Early Version – Uncorrected



Does Exercise Intensity Influence Wrist-Worn Heart Rate Accuracy? Validation of the Apple Watch Following Treadmill Walking and Running in Healthy Young Adult Males

Yan Wu, Kurusart Konharn, Lee David Johnson

DOI: http://dx.doi.org/10.31584/jhsmr.20261523

Received: 18 March 2026

Revised: 14 June 2026

Accepted: 22 June 2026

Published online: 23 November 2026

Please cite this article as: Wu Y, Konharn K, Johnson LD. Does exercise intensity influence wrist-worn heart rate accuracy? Validation of the Apple Watch following treadmill walking and running in healthy young adult males. J Health Sci Med Res doi: http://dx.doi.org/10.31584/jhsmr.20261523

Early Version — Uncorrected: This accepted manuscript has not been typeset or proofread. Content and formatting may change prior to final publication and should not be regarded as the final Version of Record.

© 2026 JHSMR. Hosted by Prince of Songkla University.

This is an open access article under the CC BY-NC-ND license (https://creativecommons.org/licenses/by-nc-nd/4.0/).

Original Article

Does Exercise Intensity Influence Wrist-Worn Heart Rate Accuracy? Validation of the Apple Watch Following Treadmill Walking and Running in Healthy Young Adult Males

Yan Wu, M.Sc.[1], Kurusart Konharn, Ph.D.[2,3], Lee David Johnson, Ph.D. [2,3]

[1]Exercise and Sport Sciences Program, Graduate School, Khon Kaen University, Khon Kaen 40002, Thailand.
[2]School of Physical Therapy, Faculty of Associated Medical Sciences, Khon Kaen University, Khon Kaen 40002, Thailand.
[3]Research Center in Back Neck Other Joint Pain and Human Performance (BNOJPH).
Khon Kaen University, Khon Kaen 40002, Thailand.

*** Corresponding author.**

Kurusart Konharn, Ph.D., PT.

School of Physical Therapy, Faculty of Associated Medical Sciences, Khon Kaen University, Khon Kaen 40002, Thailand.

E-mail: kuruko@kku.ac.th.

ORCID and Email:

Yan Wu

ORCID no: https://orcid.org/0009-0002-0415-8529

E-mail address: yan.w@kkumail.com

Kurusart Konharn

ORCID no.: https://orcid.org/0000-0003-3312-8740

Email: kuruko@kku.ac.th

Early Version - Uncorrected

Lee David Johnson

ORCID no.: https://orcid.org/0009-0004-5627-4606

Email: ljrehabilitation@gmail.com

Running title: HR accuracy in Smartwatch

Abstract

Objective: Wrist-worn wearable technologies are widely used for heart rate (HR) monitoring in both fitness and research settings. However, their measurement performance may vary according to exercise intensity, and evidence regarding the consistency of this intensity dependence remains limited. This study evaluated the agreement of post-exercise HR measurements obtained from the Apple Watch Series 8 compared with those obtained from a fingertip pulse oximeter used in clinical practice following standardized 6-minute treadmill exercise performed at three intensities in healthy young adult males.

Materials and Methods: This method comparison study enrolled fifteen healthy young adult males who completed three standardized 6-minute treadmill exercise bouts of walking (4.0 km/h), brisk walking (7.2 km/h), and running (9.0 km/h). Immediately after each bout, HR was measured using both devices. Agreement was assessed using Pearson's correlation coefficient (r), Lin's concordance correlation coefficient (CCC), Bland-Altman analysis, and absolute error metrics.

Results: Good-to-excellent correlations were observed across all intensities (r = 0.981-0.993), with high concordance (CCC = 0.95-0.99). Bland-Altman analysis demonstrated small systematic bias (within ±2 bpm) and narrow limits of agreement. Absolute error remained low across conditions (MAE: 1.3-1.5 bpm; MAPE < 2%). Although overall agreement was high, concordance was slightly reduced during running compared with lower intensities.

