



OPEN ACCESS

EDITED BY

James Navalta,
University of Nevada, Las Vegas, United States

REVIEWED BY

Bruce Rogers,
University of Central Florida, United States
Alexander Montoyo,
Alma College, United States
Joel Reece,
Brigham Young University–Hawaii,
United States

*CORRESPONDENCE

Bingyu Pan,
✉ panbingyu@bsu.edu.cn
Yanfei Shen,
✉ syf@bsu.edu.cn

RECEIVED 02 May 2025

ACCEPTED 23 June 2025

PUBLISHED 04 July 2025

CORRECTED 16 July 2025

CITATION

Lu C, Cui W, Zhu Z, Wu Y, Xing Q, Pan B and Shen Y (2025) Validity of smartwatch-derived estimates of lactate threshold heart rate and pace compared to graded exercise testing. *Front. Physiol.* 16:1621996. doi: 10.3389/fphys.2025.1621996

COPYRIGHT

© 2025 Lu, Cui, Zhu, Wu, Xing, Pan and Shen. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](#). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Validity of smartwatch-derived estimates of lactate threshold heart rate and pace compared to graded exercise testing

Changda Lu¹, Wei Cui², Zheng Zhu², Yiwei Wu¹, Qingjun Xing¹, Bingyu Pan^{2,3*} and Yanfei Shen^{2,4*}

¹School of Sport Science, Beijing Sport University, Beijing, China, ²School of Sports Engineering, Beijing Sport University, Beijing, China, ³China Sports Big Data Center, Beijing Sport University, Beijing, China, ⁴Engineering Research Center of Strength and Conditioning Training Key Core Technology Integrated System and Equipment, Ministry of Education, Beijing, China

Background: Quantifying exercise intensity through the lactate threshold (LT) is crucial for optimizing athletic training regimes. Traditional methods like maximal lactate steady state and graded exercise testing are valid but invasive and costly. Advances in smartwatch technology offer a non-invasive alternative for monitoring LT, though their measurement protocols and outcomes have been less validated.

Methods: This study evaluates the validity of three mainstream smartwatches (Huawei GT Runner[®], Garmin Forerunner 265[®] or 265s[®], and Coros Pace3[®]) in estimating lactate threshold heart rate (LT HR) and pace (LT Pace), comparing these to measurement protocols and results from the modified Dmax method in laboratory standards. One hundred healthy recreational runners underwent indoor graded exercise tests followed by outdoor tests using Huawei (n = 100), Garmin (n = 23), and Coros (n = 17) smartwatches to compare differences in testing protocols and LT HR and LT Pace.

Results: The success rates for a single test were 78% for Huawei[®], 65.22% for Garmin[®], and 47.06% for Coros[®]. For LT HR, no significant differences were observed between smartwatch and DmaxMod estimates across all devices ($p > 0.05$). The Huawei[®] watch showed MAE = 10.66 bpm, MAPE = 6.32%; Garmin[®]: MAE = 11.44 bpm, MAPE = 7.15%; Coros[®]: MAE = 8.93 bpm, MAPE = 5.95%. Corresponding Pearson correlation coefficients ranged from $r = 0.13$ to 0.67, and R^2 values ranged from 0.02 to 0.45. In contrast, LT Pace predictions demonstrated significant overestimation for all devices. Huawei[®] reported the smallest error (MAE = 1.22 km/h, MAPE = 12.70%, $p = 0.01$, $r = 0.88$, $R^2 = 0.78$), followed by Garmin[®] (MAE = 2.17 km/h, MAPE = 25.78%, $p < 0.01$, $r = 0.73$, $R^2 = 0.53$), and Coros[®] (MAE = 1.93 km/h, MAPE = 22.63%, $p = 0.08$, $r = 0.79$, $R^2 = 0.62$). Bland–Altman plots confirmed systematic biases and variable agreement patterns, particularly for LT Pace.

Conclusion: Smartwatches are capable of providing estimates of LT HR and LT Pace in recreational runners, although they tend to overestimate LT Pace and overall accuracy remains to be improved.

KEYWORDS

exercise intensity, running, lactate threshold, smartwatch, validity

1 Introduction

Quantifying exercise intensity in endurance sports is essential for enhancing athletic performance (Mujika, 2017). As health awareness increases and activities like running gain popularity, more athletes seek precise methods to monitor and measure their training. Among various quantification techniques, the anaerobic lactate threshold (LT) (Svedahl and MacIntosh, 2003; Faude et al., 2009) is considered the key reference metric for assessing personalized training intensities (Svedahl and MacIntosh, 2003; Heuberger et al., 2018). The LT is characterized by a rapid and distinct change in the slope of the blood lactate concentration curve, indicating the transition from predominantly aerobic to anaerobic energy metabolism (Faude et al., 2009).

In practical training scenarios, lactate threshold heart rate (LT HR) and lactate threshold pace (LT Pace) are commonly used to quantify training intensity. LT HR refers to the heart rate (HR) at the LT (Pfitzinger and Freedson, 1998), while LT Pace refers to the sustainable running speed at this threshold (Yoshida et al., 1987). These indicators are crucial for developing effective and safe training plans, enabling athletes to maintain optimal performance over time, directly affecting their speed and endurance in competitions (Ziogas et al., 2011).

The primary methods for testing LT include the maximal lactate steady state (MLSS) and its more accessible variant, the graded exercise testing (GXT) (Faude et al., 2009; Jones and Ehrsam, 1982; Llodio et al., 2016; Jamnick et al., 2018; Cerezuela-Espejo et al., 2018), which are widely used to determine personalized thresholds for runners (Davis et al., 2007; Antonutto and Di Prampero, 1995; Faude et al., 2009). While these traditional methods are valid (Chalmers et al., 2024), they are invasive, costly, and rely on specialized equipment and trained personnel (Ramos-Campo et al., 2017; Nascimento et al., 2017; Shiraishi et al., 2018; Di Michele et al., 2012). With the advancement of wearable technology, smartwatches have emerged as integrated solutions incorporating various advanced wearable measurement technologies. These devices, equipped with sensors like GPS, PPG, and IMU, gather kinematic and physiological data, offering non-invasive means to predict a runner's LT. This technological progress significantly enhances the implementation of training plans (Siepmann and Kowalczyk, 2021; Paul et al., 2024), the monitoring of training sessions (Jubair and Mehenaz, 2024), and fatigue recovery (Jubair and Mehenaz, 2024). However, despite the convenience provided by smartwatches, the validity of their measurements has not been widely validated (Seshadri et al., 2019). Although numerous studies have explored the validity of smartwatches in measuring steps (Gaz et al., 2018; Sears et al., 2017; Veerabhadrapa et al., 2018), energy expenditure (Le et al., 2022; Lee et al., 2018; Nuss et al., 2019), HR (Fuller et al.,

2020; Gillinov et al., 2017; Støve et al., 2019; Boolani et al., 2019) and maximum oxygen uptake ($\text{VO}_{2\text{max}}$) (Carrier et al., 2023; Carrier et al., 2025; Caserman et al., 2024; Jamieson et al., 2024; Muthusamy et al., 2021; Cooper and Shafer, 2019; Kraft and Roberts, 2017), but research on their accuracy in estimating LTs is comparatively limited. Among the limited studies, Schlie observed that the Garmin Fenix 7[®] tends to underestimate LT HR and LT Pace (Schlie et al., 2024), a finding echoed by Carrier's research, which also reported an underestimation of LT performance by the Garmin Fenix 6[®] (Carrier et al., 2021). Additionally, few studies have investigated how the testing protocols of these watches might influence the results. Therefore, this study aims to compare the LT measurement protocols and their results from three mainstream smartwatches against LTs determined using the modified Dmax method (DmaxMod) in the laboratory setting (Siahkhouhian et al., 2012), assessing the validity of these wearable devices in terms of their measurement protocols and outcomes for LTs.

2 Materials and methods

2.1 Participants

This study recruited 100 healthy recreational runners aged 18–65 years as participants. All participants had no history of acute or chronic heart disease, hypertension, joint or musculoskeletal disorders, acute illnesses, pregnancy, or suspected pregnancy. They also had no lower limb injuries or surgeries in the past year. The research protocol adhered to the ethical principles of the Declaration of Helsinki and was approved by the Sports Science Experimental Ethics Committee of Beijing Sport University (2024136H). Before the experiment, all participants were fully informed about the procedures and provided their written informed consent. The consent form confirmed the participants' voluntary participation, understanding of the study protocol, and right to withdraw from the experiment at any time.

