_1	HR-VT_2	HRpeak
Polar H10	139 ± 15	169 ± 11	185 ± 10
Polar Verity Sense	138 ± 15	167 ± 10*	183 ± 11*
Polar Vantage V	137 ± 15	166 ± 10*	181 ± 10*
Hera Leto	138 ± 14	167 ± 11*	182 ± 11*

Table 5. HRpeak, HR-VT1 and HR-VT2 across the HR devices. "*" indicates that compared to Polar H10, $P < 0.05$.

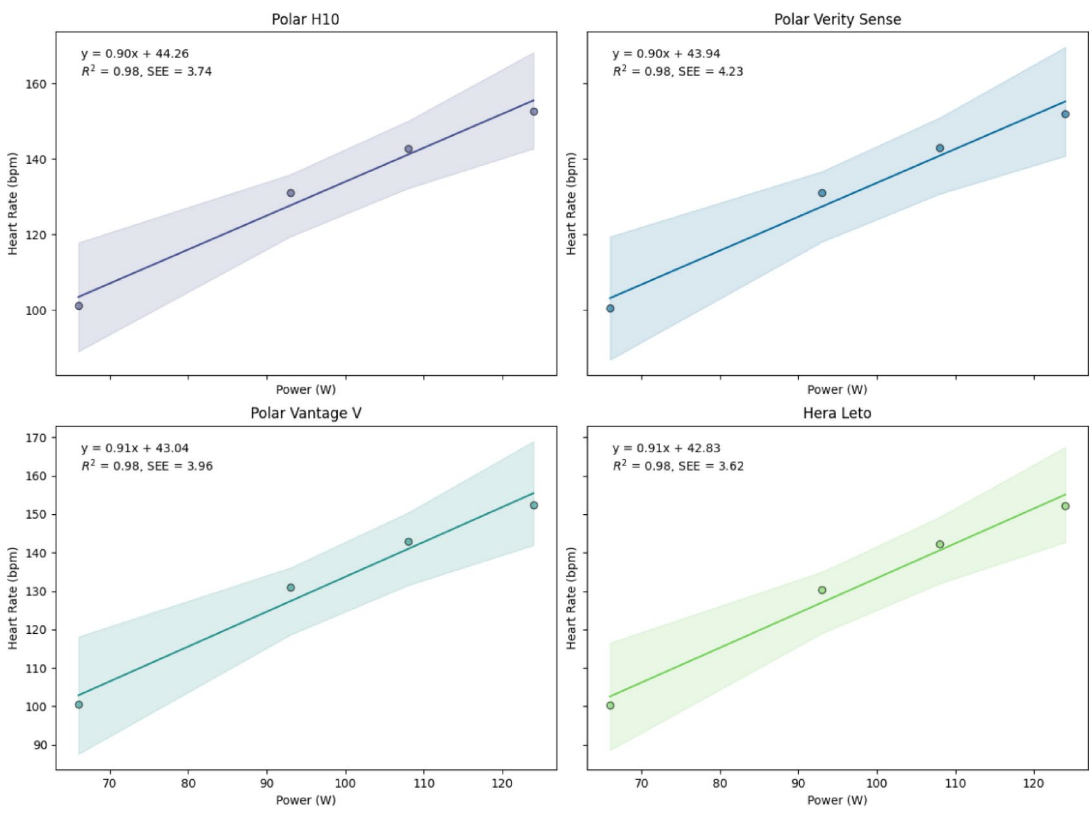

Fig. 3. The linear equations of HR versus power measured by different devices.

that throughout the incremental load exercises, the HR validity measurements of Hera Leto, Polar Verity Sense, and Polar Vantage V all remained within acceptable ranges, consistent with the findings reported by Wallen et al.[32]. We only tentatively speculate that VT_2 might represent the turning point where validity begins to decline, given the large intervals between our incremental intensities. Referring to Jo et al.[29], who found this turning point to be at 116 bpm (ECG HR), our study suggests it may occur after the HR corresponding to VT_1 (Polar H10: 139 ± 15). Lastly, notably, Hera Leto exhibits a later turning point for error reduction compared to Polar