Conclusion: The Apple Watch Series 8 provided accurate post-exercise HR measurements across walking, brisk walking, and running conditions, with a modest intensity-dependent decline in agreement at higher workloads.

Keywords: agreement, heart rate recovery, photoplethysmography, wearable sensors

Early Version - Uncorrected

Introduction

Physiological load is commonly quantified through heart rate (HR) monitoring, which serves as an important indicator for optimizing exercise training, guiding rehabilitation programs, and ensuring safety in clinical settings[1,2]. With rapid advances in sensor technology, wrist-worn wearable devices that use photoplethysmography (PPG), including the Apple Watch and Fitbit, have become increasingly popular, particularly among young adults engaged in physical exercise[3,4]. A major advantage of these devices is their ability to provide continuous, non-invasive physiological monitoring during physical activity under free-living conditions. However, because wrist-based PPG measurements are susceptible to motion artifacts and fluctuations in skin-sensor contact, verification of their accuracy against established reference methods remains essential for valid data interpretation. Accordingly, the fingertip pulse oximeter, a widely accepted clinical tool for measuring oxygen saturation and pulse rate, is frequently used as a reliable reference method for evaluating the performance of wearable HR monitoring technologies in both field and clinical settings[5]. In the present study, only healthy young adult males were included to reduce physiological heterogeneity, as sex-related differences in photoplethysmography signal characteristics have been documented[6]. This allowed a more controlled evaluation of intensity-specific agreement between devices.

Despite the widespread use of wrist-worn PPG devices, their accuracy remains vulnerable to motion artifacts during dynamic activities[3,7]. Factors such as arm swing rhythm, perspiration, and fluctuations in skin-sensor contact pressure may introduce substantial noise into optical signals, potentially compromising data validity. Moreover, empirical evidence suggests that measurement error is not uniformly distributed but may increase with exercise intensity, resulting in reduced reliability at higher heart rates[8]. However, many existing validation studies have been conducted primarily under resting or steady-state conditions[4].

Although previous studies have evaluated the overall accuracy of wearable heart-rate monitors[9], few have systematically examined device agreement across stratified exercise intensities under standardized exercise protocols. These protocols may include clinical graded exercise tests, such as the Bruce or modified Bruce protocols, as well as multi-stage constant-load

Early Version - Uncorrected

submaximal protocols commonly used in wearable method-comparison studies[4]. Given that the primary objective of the present study was to evaluate device agreement across stabilized exercise intensities, a multi-stage constant-load submaximal framework was selected for this investigation. Cumulative evidence comparing different wearable modalities indicates that chest-strap monitors generally demonstrate higher dynamic accuracy than the optical sensors in forearm and wrist-worn devices, whose performance may vary according to both exercise modality and intensity[8-12]. Specifically, previous comparative trials have shown that although wrist-worn PPG sensors achieve high agreement during lower-intensity conditions such as treadmill walking, their measurement consistency tends to decline during vigorous-intensity phases or more dynamic exercise modalities, leading to notable intensity-dependent variability in wearable-derived heart rate data[8-12].

Importantly, 6-minute protocols are widely used as practical submaximal exercise assessments because they provide sufficient physiological loading to elicit meaningful cardiovascular responses while remaining feasible and safe for implementation in both exercise and clinical settings[13,14]. Therefore, the lack of intensity-stratified validation evidence within a standardized 6-minute testing framework represents an important research gap, particularly given the central role of accurate HR monitoring in exercise prescription, physiological monitoring, and safety assessment.

The primary objective of this study was to evaluate the intensity-stratified agreement and reliability of heart rate measurements obtained from the Apple Watch Series 8 compared with a clinical fingertip pulse oximeter following standardized 6-minute treadmill exercise (walking, brisk walking, and running) in healthy young adult males, thereby guiding appropriate device selection for training prescription and safety monitoring. We hypothesized that device agreement would exhibit an intensity-dependent pattern, maintaining high concordance during walking and brisk walking but showing a reduction during running due to increased motion-related optical signal disturbances.