2.2 Experimental equipment

- Treadmill: Technogym Excite Run 700[®], Technogym.
- Blood Lactate Analyzer: Biosen C-line[®], EKF[®].
- Heart Rate Monitor Strap: POLAR H10[®], POLAR Electro Oy.
- Huawei Watch: HUAWEI GT Runner[®], version HarmonyOS 4.0, Huawei Terminal Co., Ltd.
- Garmin Watches: Garmin Forerunner 265[®] (used for male participants) and Garmin Forerunner 265s[®] (used for female

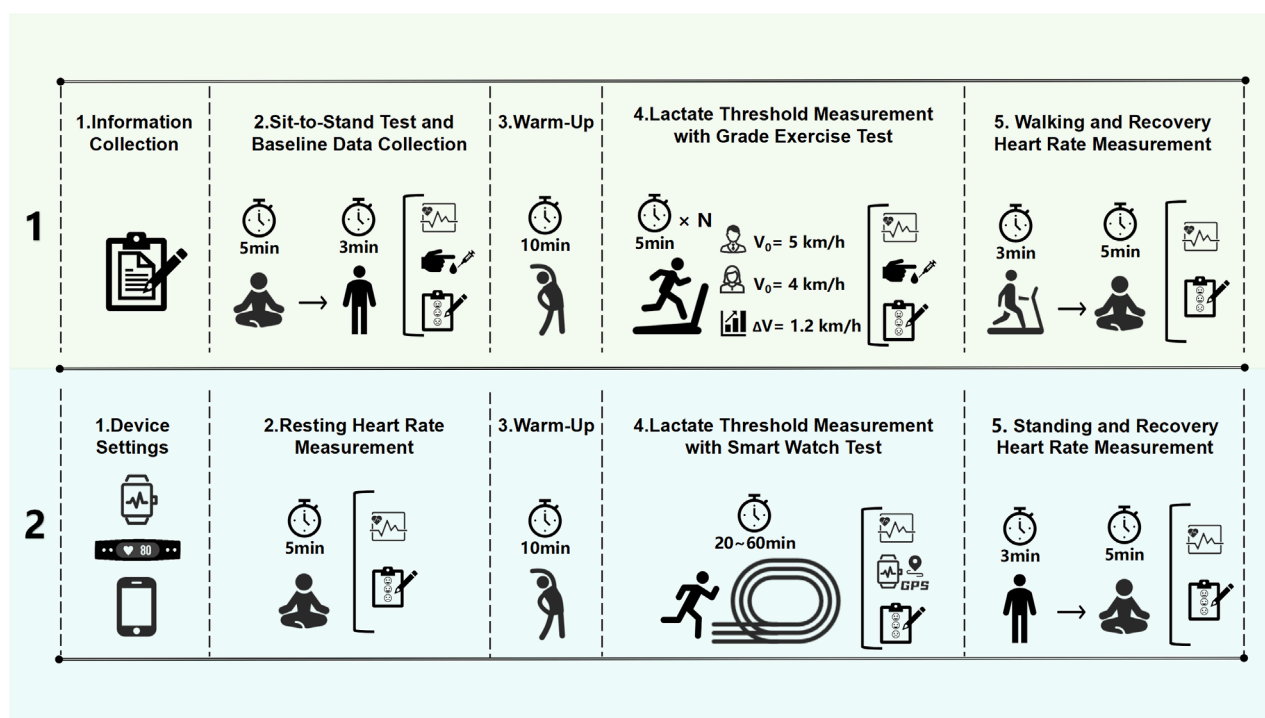


FIGURE 1
Experimental procedures for lactate threshold heart rate (LT HR) and lactate threshold pace (LT Pace) measurement protocols.

participants), version 19.18 Garmin International Inc. These two models are functionally identical, differing only in case size and strap length.

- Coros Watch: COROS Pace 3®, version v3.0309, Coros Sports Technology Co., Ltd.

2.3 Study design

The study was conducted in two phases: an indoor GXT and an outdoor test using smartwatches to assess LT levels. The testing process is depicted in Figure 1. Participants could engage in the outdoor test with one to three smartwatches. A minimum interval of 48 h was maintained between the indoor and outdoor tests and the various outdoor tests to ensure sufficient recovery and accuracy.

2.4 Indoor graded exercise test

Researchers collected baseline demographic data, including the participants' age, height, and weight. Before the exercise test commenced, participants completed a sit-to-stand test, which involved sitting quietly for 5 minutes and standing for 3 minutes; HR measured during this period was recorded as resting HR. Immediately after the sit-to-stand test, fingertip blood samples were taken to assess resting blood lactate levels. Participants then performed a 10-min dynamic warm-up routine by following a standardized instructional video provided by the researcher. Following the warm-up, the treadmill test began at an initial speed of

5 km/h for males and 4 km/h for females, with a 1% incline to offset the external wind resistance encountered outdoors (Davies, 1980). Each stage of the test lasted 5 minutes, followed by a 30-s pause for collecting fingertip blood samples and assessing fatigue using Borg's Rating of Perceived Exertion (RPE), which ranges from zero (no exhaustion) to ten (maximum exhaustion). The treadmill speed was then increased by 1.2 km/h for the next stage, continuing until the participant reached exhaustion or could not proceed further. Upon completion of the test, participants strolled for 3 minutes, followed by 5 minutes of seated rest, during which recovery HR was measured and RPE scores were reassessed.

2.5 Outdoor lactate threshold test with smartwatches

After completing the indoor graded exercise test, participants were informed of the outdoor testing procedures and duration for each smartwatch model. The selection of which smartwatch to test was made voluntarily by each participant based on personal preference and time availability. Participants were allowed to take part in tests for multiple smartwatch models, but a minimum interval of 48 h was required between each test. All outdoor tests were conducted on a standardized 400-m track under non-extreme weather conditions. Testing sessions were scheduled according to participant availability and were typically aligned with their usual training days. All indoor and outdoor tests were scheduled during morning or early afternoon hours to minimize potential circadian effects.

Researchers reset the data on each smartwatch and entered basic participant information such as gender, age, height, and weight into the respective software of each device. Participants were then fitted with the HR strap and the appropriate smartwatch. The smartwatch was placed on the dorsal side of the non-dominant wrist by trained researchers. To ensure consistent data collection conditions, all participants wore a Polar H10 chest strap during testing. For the Garmin® watch, the chest strap was a mandatory requirement to initiate its LT test. For the Huawei® and Coros® watches, the strap was not required for testing. It was connected via Bluetooth to a smartphone with GPS manually disabled, which was used solely for HR logging. The actual HR data for analysis was collected from the watches' built-in optical PPG sensors. Following this setup, they sat quietly for 5 min to measure their resting HR and assess their initial level of physical and psychological fatigue using Borg's RPE scale. Subsequently, participants completed a standardized 10-min dynamic warm-up routine via a pre-recorded instructional video. After the warm-up, they ran for 20–60 min, following the HR and pace instructions for each smartwatch's testing protocol. Throughout the run, all watches guided voice prompts and screen displays to maintain the stability of the HR and pace as much as possible. Upon completion of the run, researchers immediately recorded the participants' RPE scores.

2.5.1 Huawei® watch testing protocol

The Huawei watch testing required setting the participant's age and maximum heart rate (HR_{max} was determined using the peak HR obtained during the indoor test) before the test. The test included an 8-min warm-up phase, followed by gradual increases in HR every 3–5 min. The test was divided into six stages, lasting approximately 20–30 min. Participants were required to complete at least four stages.

2.5.2 Garmin® watch testing protocol

The Garmin® watch test required participants to perform an outdoor run to determine their VO₂max before initiating the LT test. Employing the Firstbeat® technology, the watch estimated the LT HR and LT Pace by analyzing changes in HR and respiratory rate (which affect heart rate variability). The test included a 10-min warm-up, then progressively increased the HR zones over six stages, each 3–4 min long, totaling approximately 25–30 min.

2.5.3 Coros® watch testing protocol

The Coros watch test required participants to input their 10-km run time. The test commenced with a 5-min warm-up, followed by 25 min of running at a marathon pace, then 3 min at a 10 km pace, and 3 min at a 5 km pace. If this point successfully captured the LT data, the test transitioned to a cool-down phase; if not, one to two additional sets at increased speeds were required. The test included up to 7 phases and lasted approximately 30–60 min. Participants were required to complete at least three running stages.

2.6 Determination of the lactate threshold

The LT of the indoor GXT was calculated using the DmaxMod method (Fabre et al., 2010). Blood lactate concentrations at each stage were interpolated using cubic spline interpolation to generate

a smooth and standardized curve for subsequent analysis. A third-order polynomial was then applied to fit the curve of blood lactate concentration as a function of running speed (Zwingmann et al., 2019). The LT was determined as the point on the blood lactate curve that maximizes the distance to the line connecting the first data point where the lactate concentration increases by more than 0.4 mmol/L to the last sample point. The heart rate and running speed at the LT are determined as the values corresponding to the point, as shown in Figure 2. The LT HR and LT Pace for each smartwatch were obtained from data downloaded from the cloud following the successful completion of the outdoor test.

2.7 Statistical analysis

Statistical analyses were conducted using Excel (Microsoft, 2021), SPSS version 22 (IBM Statistics, 2021), OriginPro (OriginLab, 2021), and Python version 3.11 to evaluate the predictive accuracy and measurement success rates of three smartwatches in estimating LT HR and LT Pace. Success rates were calculated for each device, and histograms were used to analyze the distribution of results across different age groups and LTs. Data normality was assessed by the Shapiro–Wilk test. For normally distributed variables, independent-samples t-tests were used to compare values between participants with successful and failed tests. The mean absolute error (MAE) and mean absolute percentage error (MAPE) between smartwatch estimates and DmaxMod results were calculated for each device, and Welch's t-tests were used to determine whether these differences were statistically significant. For variables that did not meet normality assumptions, the Mann-Whitney U test was applied as a nonparametric alternative. Line plots were employed to depict deviations and trends between the estimates provided by the devices and the measurements obtained using the DmaxMod method. Linear regression analyses were executed to plot estimates from the devices against measurements using the DmaxMod method. The coefficient of determination (R^2) was used to evaluate the model fit, and the Pearson correlation coefficient assessed the strength of the correlation. Bland-Altman plots were used to analyze the agreement and consistency between the estimates from the smartwatches and those obtained using the DmaxMod method, emphasizing the error range and the reliability of the estimation methods employed by the smartwatches. In addition, the two one-sided tests (TOST) procedure was employed to assess the statistical equivalence of LT HR and LT Pace values between smartwatch estimates and DmaxMod measurements. Equivalence bounds were defined using Cohen's $d = \pm 0.5$, with pooled standard deviations used to derive absolute margins.