Type	Parameter	Polar H10	Polar Verity Sense	Polar Vantage V	Hera Leto
45%VO_2max	$HRbaseline$ (bpm)	98 ± 9	97 ± 10	97 ± 10	97 ± 9
	A (bpm)	27.22 ± 9.52	28.04 ± 10.41	27.67 ± 9.68	27.46 ± 9.98
	τ (s)	111.09 ± 66.20	124.97 ± 61.97 *	121.88 ± 60.60	119.72 ± 55.16
	RMSE	3 ± 1	3 ± 1	3 ± 1	3 ± 1
	R^2	0.73 ± 0.15*	0.79 ± 0.12*	0.77 ± 0.13*	0.75 ± 0.15*
55%VO_2max	$HRbaseline$ (bpm)	98 ± 9	96 ± 8	97 ± 8	97 ± 8
	A (bpm)	44.05 ± 6.75	46.46 ± 6.98	45.76 ± 6.88	44.66 ± 6.76
	τ (s)	91.21 ± 36.09	112.67 ± 50.72	112.97 ± 46.61*	106.18 ± 46.79
	RMSE	3 ± 1	3 ± 1	3 ± 1	3 ± 1
	R^2	0.89 ± 0.06	0.92 ± 0.04*	0.92 ± 0.05*	0.91 ± 0.05
65%VO_2max	$HRbaseline$ (bpm)	98 ± 9	97 ± 8	97 ± 7	97 ± 8
	A (bpm)	53.50 ± 10.44	54.30 ± 10.78	54.69 ± 9.81	54.52 ± 10.16
	τ (s)	102.37 ± 31.96	114.60 ± 36.79*	119.15 ± 40.30*	110.24 ± 37.61
	RMSE	3 ± 1	3 ± 1*	3 ± 1	3 ± 1
	R^2	0.93 ± 0.03	0.95 ± 0.02*	0.95 ± 0.02*	0.93 ± 0.03*
75%VO_2max	$HRbaseline$ (bpm)	105 ± 9	104 ± 9	104 ± 9	104 ± 9
	A (bpm)	54.02 ± 8.56	55.65 ± 8.62*	55.73 ± 8.21*	54.02 ± 8.56
	τ (s)	97.31 ± 31.84	112.67 ± 32.49*	112.53 ± 28.08*	104.27 ± 30.46*
	RMSE	3 ± 1	3 ± 1	3 ± 1	3 ± 1
	R^2	0.94 ± 0.03	0.95 ± 0.03	0.96 ± 0.02*	0.94 ± 0.03

Table 6. Fitting parameters for HR kinetics during the ON phase of steady-state exercise. "*" indicates that compared to Polar H10, $P < 0.05$.

Verity Sense and Polar Vantage V. This suggests that ear-based devices like Hera Leto may offer advantages in sustaining HR measurement accuracy under increasing exercise intensities, particularly in high-performance and endurance training settings.

This study identified that Hera Leto demonstrated overall better validity during HIIT compared to Polar Verity Sense and Polar Vantage V. However, all three devices exhibited unacceptable measurement errors in the initial sets of HIIT. Across the three HIIT formats, the validity of HR measurements—assessed via MAPE, MAE, and CCC—remained within acceptable limits. However, a slight decline in accuracy was observed as interval duration increased (e.g., $HIIT_{30/10}$), particularly in the CCC values of Verity Sense and Vantage V. We speculate that this phenomenon may be attributed to similar factors that lead to increased HR errors during extreme exercise. Firstly, during HIIT, participants tend to grip the bike handles tightly and contract upper arm muscles, which can introduce noise into HR estimation and reduce overall HR measurement validity[19,31]. Secondly, at the onset of HIIT, devices located on the arms, wrists, and in the ears may not be as sensitive to sudden changes in exercise intensity, resulting in lower validity of HR measurement. However, ear-based devices may exhibit higher sensitivity to changes in exercise intensity compared to the other two locations[9]. Lastly, the improved validity of all three devices with increasing sets may be attributed to rapid blood flow redistribution during rest periods. Conversely, longer interval HIIT, due to prolonged ischemic periods, may not achieve sufficient recovery within the same rest time, leading to persistent HR measurement errors. This also partly explains why all three devices exhibited poorer validity in measuring HR during sustained HIIT and sprint interval exercises[16].