Early Version - Uncorrected

Materials and Methods

Design

A method comparison design was used. Data were collected from October to December 2025. The study was approved by the XXX University Ethics Committee for Human Research, Thailand (approval number: HEXXXXX).

Participants

Fifteen healthy male participants (age: 21.67 ± 2.02 years; BMI: 23.48 ± 3.96 kg/m²) were recruited from Khon Kaen University, Thailand, through university-based social media announcements. Participants aged 18-26 years without a history of cardiovascular disease were eligible. The sample size was formally estimated based on Lin's concordance correlation coefficient (CCC) framework for method comparison studies. Based on the available literature and clinical acceptability thresholds[15], we set the null hypothesis (H_0) at a minimum acceptable CCC of 0.80, and the alternative hypothesis (H_1) at an expected CCC of 0.94. With a two-sided significance level (α) of 0.05 and a statistical power ($1-\beta$) of 0.80, a minimum sample size of 13 participants was determined to be sufficient. To account for potential data loss and enhance statistical rigor, 15 participants (n=15) were ultimately enrolled in this repeated-measures framework, ensuring that the investigation was fully powered to detect meaningful device agreement. All the participants reported engaging in regular physical activity (1-3 sessions per week), as assessed using a standardized physical activity questionnaire16. Potential participants were excluded if they had sustained a lower-extremity injury within the previous 3 months or were using medications that could affect HR or exercise tolerance. Before participation, all individuals completed the Physical Activity Readiness Questionnaire (PAR-Q) and the International Physical Activity Questionnaire-Short Form (IPAQ-SF) to confirm eligibility and characterize habitual physical activity levels[16].

Early Version - Uncorrected

Anthropometry

For anthropometric assessment, participants were instructed to wear light clothing and remove their shoes. Body mass (kg) was measured using a digital scale (Xiaomi Inc., China), whereas standing height was measured using a calibrated stadiometer (Tanita Corp., Tokyo, Japan). Body mass index (BMI) was calculated using the standard formula: body mass (kg)/height² (m^2). All walking and running tests were performed on a Lode Valiant CPET treadmill (Lode B.V., Groningen, Netherlands).

Heart Rate Measurement Devices and Testing Schedule

Heart rate was assessed using a wrist-based wearable device (Apple Watch Series 8; watchOS 11.6.1, build 22U.90; Apple Inc., Cupertino, CA, USA) and a fingertip pulse oximeter (AOJ-70C; Medisana, Germany). The Apple Watch was worn on the participant's left non-dominant wrist, positioned approximately 3 cm proximal to the wrist bone and secured firmly to ensure consistent skin contact throughout testing. Fingertip pulse oximetry was performed on the participant's left middle finger, with placement standardized to the same finger and side across all sessions, and HR was recorded as the displayed average value[5].

Each participant completed three standardized 6-minute treadmill exercise bouts at predefined intensities: walking (4.0 km/h), brisk walking (7.2 km/h), and running (9.0 km/h), in a fixed order, with a 10-minute seated rest interval between bouts. Immediately upon completion of each bout (within 5 s), participants stopped exercising and were seated upright. Fingertip pulse oximeter measurements were then obtained without delay by the principal investigator, and HR values from the Apple Watch were recorded concurrently to enable direct comparison between devices at the same times[10]. To assess measurement consistency, the entire testing protocol was repeated on two separate occasions with a 24-hour interval between sessions.

Standardized Treadmill Exercise Protocol

All participants completed standardized treadmill exercise testing at three predefined intensities: walking (4.0 km/h), brisk walking (7.2 km/h), and running (9.0 km/h)10. Each bout

Early Version - Uncorrected

lasted 6 minutes, consistent with commonly used submaximal exercise protocols designed to elicit stable cardiovascular responses while maintaining feasibility in laboratory-based testing[13,14]. The protocol was performed on a research-grade treadmill (Lode B.V., Groningen, Netherlands).