3 Results

3.1 Descriptive statistics

This study measured the LTs of 100 participants through an indoor GXT. All participants took part in the outdoor running test using the Huawei® smartwatch. 23 participants engaged in the Garmin® smartwatch test, and 17 tested with the Coros smartwatch. Among all participants, 71 completed outdoor testing

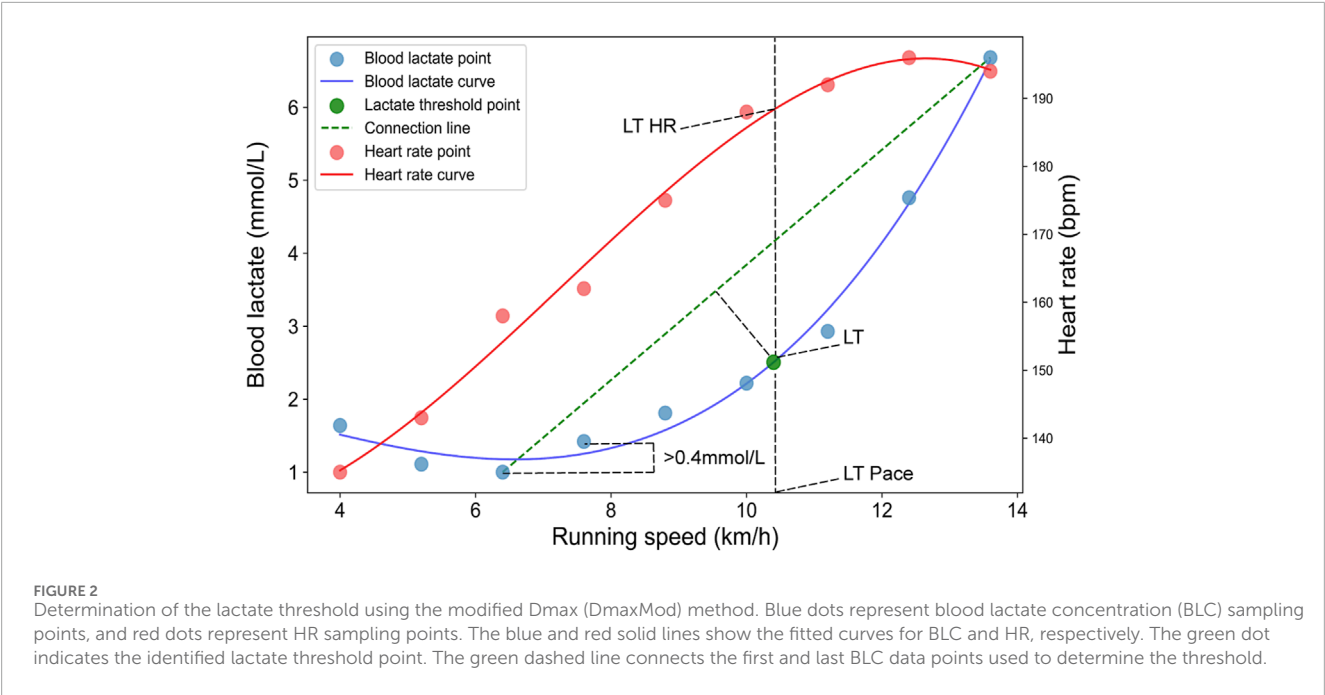


TABLE 1 Basic information of participants.

Watch	Gender ratio (male: Female)	Age (years)	Height (cm)	Weight (kg)
Huawei	61:39	28.0 ± 9.6	170.79 ± 8.34	64.30 ± 11.29
Garmin	13:10	29.13 ± 9.93	170.78 ± 9.58	64.96 ± 11.45
Coros	12:5	25.88 ± 7.11	173.35 ± 6.70	63.26 ± 10.46

with one smartwatch, 17 with two smartwatches, and 11 with all three smartwatches. The basic information of the participants is shown in [Table 1](#).

3.2 Measurement success rates

Since the LT in indoor tests was determined using the DmaxMod method, successful threshold identification was guaranteed as long as data acquisition was completed. In contrast, during the outdoor LT testing using smartwatches, there were instances in which participants completed all testing procedures, but the devices failed to generate LT HR or LT Pace values due to algorithmic limitations. Therefore, we analyzed the success rates of each smartwatch, as well as the distribution of successful and failed tests across different age groups, LT HR, and LT Pace. According to [Table 2](#), Huawei® watches exhibited a higher success rate of 78%, while Garmin® watches showed a moderate success rate of 65.22%. The Coros® watch recorded a lower success rate at only 47.06%. These varied outcomes are likely influenced by the different testing protocols and the complexities associated with each device.

[Table 3](#) shows the distribution of participants with successful and failed tests. Based on the results of Welch's t-test, we observed no significant differences between the successful and failed test groups

TABLE 2 Measurement success rate by smartwatches.

Indicator	Huawei	Garmin	Coros
Total	100	23	17
Successful	78	15	8
Failed	22	8	9
Success rate	78%	65.22%	47.06%

However, for the Coros® watch, statistically significant differences were found in both LT HR and LT Pace ($p < 0.01$), indicating that participants who failed the test tended to have lower LT HR and slower LT Pace compared to those with successful results.

[Figures 3A1–A3](#) show that the successful tests of the Huawei® watch primarily occurred in participants below 30 years of age, with an LT HR of approximately 170 bpm and an LT Pace of around 10–11 km/h. Failed tests were predominantly observed in participants with lower LT HR and LT Pace. [Figures 3B1–B3](#) displayed a moderate success rate across a wide range of ages and HR, particularly achieving higher success rates in participants

TABLE 3 Test outcomes by smartwatches.

Indicator	Group	Huawei	P	Garmin	p	Coros	p
Age (years)	Successful tests	28.47 ± 9.53	0.22	30.93 ± 10.38	0.12	23.88 ± 3.56	0.47
	Failed tests	27.05 ± 10.08		25.75 ± 8.61		27.67 ± 9.07	
LT HR (bpm)	Successful tests	165.12 ± 15.11	0.46	162.14 ± 14.80	0.67	167.45 ± 12.50	0.74
	Failed tests	163.33 ± 16.95		155.66 ± 23.78		159.10 ± 24.38	
LT Pace (km/h)	Successful tests	9.75 ± 2.28	0.31	8.93 ± 1.63	0.46	9.75 ± 2.12	0.36
	Failed tests	9.29 ± 2.86		9.53 ± 2.00		8.94 ± 1.81	

aged 20 to 35, with LT HR between 160–180 bpm and LT Pace of 9–11 km/h. However, there was a higher rate of test failures among the age groups under 20 and those with lower LT HR and LT Pace. As depicted in **Figures 3C1–C3**, the Coros® watch's successful tests were mainly concentrated among participants aged 18 to 30, with LT HR around 160–170 bpm and LT Pace primarily around 9–11 km/h. Despite these successes, the overall success rate was relatively low, mainly due to a higher failure rate at atypical LT HR and LT Pace, similar to the other smartwatches.

3.3 Predictive accuracy

In this study, we evaluated the prediction accuracy of the Huawei®, Garmin®, and Coros® smartwatches in terms of LT HR and LT Pace by comparing the results of these devices in an outdoor test with the results of an indoor GXT. Regarding LT HR prediction, as shown in **Table 4**, despite variations in the sample sizes tested by each watch, estimations from all three devices showed no significant differences ($p > 0.05$) with the results from the DmaxMod method and maintained MAE ranged from 8.93 to 11.44 bpm, and the MAPE ranged from 5.95% to 7.15%. Regarding LT Pace prediction, all devices tended to overestimate LT Pace compared to the DmaxMod method. The Huawei® watch showed the smallest error (MAE = 1.22 km/h, MAPE = 12.70%, $p = 0.01$), followed by the Garmin® watch (MAE = 2.17 km/h, MAPE = 25.78%, $p < 0.01$), and the Coros® watch (MAE = 1.93 km/h, MAPE = 22.63%, $p = 0.08$).

Figure 4 shows that most data points for both LT HR and LT Pace measured by the Huawei® Watch fall within the ± 1.96 standard deviation range from the DmaxMod method reference. The mean deviation is 2.86 bpm for LT HR and 0.98 km/h for LT Pace.

Figure 5 illustrates that, compared to the results of the DmaxMod method, the LT HR measurements from the Huawei® Watch are generally more concentrated. There is an underestimation at higher LT HR values and occasional overestimation in some subjects at lower LT HR levels. The distribution of LT Pace measurements is uniform, though it generally trends slightly above the DmaxMod method results, indicating a minor overestimation in Huawei® watch's LT Pace estimations.

Figure 6 shows that the prediction of LT HR showed an overestimation, with a mean deviation of 6.28 bpm and a

distribution of data points within $\pm 1.96SD$, but with a significant degree of dispersion, especially in the low HR interval, showing a considerable deviation. Regarding LT Pace, the mean deviation was 2.06 km/h, with a relatively dispersed distribution of data points, and most of them were located in the positive deviation range.

Figure 7 shows that the Gamin Watch exhibited underestimation in LT HR prediction, particularly in subjects with lower LT HR. Regarding LT Pace, the prediction results of the Gamin Watch were higher than the results of the DmaxMod method overall.

As shown in **Figure 8**, the mean deviation of LT HR was 5.44 bpm. In contrast, the mean deviation of LT Pace was 1.77 km/h, with the data points in the positive deviation range showing a tendency of overestimation.

As shown in **Figure 9**, the distribution of data points of LT HR of the Coros® watch is concentrated, showing a tendency of overestimation in individuals with lower LT HR and underestimation in those with higher LT HR. As for LT Pace, the prediction results of the Coros® watch are significantly higher than the results of the DmaxMod method, and the deviation of some data points is more significant, reflecting the systematic overestimation of pace prediction by the Coros® watch.