The evaluation of Hera Leto's HR measurement using dynamic HR equations showed that it yielded more favorable validity metrics—such as lower MAPE and higher CCC—compared to Polar Verity Sense and Polar Vantage V, particularly during certain HIIT sessions. Additionally, we found that the HR kinetics during steady-state exercise and post-exercise, as measured by Hera Leto, Polar Verity Sense, and Polar Vantage V, could be fitted to HR dynamic equations. For the majority of activities, especially during rest, low to moderate intensity exercises, and post-exercise recovery phases, the HR measurement devices were generally effective[18]. Ear-based PPG devices, such as Hera Leto, hold potential as a solution to improve HR measurement validity during HIIT. This is because ear-based PPG devices typically employ multi-channel monitoring, which enhances the signal-to-noise ratio and coverage of PPG[9]. Moreover, during severe stress conditions, the body maximizes blood flow to maintain adequate cerebral perfusion, and in-ear PPG devices, being situated near the highly sensitive brain, can measure HR with minimal delay[12]. Although PPG is susceptible to constriction of peripheral blood flow in low temperatures, in-ear PPG has demonstrated resistance to these effects due to the unique characteristics of the ear canal[33].

Limitations of the study

The validity of the Hera Leto device was comprehensively evaluated across various activity types and intensities, comparing its performance with commonly used PPG devices. However, this study has several limitations. One key limitation is the relatively small sample size, primarily due to the high dropout rate associated with the Hera Leto device, which reduced the number of paired samples available for effectiveness analysis. Additionally, the study lacks reliability data, making it difficult to draw definitive conclusions about the validity of the tested