Participants were instructed to maintain the assigned treadmill speed throughout each 6-minute bout and to avoid unnecessary upper-limb movement that might interfere with wrist-based sensor recordings. To assess measurement reliability, the full protocol was repeated on two separate testing occasions with a 24-hour interval, a design commonly used in wearable validation studies to reduce fatigue-related confounding while enabling short-term test-retest assessment[17]. The order of exercise intensities was kept consistent across participants (walking, followed by brisk walking and running) to ensure a standardized physiological progression and minimize abrupt workload transitions. A 10-minute seated recovery period was provided between stages to reduce carryover effects and allow HR to return toward baseline[18].

Safety Monitoring

All exercise sessions were supervised by trained personnel. Participants were monitored for signs of discomfort, dizziness, or abnormal responses during treadmill testing. The protocol followed general safety recommendations for exercise testing in young adults[19], and testing was discontinued if participants requested to stop or if any adverse symptoms occurred.

Statistical Analysis

Statistical analyses were performed using MedCalc (version 23.4.8). Descriptive statistics are presented as mean ± standard deviation (SD). Normality of HR data was assessed using the Shapiro-Wilk test and visual inspection of distribution plots.

Agreement between wrist-worn and fingertip-based average HR measurements was examined separately for each exercise intensity condition (walking, brisk walking, and running). Linear association between methods was evaluated using Pearson's correlation coefficient (r). Method agreement was further quantified using Lin's concordance correlation coefficient (CCC), which accounts for both precision and accuracy. Bland-Altman analysis was performed to assess

Early Version - Uncorrected

systematic bias and 95% limits of agreement (LoA), calculated as the mean difference (wrist - fingertip) ± 1.96 × SD of the differences. Bias was interpreted as the average directional difference between measurement methods. All analyses were stratified by exercise intensity to evaluate intensity-dependent agreement.

In addition, absolute error was evaluated using mean absolute error (MAE), root mean square error (RMSE), and mean absolute percentage error (MAPE), with fingertip pulse oximeter values treated as the reference method. These metrics were calculated for each intensity condition to determine whether measurement error varied with exercise intensity.

For the primary agreement analysis, heart rate values from the two testing sessions were averaged for each participant within each exercise intensity before calculating agreement and error metrics. This approach yielded one paired wrist-worn and fingertip heart rate value per participant for each intensity condition (n = 15), reducing within-participant day-to-day variation while preserving participant-level independence for the method-comparison analysis.

Table 1. Agreement between wrist-worn and fingertip heart rate measurements after 6-minute exercise across exercise intensities

Table 2. Absolute error metrics between wrist-worn and fingertip heart rate measurements after 6-minute exercise

Figure 1 Bland-Altman plots showing the agreement between wrist-worn and fingertip HR measurements across exercise intensities. The solid line represents the mean bias, and the dashed lines represent the 95% limits of agreement (n = 15 for each condition).

Figure 2 Comparison of measurement bias and stability across activity intensities. The bar chart illustrates the mean difference (wrist - fingertip), and error bars indicate the standard deviation (n = 15 per condition).

Early Version - Uncorrected

Results

Participant Characteristics

Fifteen healthy young adult males completed all treadmill testing sessions without adverse events. Participant characteristics were as follows: age 21.67 ± 2.02 years, height 173.27 ± 5.19 cm, weight 70.67 ± 13.33 kg, and BMI 23.48 ± 3.96 kg/m². Post-exercise HR increased progressively with exercise intensity, ranging from 84.27 ± 15.89 bpm versus 83.93 ± 14.89 bpm during walking, to 140.93 ± 10.10 bpm versus 141.93 ± 10.42 bpm during brisk walking, and 167.80 ± 5.67 bpm versus 169.33 ± 7.07 bpm during running for the wrist-worn and fingertip devices, respectively. Across all intensities, mean HR values were highly comparable between devices, with only minimal between-device differences. Agreement and absolute error indices for the wrist-worn wearable device and fingertip pulse oximeter across exercise intensities are summarized in Tables 1 and 2; these values were calculated from the participant-level average of the two testing sessions, resulting in n = 15 paired observations per intensity. Individual agreement across the HR range for each intensity condition is illustrated in the Bland-Altman plots (Figure 1). In addition, Figure 2 shows the comparison of measurement bias and stability across the three activity conditions, indicating a trend toward greater average bias but reduced variability during higher-intensity exercise.