As shown in **Figure 10 A1 and A1**, the Huawei® Watch model for predicting LT HR has a coefficient of determination ($R^2 = 0.13$) and a Pearson correlation coefficient ($r = 0.36$). For LT pace, the coefficients are $R^2 = 0.78$ and $r = 0.88$. As shown in **Figure 10B1,B2**, the Garmin® Watch model shows a coefficient of determination ($R^2 = 0.45$) and a Pearson correlation coefficient ($r = 0.67$) for LT HR estimations. For LT Pace estimations, the coefficients are $R^2 = 0.53$ and $r = 0.73$. As shown in **Figure 12C1,C2**, the Coros® Watch model for predicting LT HR has $R^2 = 0.02$ and $r = 0.13$ coefficients. For LT Pace, the coefficients are $R^2 = 0.62$ and $r = 0.79$.

The TOST results indicated that none of the three devices met the equivalence criteria for LT HR. The mean differences were -2.44 bpm for Huawei®, -2.65 bpm for Coros®, and -6.52 bpm for Garmin®. Similarly, for LT Pace, the mean differences were -0.95 km/h for Huawei®, -1.53 km/h for Coros®, and -2.06 km/h for Garmin®. These results suggest that although average deviations were modest in some cases, none of the smartwatch-based estimates could be statistically considered equivalent to laboratory-derived reference values under the predefined thresholds.

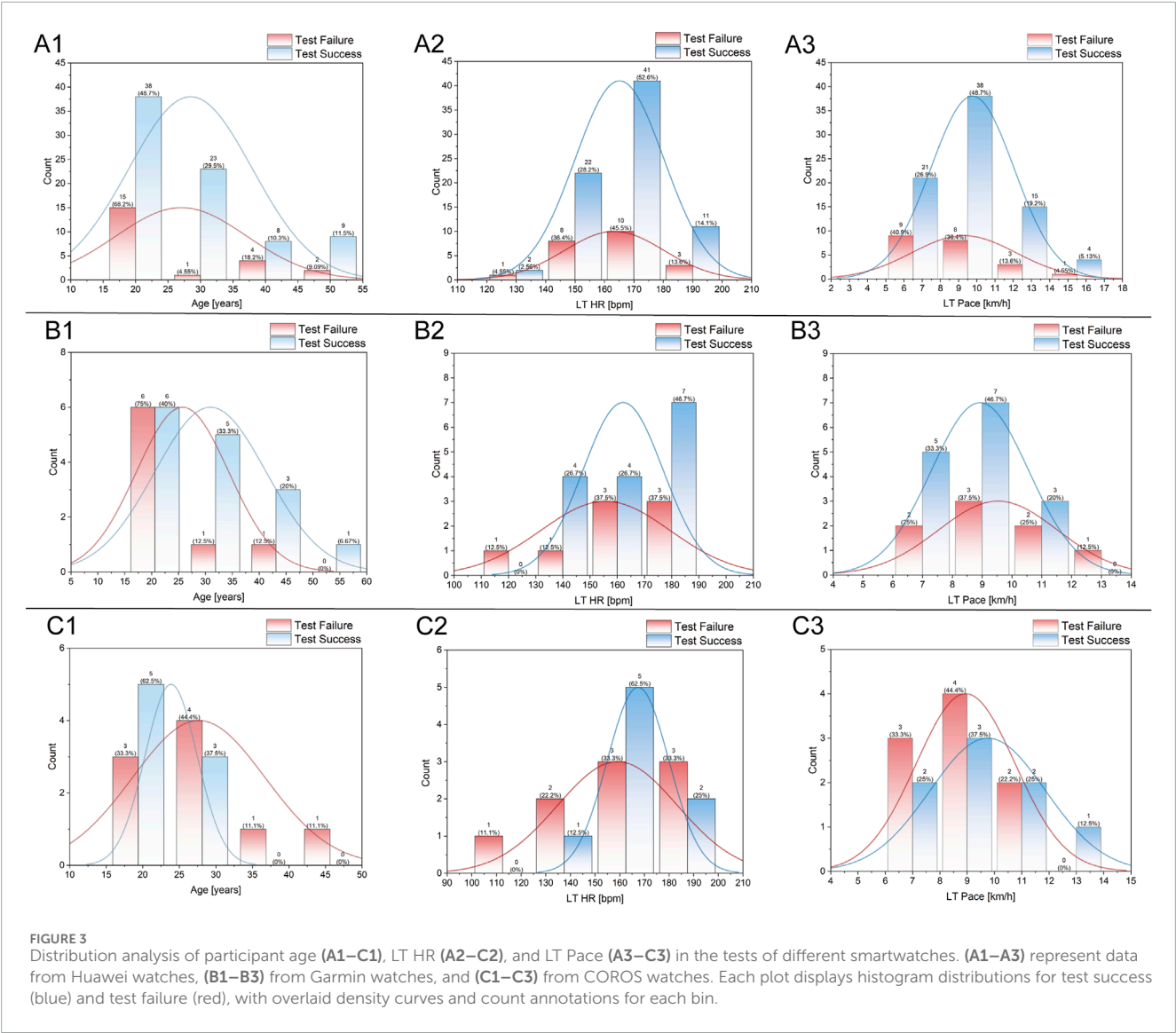


TABLE 4 Predictive accuracy of lactate threshold heart rate and lactate threshold pace.

Watch	Type	LT HR (bpm)	MAE (bpm)	MAPE (%)	p	LT pace (km/h)	MAE (km/h)	MAPE (%)	p
Huawei	DmaxMod	169.20 ± 15.12	10.13	6.63%	0.24	9.76 ± 2.28	1.22	12.70%	0.01
	Watch test	167.60 ± 9.93				10.71 ± 2.19			
Garmin	DmaxMod	162.14 ± 14.81	8.93	5.95%	0.15	8.93 ± 1.63	2.17	25.78%	0.00
	Watch test	168.67 ± 8.25				10.99 ± 1.60			
Coros	DmaxMod	167.46 ± 12.50	11.44	7.15%	0.13	9.75 ± 2.12	1.93	22.63%	0.08
	Watch test	175.38 ± 5.18				11.69 ± 2.01			

4 Discussion

4.1 Test protocols of each smartwatch

The basic principle of threshold tests is to gradually increase the external load on the runner, which causes a corresponding change

in the internal load, and then to monitor the threshold at which significant changes in the runner’s physiological and biochemical parameters occur (Mader and Heck, 1986; Poole et al., 2021). Depending on the indicators used, these tests can be categorized into ventilatory threshold (Hollmann, 2001), LT (Poole et al., 2021; Faude et al., 2009), and heart rate variability threshold

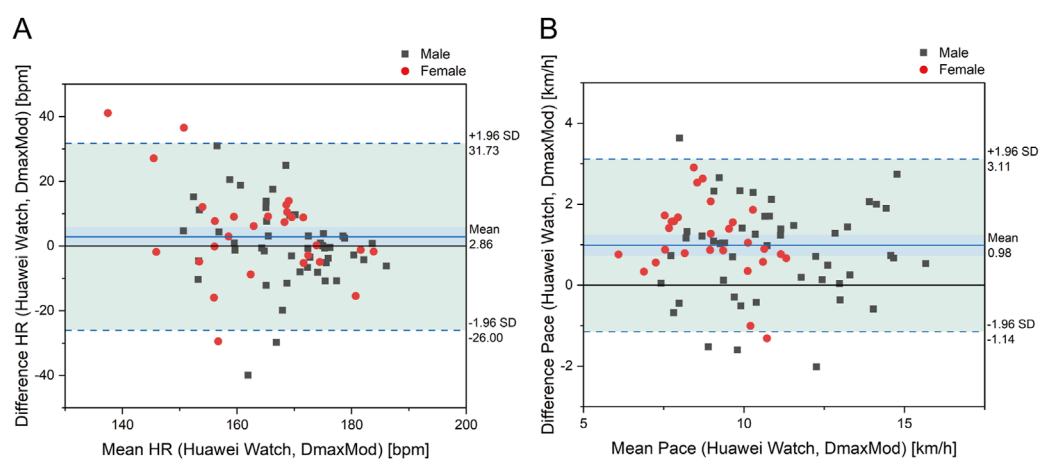


FIGURE 4
Bland-Altman plot for differences in (A) LT HR and (B) LT pace between Huawei watch and DmaxMod method.

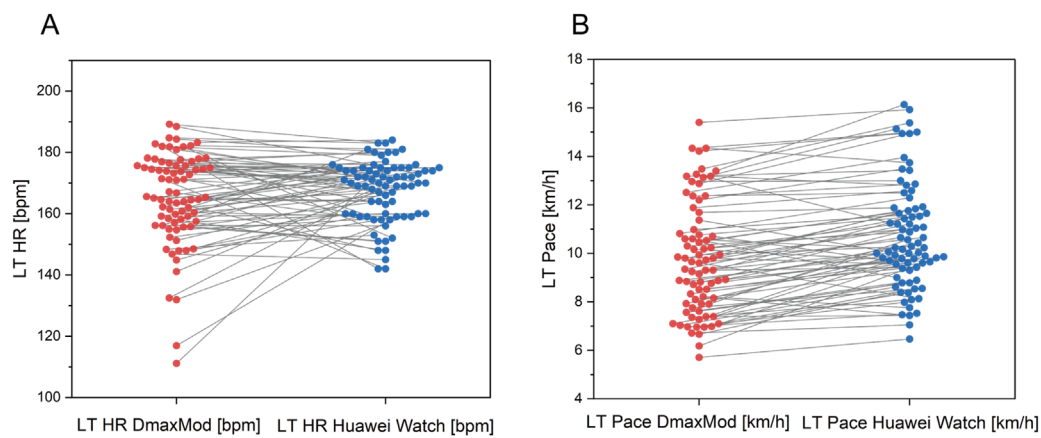


FIGURE 5
Comparison of (A) LT HR and (B) LT pace between Huawei watch and DmaxMod method.