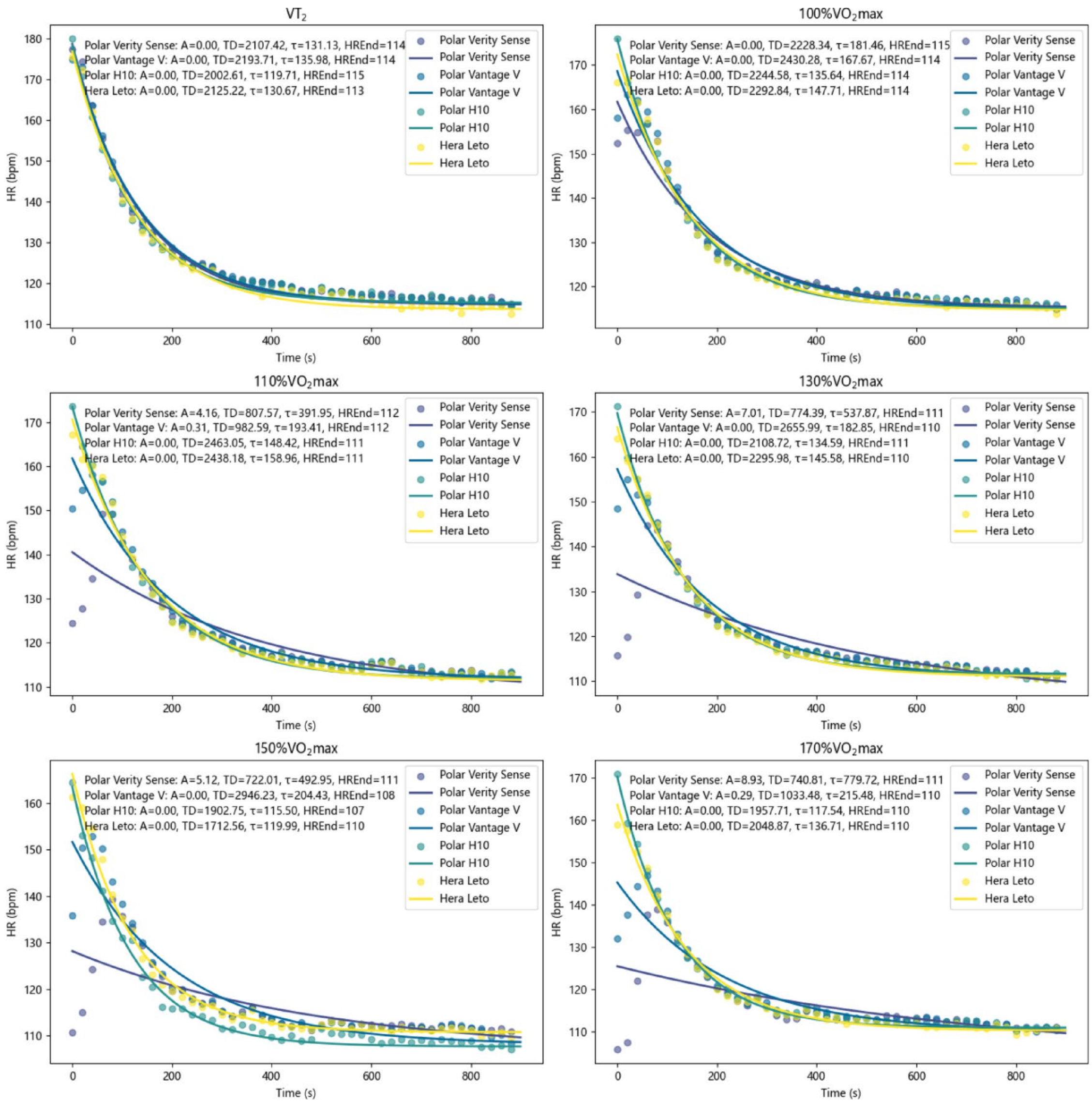

Fig. 4. Fitting Parameters A, TD, and τ1 for HR kinetics during the OFF phase of severe intensity exhaustive exercise.

devices. While the controlled laboratory setting allowed for precise comparison across devices, it may not fully capture the complexity of real-world outdoor exercise scenarios, such as road cycling. Additionally, the study sample consisted of trained university athletes, which may limit generalizability to clinical populations or older individuals. Factors such as skin temperature and skin tone, which may influence PPG signal quality, were not specifically controlled and should be considered in future studies[9]. Moreover, the absence of detailed information on the engineering design and algorithmic processing of the Hera Leto device prevents a thorough analysis of the observed differences from a technical perspective. To minimize variability, all devices were kept on the same firmware version throughout the study, and automatic updates were disabled.

Type	Phase	Metrics	Polar Verity Sense	Polar Vantage V	Hera Leto
$HIIT_{10/10}$	ON	MAE (bpm)	2	3	2
		MAPE (%)	1.93	2.65	1.44
		CCC	1	0.99	1
		Dropout Rate (%)	7.41	7.14	20.83
	OFF	MAE (bpm)	3	2	2
		MAPE (%)	2.22	1.42	1.87
		CCC	0.99	1	1
		Dropout Rate (%)	3.57	16.67	3.7
$HIIT_{20/10}$	ON	MAE (bpm)	4	5	5
		MAPE (%)	3.16	3.39	3.58
		CCC	0.91	0.92	0.87
		Dropout Rate (%)	21.43	0	4.17
	OFF	MAE (bpm)	2	3	2
		MAPE (%)	2.37	3.42	2.04
		CCC	0.98	0.97	0.99
		Dropout Rate (%)	11.11	0	4.35
$HIIT_{30/10}$	ON	MAE (bpm)	8	6	5
		MAPE (%)	5.42	4.54	3.4
		CCC	0.78	0.86	0.91
		Dropout Rate (%)	3.7	0	8
	OFF	MAE (bpm)	2	3	2
		MAPE (%)	2	2.98	1.86
		CCC	0.99	0.98	0.99
		Dropout Rate (%)	3.7	0	16

Table 7. Validity of HR measurement in HIIT.