Walking

During walking, the mean HR measured by the wrist device was 84.27 ± 15.89 bpm, whereas the fingertip pulse oximeter recorded 83.93 ± 14.89 bpm. HR measurements from the wrist and fingertip devices showed a very strong correlation (Pearson r = 0.9811) and excellent agreement (Lin's CCC = 0.9788) (Table 1). Bland-Altman analysis (Figure 1A) revealed an average bias of 0.33 bpm, with 95% LoA ranging from -5.85 to 6.52 bpm. The absolute error metrics further supported strong consistency, with an MAE of 1.53 bpm, RMSE of 3.07 bpm, and MAPE of 1.69% (Table 2).

Early Version - Uncorrected

Brisk Walking

At the brisk walking intensity, the mean HR measured by the wrist device was 140.93 ± 10.10 bpm, whereas the fingertip device recorded 141.93 ± 10.42 bpm. Agreement remained very high, with Pearson r = 0.9930 and Lin's CCC = 0.9875 (Table 1). As shown in Figure 1B, the average bias was -1.00 bpm, indicating that the wrist-based measurement was slightly lower than the fingertip measurement, and the 95% LoA were narrow (-3.46 to 1.46 bpm). Error analysis showed that the overall absolute difference was the lowest among the three intensities, with an MAE of 1.27 bpm, RMSE of 1.57 bpm, and MAPE of 0.89% (Table 2).

Running

During running, HR increased substantially, with mean values of 167.80 ± 5.67 bpm for the wrist device and 169.33 ± 7.07 bpm for the fingertip device. Although the correlation remained strong (Pearson r = 0.9839), agreement was lower than that observed during walking and brisk walking (Lin's CCC = 0.9471) (Table 1). Bland-Altman analysis (Figure 1C) showed an average bias of -1.53 bpm, with 95% LoA from -3.74 to 0.67 bpm, indicating a slight underestimation of HR by the wrist device compared with the fingertip pulse oximeter. Absolute error remained low, with an MAE of 1.53 bpm, RMSE of 1.88 bpm, and MAPE of 0.90% (Table 2).

Discussion

The present study evaluated the agreement of post-exercise HR measurements obtained from the Apple Watch Series 8 compared with those obtained from a fingertip pulse oximeter following standardized 6-minute treadmill exercise at three intensities in healthy young adult males. The principal finding was that the Apple Watch demonstrated consistently high agreement with the fingertip reference method across walking, brisk walking, and running conditions. Strong correlations were observed at all intensities (Pearson r = 0.981-0.993), accompanied by excellent concordance (Lin's CCC = 0.947-0.988), supporting the agreement of the wrist-worn device with the fingertip pulse oximeter for post-exercise HR assessment.

Early Version - Uncorrected

Bland-Altman analysis showed that systematic bias was minimal across intensities, with the mean difference ranging from 0.33 bpm during walking to -1.53 bpm during running. In addition, absolute error remained low under all conditions, with MAE values ranging from 1.27 to 1.53 bpm and MAPE values consistently below 2%. Notably, agreement showed an intensity-dependent pattern, with the highest concordance during brisk walking (CCC = 0.9875) and a modest reduction during running (CCC = 0.9471). Collectively, these findings indicate that the Apple Watch Series 8 provides accurate post-exercise HR measurements across moderate-to-vigorous treadmill exercise intensities, although performance may decline slightly as intensity increases.

The observed intensity-dependent decline in concordance is likely attributable to physiological and biomechanical factors inherent to wrist-based PPG sensing. As exercise intensity increases, arm swing amplitude, impact-related vibration, and rapid changes in peripheral perfusion may introduce motion artifacts and transient disruptions in sensor-skin contact, thereby reducing concordance between devices. In the present study, this was reflected mainly by a slightly greater negative bias during running rather than by wider limits of agreement, as variability was not greatest at the highest intensity. The negative bias observed during brisk walking and running suggests slight underestimation by the wearable device, which may reflect reduced optical signal stability during dynamic movement and the difficulty of accurately tracking higher HR values using reflective PPG technology.