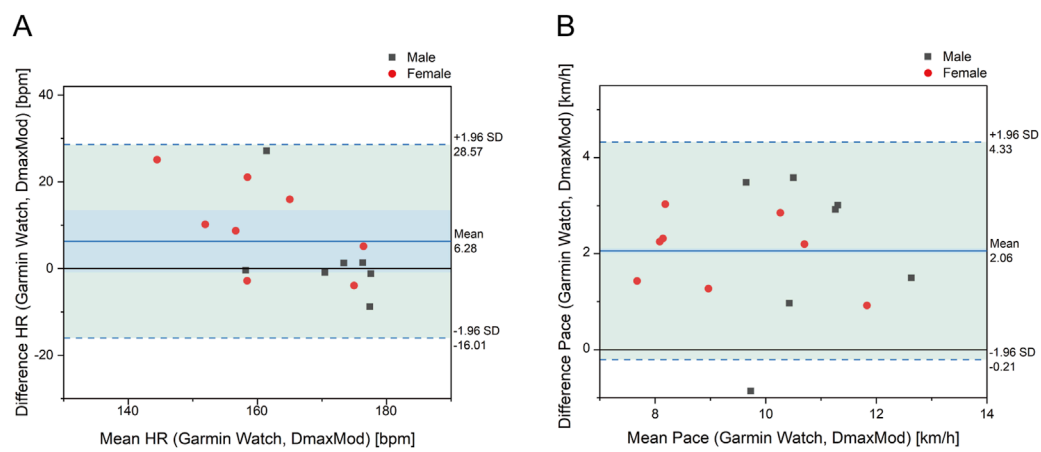


FIGURE 6
Bland-Altman plot for differences in (A) LT HR and LT (B) pace between Garmin watch and DmaxMod method.

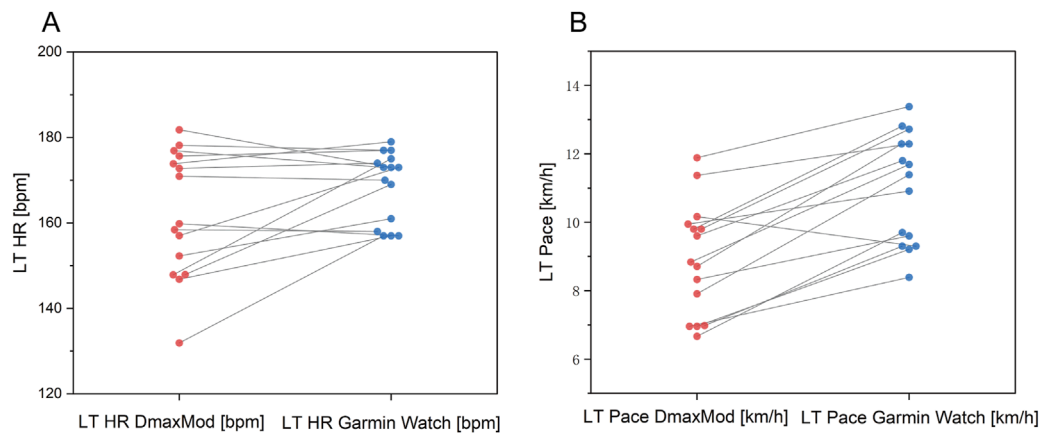


FIGURE 7
Comparison of (A) LT HR and (B) LT pace between Garmin watch and DmaxMod method.

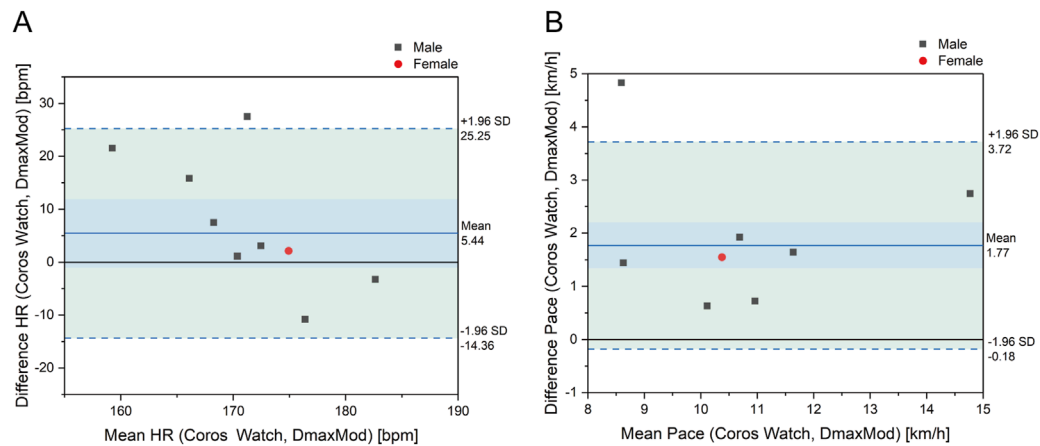


FIGURE 8
Bland-Altman plot for differences in (A) LT HR and (B) LT pace between Coros watch and DmaxMod method.

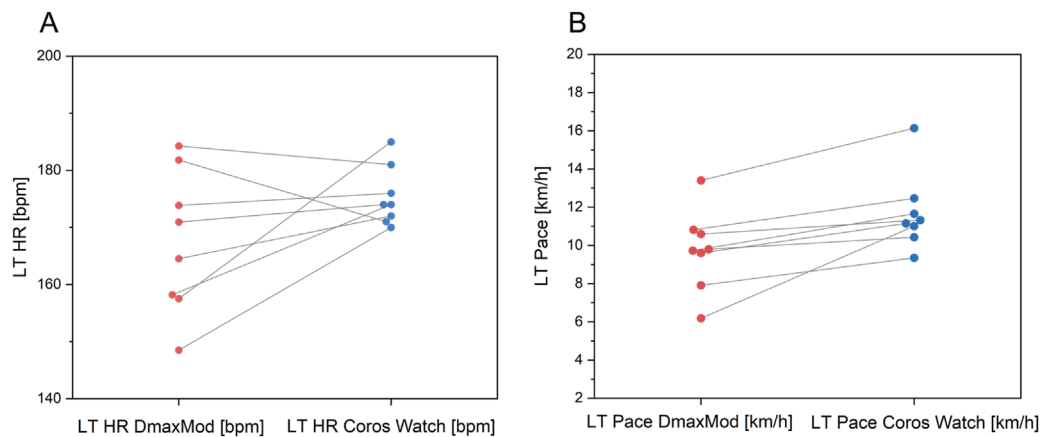


FIGURE 9
Comparison of LT (A) HR and (B) LT pace between Coros watch and DmaxMod method.

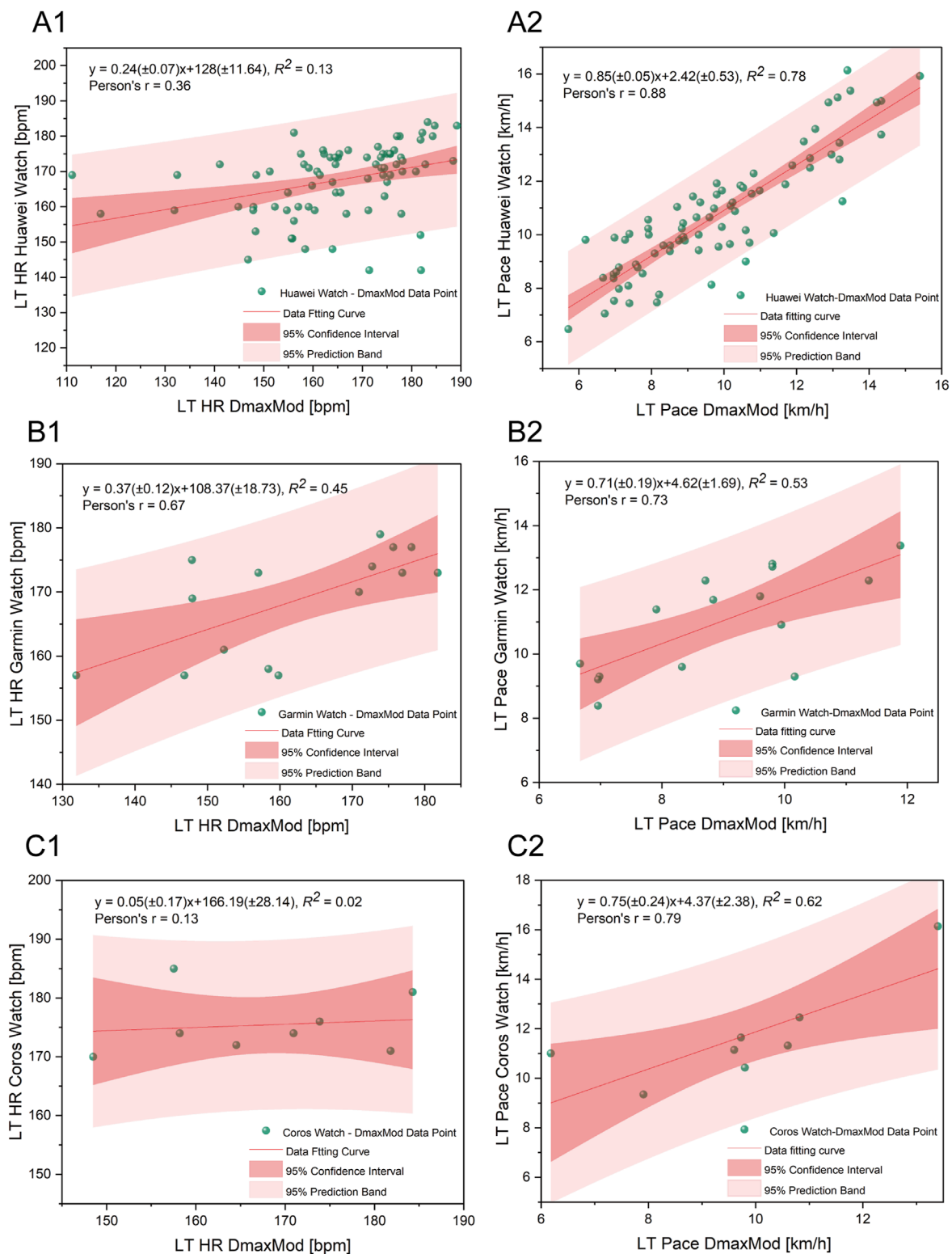


FIGURE 10

Regression analysis comparing LT HR (A1–C1) and LT Pace (A2–C2) measured by smartwatches with those calculated using the DmaxMod method. (A1,A2) represent data from Huawei watches, (B1,B2) from Garmin watches, and (C1,C2) from COROS watches. Each plot includes the fitted regression line, 95% confidence interval (dark red band), and 95% prediction band (light red band), with green dots representing individual data points.