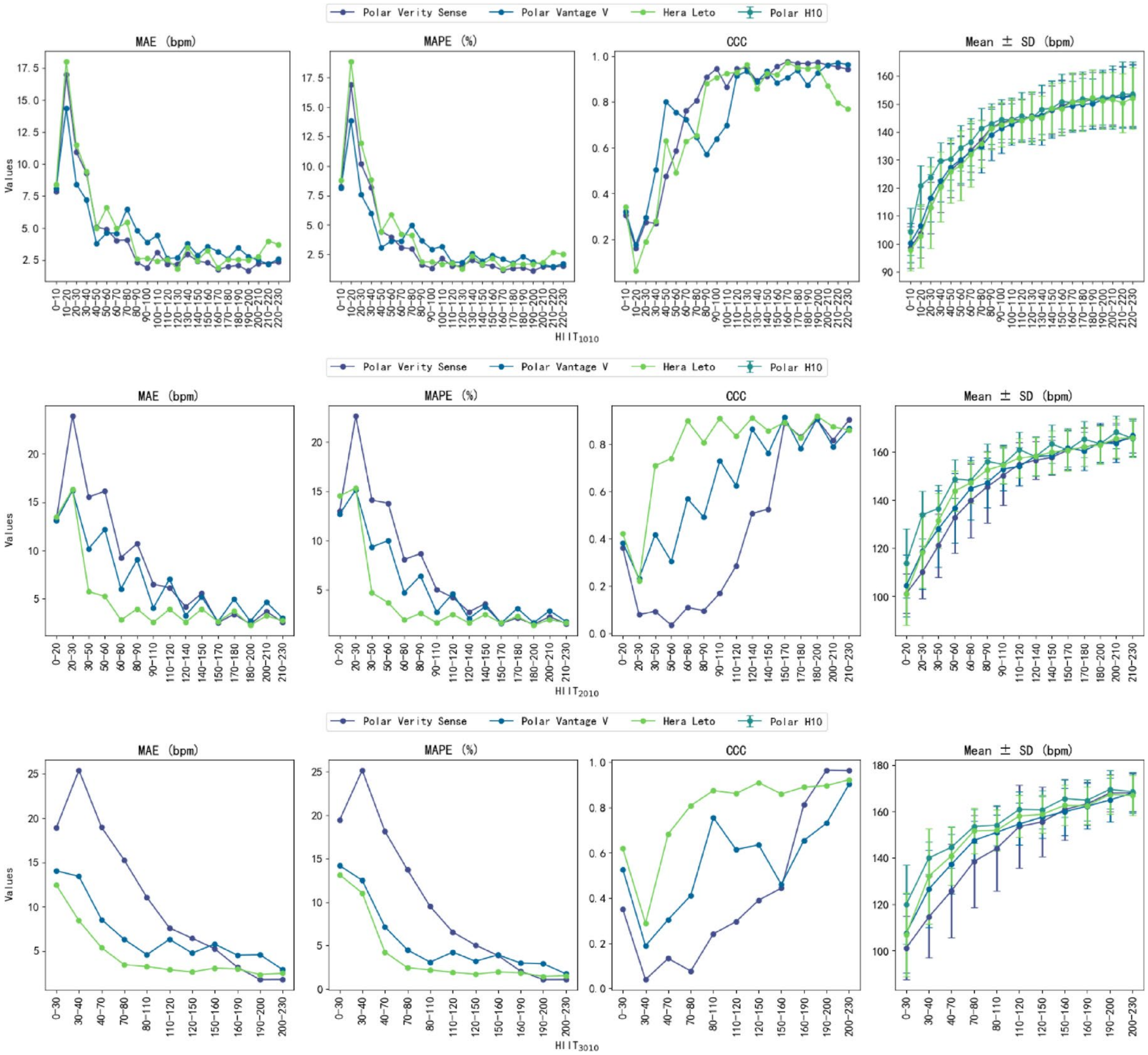

Fig. 5. The validity of different devices in measuring HR during the intervals and rest periods of HIIT.