(Cottin et al., 2006; Karapetian et al., 2012; Queiroz et al., 2018). The differences in testing methods primarily lie in three areas: 1) the manner in which external load is imposed, 2) the amount and granularity of these loads, and 3) the algorithm used to detect changes in internal load indicators. The LT measurement technologies used in smartwatches varies, but the main differences also fall within the abovementioned aspects. Because manufacturers have not disclosed detailed information about their detection algorithms, this discussion will instead concentrate on the methods of imposing external loads and the amount and granularity of these loads.

4.2 Methods of imposing external load

When using smartwatches for LT measurement, hardware capabilities and software interactions limit the methods of imposing external load (Seshadri et al., 2019). Although external load in running tests typically appears as changes in speed, smartwatches cannot directly control a runner's speed but guide adjustments through prompts. Different prompt methods can directly affect the validity of the test. For instance, the Coros® watch test instructs participants to increase their pace. This approach is straightforward and similar to an indoor GXT (Faude et al., 2009; Jones and Ehrsam, 1982; Llodio et al., 2016; Jamnick et al., 2018; Cerezuela-Espejo et al., 2018). However, the issue with this method is that, although the pace is stable during a test stage, achieving a steady state in HR within about 3 min is challenging (Bailey and Ratcliffe, 1995; Zakyntinaki, 2015; Marini et al., 2022). This increases the difficulty of subsequent detection algorithms. Indeed, due to these reasons, the Coros® watch's test failure rate is comparatively higher. On the other hand, Huawei® and Garmin® watches employ a different strategy by asking runners to maintain a specific HR range. This subtler method does not require runners to keep a certain pace but enables them to manage their HR by adjusting their running pose. Since the heart muscle is involuntary, runners cannot control their HR directly. It allows runners to maintain it by focusing on their running posture and adjusting their breathing and other physiological processes (Meir et al., 2014; Gu et al., 2017). Therefore, using HR as the control variable may be a more optimal choice, as it can better maintain physiological stability throughout each phase and reduce the complexity of the detection algorithm (Hoppe et al., 2018; Wallace et al., 2014).

4.3 Amount and granularity of external load

The amount and granularity of the external load applied in LT tests are influenced by the input indicators provided before the test. For Huawei® watches, the input data can be either the runner's HRmax or their age. If age is input, the HRmax is calculated using the "220-age." Garmin® watches use the VO₂max predicted during outdoor running tests to set parameters. For Coros® watches, the input is based on the runner's 10-km running performance. Each of these inputs has its limitations. For inputting HRmax, without laboratory measurements, it is difficult for runners to determine

their actual HRmax accurately (Abt et al., 2018), and reliance on age for estimation is inaccurate (Shookster et al., 2020; Robergs and Landwehr, 2002). The problem with inputting VO₂max is similar to that of HRmax. The Garmin® watches do not support manual input, but are measured only through outdoor running tests. Research has shown a discrepancy between the maximum oxygen uptake measured by watches and that measured in laboratories (Molina-Garcia et al., 2022; Kraft and Roberts, 2017; Snyder et al., 2021). This discrepancy can significantly impact the accuracy of the test. When the input indicator is a 10-km performance, runners often enter their best past 10-km time, which could lead to higher pacing in various test phases, potentially increasing the difficulty of the test. Additionally, when considering the input indicators themselves, the stability of a 10-km performance is significantly lower than that of physiological indicators such as HRmax (Düking et al., 2022) and VO₂max (Molina-Garcia et al., 2022). This can affect the reliability and accuracy of the tests.

4.4 Validity of the measurements

In this study, the three smartwatches demonstrated high validity in estimating LT HR, with no significant differences from the results obtained using the DmaxMod method and MAPE ranging between 5.95% and 7.15%. These findings are consistent with those of Carrier et al., who reported a 6.20% error in HR measurements for the Garmin Fenix 6® (Carrier et al., 2021). However, they are slightly higher than the 1.71% error reported by Schlie's study using the Garmin Fenix 7® during outdoor tests (Schlie et al., 2024). Although the MAPE was low, the MAE of 8.93–11.44 bpm may still affect the precision of training zone determination. Therefore, runners are advised to interpret smartwatch-estimated LT HR in conjunction with perceived exertion to better guide endurance training.

Moreover, the data points from the three watches were quite concentrated, tending to overestimate the values for individuals with lower LT HR and underestimate those with higher LT HR, unlike the DmaxMod method, which can more accurately differentiate HR in marginal zones. This tendency toward that smartwatch algorithms tend to perform best when predicting values near the population average, but may produce less accurate estimations for individuals at the lower or upper extremes of fitness. Such a limitation could reduce the utility of these devices for high-level athletes or individuals with unusually low or high thresholds.

In estimating LT Pace, all three watches demonstrated significant deviations from the DmaxMod results, with MAPE ranging from 12.70% to 25.63%. This finding contrasts with Schlie's results using the Garmin Fenix 7 in outdoor tests, where the watch slightly underestimated LT Pace (Schlie et al., 2024). This discrepancy may stem from the differences in testing environments. Schlie's measurements were taken on an outdoor track, while this study was conducted on a treadmill within a laboratory setting. Previous evidence suggests that treadmill running tends to elicit lower blood lactate concentrations at submaximal speeds. And a subtle mismatch between actual pacing and perceived exertion has been reported, indicating that treadmill-based pacing may feel

different to participants despite similar physiological responses (Miller et al., 2019).

Additionally, the study found that although the three smartwatches had a high accuracy in predicting LT HR, their regression analysis performance was poor ($R^2 = 0.02\text{--}0.45$, $r = 0.13\text{--}0.67$) and notably wide limits of agreement. This phenomenon suggests that the high accuracy might not all stem from the validity of the predictive algorithms in the watches but rather from the concentration of different LT HR data distributions, potentially overlooking individual differences among runners. Moreover, this could be influenced by smartwatch manufacturers' strategies, as showing higher LT HR may push training intensities beyond what runners can comfortably sustain. The conservative approach likely aims to ensure safety. For LT Pace, the watches showed a moderate correlation in regression analysis ($r = 0.73\text{--}0.88$), but the actual prediction accuracy was not high. Therefore, although the smartwatches demonstrated acceptable average accuracy in estimating LT HR, the wide individual variability and limits of agreement suggest that their outputs may still lead to misclassification of physiological training zones, particularly near threshold boundaries (Iannetta et al., 2018). For LT Pace, a systematic overestimation was observed, which further contributes to the risk of distorted intensity guidance. For runners who require precise intensity control, such inaccuracies could result in excessive training loads or insufficient recovery, potentially impairing long-term performance and increasing the risk of injury or overtraining. While smartwatches offer convenience and the ability to monitor general training trends, their outputs should be interpreted with caution when applied to individualized training programs. Given the potential for zone boundary distortion and physiological variability (Iannetta et al., 2023), relying solely on device-generated thresholds may result in inappropriate intensity prescription. Athletes and recreational runners alike are advised to complement wearable data with subjective indicators, ensuring a more holistic and adaptive approach to training regulation.

4.5 Limitation

The present study encounters some challenges with variability in data samples among different watches. First, the data for the Coros® watch, although providing valuable insights, is relatively minor compared to those for other watches, which might influence the extrapolation of the findings. Second, a portion of participants experienced smartwatch test failures, in which the device failed to generate LT HR or LT Pace values despite completing the test. This may have influenced the accuracy and completeness of the results. Third, the study measures runners' LTs using an indoor incremental load test, whereas the watch tests were conducted outdoors. Environmental variables could have impacted the results. Additionally, short-term physiological fluctuations in participants may have contributed to day-to-day variability in LT. Fourth, the devices differed in the methods of HR acquisition during LT testing. Huawei® and Coros® relied on optical PPG sensors, while Garmin® used an ECG-based chest strap. LT estimation depends on accurate HR and speed dynamics. PPG sensors are more susceptible to motion artifacts, which

may have affected detection performance during running. This methodological variation could contribute to differences in device accuracy. Future studies should standardize testing protocols, harmonize heart rate monitoring methods, control for internal and external sources of variability, and ensure balanced sample sizes for each device.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving humans were approved by Sports Science Experimental Ethics Committee of Beijing Sport University. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

CL: Data curation, Writing – review and editing, Conceptualization, Methodology, Visualization, Formal Analysis, Validation, Writing – original draft. WC: Software, Investigation, Validation, Writing – review and editing, Data curation. ZZ: Writing – review and editing, Validation, Data curation, Investigation. YW: Data curation, Writing – review and editing, Investigation. QX: Investigation, Data curation, Writing – review and editing. BP: Writing – review and editing, Methodology, Funding acquisition, Supervision, Conceptualization. YS: Writing – review and editing, Supervision, Conceptualization, Funding acquisition, Methodology.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. This study has been supported by the National Natural Science Foundation of China Grant (72071018), the Fundamental Research Funds for the Central Universities of China, (2023GCZX003) (Research on the Nonlinear Accurate Measurement System of Exercise Loads), and the Fundamental Research Funds for the Central Universities, (2025025), (Research and Application of a Multimodal Data-Based Estimation Model for Running Lactate Threshold Intensity).