Data availability

Data are available from the corresponding author upon reasonable request.

Received: 30 September 2024; Accepted: 29 April 2025
Published online: 07 May 2025

References

1. Keith, A. & Flack, M. The form and nature of the muscular connections between the primary divisions of the vertebrate Heart. *J. Anat. Physiol.* **41** (Pt 3), 172–189 (1907).
2. Ross, R. et al. Importance of assessing cardiorespiratory fitness in clinical practice: A case for fitness as a clinical vital sign: A scientific statement from the American heart Association. *Circulation* **134** (24), e653–e699 (2016).
3. Buchheit, M. Monitoring training status with HR measures: do all roads lead to Rome?. *Front. Physiol.* **5**, 73 (2014).
4. Tao, K. et al. Estimation of heart rate using regression models and artificial neural network in Middle-Aged Adults. *Front. Physiol.* **12**, 742754 (2021).
5. Mu, S. et al. Intelligent fatigue detection based on hierarchical multi-scale ECG representations and HRV measures. *Biomed. Signal Process. Control.* **92**, 106127 (2024).
6. Mühlen, J. M. et al. Recommendations for determining the validity of consumer wearable heart rate devices: expert statement and checklist of the INTERLIVE Network. *Br. J. Sports Med.* **55** (14), 767–779 (2021).
7. Maeda, Y., Sekine, M. & Tamura, T. The advantages of wearable green reflected photoplethysmography. *J. Med. Syst.* **35** (5), 829–834 (2011).
8. He, D. D. & Winokur ES,Sodini, C. G. An Ear-Worn vital signs Monitor. *IEEE Trans. Biomed. Eng.* **62** (11), 2547–2552 (2015).
9. Charlton, P. H. et al. The 2023 wearable photoplethysmography roadmap. *Physiol. Meas.* **44**(11). (2023).
10. Vogel, S. Hülsbusch & Starke, M. D, et al. In-ear heart rate monitoring using a micro-optic reflective sensor. *Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.* 1375–1378. (2007).
11. He, D. D. et al. The ear as a location for wearable vital signs monitoring. *Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.* 6389–6392 (2010).