Acknowledgments

We extend our sincere gratitude to all participants from various communities who contributed to this study. We also wish to express our heartfelt thanks to Hang Liu, Litian Zhou, and Yufeng Fang of Huawei Device Co., Ltd. for their constructive feedback and

their generous support in supplying the equipment essential for the completion of this research.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Correction note

A correction has been made to this article. Details can be foundat: [10.3389/fphys.2025.1661421](https://doi.org/10.3389/fphys.2025.1661421).

References

- Abt, G., Bray, J., and Benson, A. C. (2018). The validity and inter-device variability of the apple watch™ for measuring maximal heart rate. *J. sports Sci.* 36, 1447–1452. doi:10.1080/02640414.2017.1397282
- Antonutto, G., and Di Prampero, P. E. (1995). The concept of lactate threshold. A short review. *J. Sports Med. Phys. Fit.* 35, 6–12.
- Bailey, M. J., and Ratcliffe, C. M. (1995). Reliability of physiological cost index measurements in walking normal subjects using steady-state, non-steady-state and post-exercise heart rate recording. *Physiotherapy* 81, 618–623. doi:10.1016/S0031-9406(05)66648-5
- Boolani, A., Towler, C., Lecours, B., Blank, H., Larue, J., and Fulk, G. (2019). Accuracy of 6 commercially available activity monitors in measuring heart rate, caloric expenditure, steps walked, and distance traveled. *Cardiopulm. Phys. Ther. J.* 30, 153–161. doi:10.1097/cpt.0000000000000107
- Carrier, B., Cruz, K., Farmer, H., and Navalta, J. (2021). Validation of the lactate threshold estimate from the garmin fenix 6 fitness tracker. *Med. Sci. Sports Exerc.* 53, 48. doi:10.1249/01.mss.0000759640.87686.98
- Carrier, B., Helm, M. M., Cruz, K., Barrios, B., and Navalta, J. W. (2023). Validation of aerobic capacity (VO2max) and lactate threshold in wearable technology for athletic populations. *Technologies* 11, 71. doi:10.3390/technologies11030071
- Carrier, B., Marten Chaves, S., and Navalta, J. W. (2025). Validation of aerobic capacity (VO2max) and pulse oximetry in wearable technology. *Sensors* 25, 275. doi:10.3390/s25010275
- Caserman, P., Yum, S., Göbel, S., Reif, A., and Matura, S. (2024). Assessing the accuracy of smartwatch-based estimation of maximum oxygen uptake using the apple watch series 7: validation study. *JMIR Biomed. Eng.* 9, e59459. doi:10.2196/59459
- Cerezuela-Espejo, V., Courel-Ibáñez, J., Morán-Navarro, R., MartíNEZ-Cava, A., and Pallarés, J. G. (2018). The relationship between lactate and ventilatory thresholds in runners: validity and reliability of exercise test performance parameters. *Front. physiology* 9, 1320. doi:10.3389/fphys.2018.01320
- Chalmers, S., Esterman, A., Chalmers, S., Esterman, A., Estona, R., Norton, K., et al. (2024). Standardization of the dmax method for calculating the second lactate threshold. *Int. J. Sports Physiology Perform.* 10, 921–926. doi:10.1123/ijspp.2014-0537
- Cooper, K. D., and Shafer, A. B. (2019). Validity and reliability of the polar A300's fitness test feature to predict VO2max. *Int. J. Exerc. Sci.* 12, 393–401. doi:10.70252/BDAY8376
- Cottin, F., Leprêtre, P.-M., Lopes, P., Papelier, Y., Médigue, C., and Billat, V. (2006). Assessment of ventilatory thresholds from heart rate variability in well-trained subjects during cycling. *Int. J. Sports Med.* 27, 959–967. doi:10.1055/s-2006-923849
- Davies, C. (1980). Effects of wind assistance and resistance on the forward motion of a runner. *J. Appl. Physiology* 48, 702–709. doi:10.1152/jappl.1980.48.4.702
- Davis, J. A., Rozenek, R., Decicco, D. M., Carizzi, M. T., and Pham, P. H. (2007). Comparison of three methods for detection of the lactate threshold. *Clin. physiology Funct. imaging* 27, 381–384. doi:10.1111/j.1475-097X.2007.00762.x
- Di Michele, R., Gatta, G., Di Leo, A., Cortesi, M., Andina, F., Tam, E., et al. (2012). Estimation of the anaerobic threshold from heart rate variability in an incremental swimming test. *J. Strength and Cond. Res.* 26, 3059–3066. doi:10.1519/JSC.0b013e318245bde1
- Düking, P., Van Hooren, B., and Sperlich, B. (2022). Assessment of peak oxygen uptake with a smartwatch and its usefulness for training of runners. *Int. J. Sports Med.* 43, 642–647. doi:10.1055/a-1686-9068
- Fabre, N., Balestreri, F., Pellegrini, B., and Schena, F. (2010). The modified dmax method is reliable to predict the second ventilatory threshold in elite cross-country skiers. *J. Strength and Cond. Res.* 24, 1546–1552. doi:10.1519/JSC.0b013e3181dc450a
- Faude, O., Kindermann, W., and Meyer, T. (2009). Lactate threshold concepts: how valid are they? *Sports Med.* 39, 469–490. doi:10.2165/00007256-200939060-00003
- Fuller, D., Colwell, E., Low, J., Orychock, K., Tobin, M. A., Simango, B., et al. (2020). Reliability and validity of commercially available wearable devices for measuring steps, energy expenditure, and heart rate: systematic review. *JMIR mHealth uHealth* 8, e18694. doi:10.2196/18694
- Gaz, D. V., Rieck, T. M., Peterson, N. W., Ferguson, J. A., Schroeder, D. R., Dunfee, H. A., et al. (2018). Determining the validity and accuracy of multiple activity-tracking devices in controlled and free-walking conditions. *Am. J. Health Promot.* 32, 1671–1678. doi:10.1177/0890117118763273
- Gillinov, S., Etiwy, M., Wang, R., Blackburn, G., Phelan, D., Gillinov, A. M., et al. (2017). Variable accuracy of wearable heart rate monitors during aerobic exercise. *Med. and Sci. Sports and Exerc.* 49, 1697–1703. doi:10.1249/MSS.0000000000001284
- Gu, F., Niu, J., Das, S. K., He, Z., and Jin, X. (2017). Detecting breathing frequency and maintaining a proper running rhythm. *Pervasive Mob. Comput.* 42, 498–512. doi:10.1016/j.pmcj.2017.06.015
- Heuberger, J. A., Gal, P., Stuurman, F. E., De Muinck Keizer, W. A., Mejia Miranda, Y., and Cohen, A. F. (2018). Repeatability and predictive value of lactate threshold concepts in endurance sports. *PloS one* 13, e0206846. doi:10.1371/journal.pone.0206846
- Hollmann, W. (2001). 42 years Ago—Development of the concepts of ventilatory and lactate threshold. *Sports Med.* 31, 315–320. doi:10.2165/00007256-200131050-00002
- Hoppe, M. W., Baumgart, C., and Freiwald, J. (2018). Estimating external loads and internal demands by positioning systems and innovative data processing approaches during intermittent running activities in team and racquet sports. *Sports Orthop. Traumatology* 34, 3–14. doi:10.1016/j.orthtr.2018.01.001
- Iannetta, D., Inglis, E. C., Fullerton, C., Passfield, L., and Murias, J. M. (2018). Metabolic and performance-related consequences of exercising at and slightly above MLSS. *Scand. J. Med. and Sci. sports* 28, 2481–2493. doi:10.1111/sms.13280
- Iannetta, D., Marinari, G., and Murias, J. M. (2023). Can the heart rate response at the respiratory compensation point be used to retrieve the maximal metabolic steady state? *J. Sports Sci.* 41, 1025–1032. doi:10.1080/02640414.2023.2259206
- Jamieson, A., Jones, S., Steves, C., Timpson, N., Chaturvedi, N., Hughes, A. D., et al. (2024). Comparison between smartwatch-derived and CPET-measured VO2max. *Age (years)* 215, 32–62. doi:10.22489/cinc.2024.352
- Jamnick, N. A., Botella, J., Pyne, D. B., and Bishop, D. J. (2018). Manipulating graded exercise test variables affects the validity of the lactate threshold and [formula: see text]. *PloS one* 13, e0199794. doi:10.1371/journal.pone.0199794
- Jones, N. L., and Ehsam, R. E. (1982). The anaerobic threshold. *Exerc. sport Sci. Rev.* 10, 49–83. doi:10.1249/00003677-198201000-00003
- Jubair, H., and Mehenaz, M. (2024). Smartwatch-assisted exercise prescription: utilizing machine learning algorithms for personalized workout recommendations and monitoring: a review. doi:10.21203/rs.3.rs-4871091/v1
- Karapetian, G., Engels, H., Gretebeck, K., and Gretebeck, R. (2012). Effect of caffeine on LT, VT and HRVT. *Int. J. sports Med.* 33, 507–513. doi:10.1055/s-0032-1301904

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

- Kraft, G. L., and Roberts, R. A. (2017). Validation of the garmin forerunner 920XT fitness watch VO2peak test. *Int. J. Innov. Educ. Res.* 5, 63–69. doi:10.31686/ijer.vol5.iss2.619
- Lee, M., Lee, H., and Park, S. (2018). Accuracy of swimming wearable watches for estimating energy expenditure. *Int. J. Appl. Sports Sci.* 30, 80–90. doi:10.24985/ijass.2018.30.1.80
- Le, S., Wang, X., Zhang, T., Lei, S. M., Cheng, S., Yao, W., et al. (2022). Validity of three smartwatches in estimating energy expenditure during outdoor walking and running. *Front. Physiology* 13, 995575. doi:10.3389/fphys.2022.995575
- Llodio, I., Gorostiaga, E., Garcia-Tabar, I., Granados, C., and Sánchez-Medina, L. (2016). Estimation of the maximal lactate steady state in endurance runners. *Int. J. Sports Med.* 37, 539–546. doi:10.1055/s-0042-102653
- Mader, A., and Heck, H. (1986). A theory of the metabolic origin of “anaerobic threshold”. *Int. J. Sports Med.* 7, S45–S65. doi:10.1055/s-2008-1025802
- Marini, C. F., Federici, A., Skinner, J. S., Piccoli, G., Stocchi, V., Zoffoli, L., et al. (2022). Effect of steady-state aerobic exercise intensity and duration on the relationship between reserves of heart rate and oxygen uptake. *PeerJ* 10, e13190. doi:10.7717/peerj.13190
- Meir, R., Zhao, G.-G., Zhou, S., Beavers, R., and Davie, A. (2014). The acute effect of mouth only breathing on time to completion, heart rate, rate of perceived exertion, blood lactate, and ventilatory measures during a high-intensity shuttle run sequence. *J. Strength and Cond. Res.* 28, 950–957. doi:10.1519/JSC.0000000000000246
- Miller, J. R., Van Hooren, B., Bishop, C., Buckley, J. D., Willy, R. W., and Fuller, J. T. (2019). A systematic review and meta-analysis of crossover studies comparing physiological, perceptual and performance measures between treadmill and overground running. *Sports Med.* 49, 763–782. doi:10.1007/s40279-019-01087-9
- Molina-Garcia, P., Notbohm, H. L., Schumann, M., Argent, R., Hetherington-Rauth, M., Stang, J., et al. (2022). Validity of estimating the maximal oxygen consumption by consumer wearables: a systematic review with meta-analysis and expert statement of the INTERLIVE network. *Sports Med.* 52, 1577–1597. doi:10.1007/s40279-021-01639-y
- Mujika, I. (2017). Quantification of training and competition loads in endurance sports: methods and applications. *Int. J. Sports Physiology Perform.* 12, S29–S217–17. doi:10.1123/ijsspp.2016-0403
- Muthusamy, S., Subramaniam, A., Balasubramanian, K., Purushothaman, V. K., and Vasanthi, R. K. (2021). Assessment of Vo2 max reliability with garmin smart watch among Swimmers.(2021). *Int. J. Life Sci. Pharma Res.* 11, 42–46. doi:10.22376/ijpbs/lpr.2021.11.4.L42-46
- Nascimento, E. M. F., Kiss, M. A. P. D. M., Santos, T. M., Lambert, M., and Pires, F. O. (2017). Determination of lactate thresholds in maximal running test by heart rate variability data set. *Asian J. Sports Med.* 8, doi:10.5812/asjms.58480
- Nuss, K. J., Thomson, E. A., Courtney, J. B., Comstock, A., Reinwald, S., Blake, S., et al. (2019). Assessment of accuracy of overall energy expenditure measurements for the fitbit charge HR 2 and apple watch. *Am. J. Health Behav.* 43, 498–505. doi:10.5993/AJHB.43.3.5
- Paul, D. D. A., Pusuluri, H., Dasari, R., Abhishek, S., and Anjali, T. (2024). “Advanced smartwatch data analysis and predictive modeling for health and fitness optimization,” in *2024 5th international conference on intelligent communication technologies and virtual Mobile networks (ICICV)*. IEEE, 226–233. doi:10.1109/ICICV62344.2024.00041
- Pfitzinger, P., and Freedson, P. (1998). The reliability of lactate measurements during exercise. *Int. J. Sports Med.* 19, 349–357. doi:10.1055/s-2007-971929
- Poole, D. C., Rossiter, H. B., Brooks, G. A., and Gladden, L. B. (2021). The anaerobic threshold: 50+ years of controversy. *J. physiology* 599, 737–767. doi:10.1113/JP279963
- Queiroz, M., Arsa, G., Rezende, D., Sousa, L., Oliveira, F., Araújo, G., et al. (2018). Heart rate variability estimates ventilatory threshold regardless body mass index in young people. *Sci. and sports* 33, 39–46. doi:10.1016/j.scispo.2017.06.005
- Ramos-Campo, D. J., Rubio-Arias, J. A., Ávila-Gandía, V., Marín-Pagán, C., Luque, A., and Alcaraz, P. E. (2017). Heart rate variability to assess ventilatory thresholds in professional basketball players. *J. Sport Health Sci.* 6, 468–473. doi:10.1016/j.jshs.2016.01.002
- Roberts, R. A., and Landwehr, R. (2002). The surprising history of the HRmax=220-age equation. *J. Exerc. Physiology Online* 5, 1–10.
- Schlie, J., Heiber, M., Schittenhelm, A., Beckert, M., Graf, P., and Schmidt, A. (2024). Garmin fenix 7[®] underestimates performance at the lactate threshold in comparison to standardized blood lactate field test. *Curr. Issues Sport Sci. (CISS)* 9, 47–58. doi:10.2147/OAJSM.S444568
- Sears, T., Alvalos, E., Lawson, S., Mcalister, I., Eschbach, L. C., and Bunn, J. (2017). Wrist-worn physical activity trackers tend to underestimate steps during walking. *Int. J. Exerc. Sci.* 10, 764–773. doi:10.70252/yjpp1345
- Seshadri, D. R., Li, R. T., Voos, J. E., Rowbottom, J. R., Alfes, C. M., Zorman, C. A., et al. (2019). Wearable sensors for monitoring the internal and external workload of the athlete. *NPJ Digit. Med.* 2, 71. doi:10.1038/s41746-019-0149-2
- Shiraishi, Y., Katsumata, Y., Sadahiro, T., Azuma, K., Akita, K., Isobe, S., et al. (2018). Real-time analysis of the heart rate variability during incremental exercise for the detection of the ventilatory threshold. *J. Am. Heart Assoc.* 7, e006612. doi:10.1161/JAHA.117.006612
- Shookster, D., Lindsey, B., Cortes, N., and Martin, J. R. (2020). Accuracy of commonly used age-predicted maximal heart rate equations. *Int. J. Exerc. Sci.* 13, 1242–1250. doi:10.70252/XFSJ6815
- Siahkhouian, M., Azizan, S., and Roohi, B. N. (2012). A new approach for the determination of anaerobic threshold: methodological survey on the modified dmax method. *J. Hum. Sport Exerc.* 7, 599–607. doi:10.4100/jhse.2012.72.23
- Siepmann, C., and Kowalczyk, P. (2021). Understanding continued smartwatch usage: the role of emotional as well as health and fitness factors. *Electron. Mark.* 31, 795–809. doi:10.1007/s12525-021-00458-3
- Snyder, N. C., Willoughby, C. A., and Smith, B. K. (2021). Comparison of the polar V800 and the garmin forerunner 230 to predict VO2max. *J. Strength and Cond. Res.* 35, 1403–1409. doi:10.1519/JSC.00000000000002931
- Støve, M. P., Haucke, E., Nymann, M. L., Sigurdsson, T., and Larsen, B. T. (2019). Accuracy of the wearable activity tracker garmin forerunner 235 for the assessment of heart rate during rest and activity. *J. Sports Sci.* 37, 895–901. doi:10.1080/02640414.2018.1535563
- Svedahl, K., and Macintosh, B. R. (2003). Anaerobic threshold: the concept and methods of measurement. *Can. J. Appl. physiology* 28, 299–323. doi:10.1139/h03-023
- Veerabhadrapa, P., Moran, M. D., Renninger, M. D., Rhudy, M. B., Dreisbach, S. B., and Gift, K. M. (2018). Tracking steps on Apple watch at different walking speeds. *J. general Intern. Med.* 33, 795–796. doi:10.1007/s11606-018-4332-y
- Wallace, L., Slattery, K., and Coutts, A. J. (2014). A comparison of methods for quantifying training load: relationships between modelled and actual training responses. *Eur. J. Appl. physiology* 114, 11–20. doi:10.1007/s00421-013-2745-1
- Yoshida, T., Chida, M., Ichioka, M., and Suda, Y. (1987). Blood lactate parameters related to aerobic capacity and endurance performance. *Eur. J. Appl. physiology Occup. physiology* 56, 7–11. doi:10.1007/BF00696368
- Zakynthaki, M. S. (2015). Modelling heart rate kinetics. *PloS one* 10, e0118263. doi:10.1371/journal.pone.0118263
- Ziogas, G. G., Patras, K. N., Stergiou, N., and Georgoulis, A. D. (2011). Velocity at lactate threshold and running economy must also be considered along with maximal oxygen uptake when testing elite soccer players during preseason. *J. Strength and Cond. Res.* 25, 414–419. doi:10.1519/JSC.0b013e3181bac3b9
- Zwiggmann, L., Struett, S., Martin, A., Volmar, P., Bloch, W., and Wahl, P. (2019). Modifications of the dmax method in comparison to the maximal lactate steady state in young Male athletes. *Physician Sportsmed.* 47, 174–181.