12. Venema, B. et al. Advances in reflective oxygen saturation monitoring with a novel in-ear sensor system: results of a human hypoxia study. *IEEE Trans. Biomed. Eng.* **59**(7), 2003–2010 (2012).
13. Haddad, S., Boukhayma, A. & Caizzone, A. Ear and finger PPG wearables for night and day Beat-to-Beat interval Detection. *Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.* 1686–1689. (2021).
14. Mase, M., Micarelli A,Strapazzon, G. & Hearables New perspectives and pitfalls of In-Ear devices for physiological monitoring: scoping review. *Scoping Review Front. Physiol.* **11**, 568886 (2020).
15. Ellebrecht, D. B. & Gola D,Kaschwich, M. Evaluation of a wearable in-Ear sensor for temperature and heart rate monitoring: A pilot Study. *J. Med. Syst.* **46** (12), 91 (2022).
16. Muggeridge, D. J. et al. *M*Measurement of heart rate using the polar OH1 and fitbit charge 3 wearable devices in healthy adults during light, moderate, vigorous, and sprint-based exercise: validation study. JMIR mHealth uHealth **9**(3) (2021).
17. Fuller, D. C. E. et al. Reliability and validity of commercially available wearable devices for measuring steps, energy expenditure, and heart rate: systematic review. JMIR mHealth uHealth 8(9) (2020).
18. Zhang, Y. et al. Validity of Wrist-Worn photoplethysmography devices to measure heart rate: A systematic review and meta-analysis. *J. Sports Sci.* **38** (17), 2021–2034 (2020).
19. Støve, M. P. et al. Accuracy of the wearable activity tracker Garmin forerunner 235 for the assessment of heart rate during rest and activity. *J. Sports Sci.* **37** (8), 895–901 (2018).
20. Montalvo, S. et al. Commercial smart watches and heart rate monitors: A concurrent validity Analysis. *J. Strength. Cond Res.* **37** (9), 1802–1808 (2023).
21. Gilgen-Ammann, R., Schweizer, T. & Wyss, T. RR interval signal quality of a heart rate monitor and an ECG Holter at rest and during exercise. *Eur. J. Appl. Physiol.* **119** (7), 1525–1532 (2019).
22. Keir, D. A. et al. Identification of non-invasive exercise thresholds: methods, strategies, and an online App. *Sports Med.* **52** (2), 237–255 (2021).
23. Ozkaya, O. et al. Different categories of VO_2 kinetics in the 'extreme' exercise intensity domain. *J. Sports Sci.* **41** (23), 2144–2152 (2023).
24. Engelen, M. et al. Effects of hypoxic hypoxia on O_2 uptake and heart rate kinetics during heavy exercise. *J. Appl. Physiol. (1985)*. **51** (6), 2500–2508 (1996).
25. Navalta, J. W. et al. Heart rate processing algorithms and exercise duration on reliability and validity decisions in biceps-worn Polar Verity sense and OH1 wearables. *Sci. Rep.* **13**(1). (2023).
26. Tao, K. et al. Automated stress recognition using supervised learning classifiers by interactive virtual reality Scenes. *IEEE Trans. Neural Syst. Rehabil. Eng.* **30**, 2060–2066 (2022).
27. Luo, Y. et al. Decoding invariant Spatiotemporal synergy patterns on muscle networks of lower-limb movements via surface electromyographic signals. *Biomed. Signal Process. Control.* **91**, 106033 (2024).
28. Tu, R., Lu, Y. & Tao, K. Regular physical activities inhibit risk factors of the common cold among Chinese Adults. *Front. Psychol.* **13**, 864515 (2022).
29. Jo, E. et al. Validation of biofeedback wearables for photoplethysmographic heart rate Tracking. *J. Sports Sci. Med.* **15** (3), 540–547 (2016).
30. Khushhal, A. et al. Validity and reliability of the Apple watch for measuring heart rate during Exercise. *Sports Med. Int. Open.* **1** (6), E206–E211 (2017).
31. Boudreaux, B. D. et al. Validity of wearable activity monitors during cycling and resistance Exercise. *Med. Sci. Sports Exerc.* **50** (3), 624–633 (2018).
32. Wallen, M. P. et al. Accuracy of heart rate watches: implications for weight Management. *PLoS One*. **11** (5), e0154420 (2016).
33. Budidha, K. & Kyriacou, P. A. In vivo investigation of ear Canal pulse oximetry during hypothermia. *J. Clin. Monit. Comput.* **32** (1), 97–107 (2018).

Acknowledgements
We would also like to extend our appreciation to Kow and Wallace Wong for their assistance in addressing our inquiries regarding earbud technology.

Author contributions
J.Y., K.T., and J.Q. conceived the study concept and design. Data acquisition was performed by J.Y., Z.L., and X.L. Data were analyzed by J.Y. and Z.L., and results were interpreted by J.Y., Z.L., K.T., and J.Q. The study was supervised by K.T. and J.Q. The manuscript was drafted by J.Y. and K.T. All authors edited and critically reviewed the final manuscript.

Funding
This study received funding from the National Key Research and Development Program of China (2018YFC2000601) and Innovation Project of the General Administration of Sport of China (23KJCX049).

Declarations

Competing interests
The authors declare no competing interests.

Additional information
Correspondence and requests for materials should be addressed to K.T. or J.Q.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations. Please help me add this part!!!!!! Funding: This study received funding from the National Key Research and Development Program of China (2018YFC2000601) and Innovation Project of the General Administration of Sport of China (23KJCX049).