These observations are consistent with previous validation studies. Gillinov et al. reported that wrist-based HR monitors generally demonstrate acceptable accuracy during treadmill walking but reduced consistency as exercise intensity increases[10]. Hough et al. similarly showed that the accuracy of wrist-worn HR monitors tends to decline during higher-intensity exercise because of movement-related artifacts[9]. In addition, Shcherbina et al. found greater variability and reduced precision in wrist-worn wearables at higher heart rates, supporting the concept that intensity-dependent error is an inherent limitation of optical wearable sensors[8]. Taken together, the present findings align with the broader literature indicating that commercial wearable devices can achieve acceptable HR validity in controlled laboratory settings, although their performance remains sensitive to movement characteristics and exercise intensity[4].

Early Version - Uncorrected

From a practical perspective, the small bias values (≤1.53 bpm) and narrow LoA observed across all stages suggest that the Apple Watch can provide clinically and practically comparable HR measurements to a fingertip reference during standardized 6-minute treadmill protocols, particularly at walking and brisk walking intensities. However, the modest reduction in concordance during running highlights the need for cautious interpretation when wrist-based HR monitoring is used for higher-intensity exercise prescription, performance monitoring, or threshold-related decision-making.

The present findings are also relevant for applied exercise and rehabilitation settings. Accurate HR monitoring is essential for exercise prescription, training load management, and safety evaluation1. The consistently low absolute error observed in this study suggests that the Apple Watch Series 8 may be suitable for post-exercise HR assessment during standardized treadmill testing, particularly when measurements are obtained immediately after exercise rather than continuously during movement. Agreement was strongest during brisk walking, supporting its use in submaximal monitoring contexts where relatively precise HR-based regulation is required, such as controlled aerobic training, exercise prescription, and standardized fitness testing.

At lower intensity (walking), the small mean difference and acceptable LoA suggest that wrist-worn monitoring may also be suitable for routine clinical and rehabilitation applications, including safety monitoring and recovery HR tracking after low-load exercise. In contrast, although agreement remained high during running, concordance was modestly reduced, and the wearable device slightly underestimated HR. Therefore, wrist-derived HR values should be interpreted with greater caution during higher-intensity exercise. Although the observed magnitude of error is unlikely to be clinically meaningful for general monitoring, the lower concordance at running intensity may limit the suitability of wrist-based HR measurements for precision-dependent applications, such as threshold identification or high-intensity training zone prescription.

Overall, these results support the practical use of the Apple Watch for post-exercise HR monitoring during standardized 6-minute treadmill protocols, particularly under walking and brisk

Early Version - Uncorrected

walking conditions, while highlighting that higher-intensity running may require more conservative interpretation when accurate HR thresholds are important.

Strengths and Limitations

A key strength of the present study is the standardized repeated-measures design, which evaluated agreement across three controlled treadmill intensities and repeated the protocol across two sessions separated by 24 hours. This design allowed participant-level averaging across sessions, reducing day-to-day variation while maintaining consistent laboratory conditions. In addition, the study used multiple complementary statistical approaches, including the Pearson correlation, Lin's CCC, Bland-Altman analysis, and absolute error metrics, thereby providing a comprehensive evaluation of wearable performance.

Nevertheless, several limitations should be acknowledged. First, the sample size was relatively small (n = 15), although it was comparable to that of previous wearable validation studies[9,10]. Second, participants were limited to healthy young adult males, which restricts the generalizability of the findings to females, older adults, and clinical populations. Sex-related differences in skin thickness, vascular characteristics, and body composition may influence PPG accuracy and should therefore be examined in future studies. Third, the fingertip pulse oximeter was used as the reference method rather than electrocardiography (ECG), which remains the gold standard for HR measurement. Although fingertip pulse oximetry is widely accepted as a clinically reliable method, future studies should include ECG comparisons to strengthen conclusions regarding validity.

In addition, the present study measured HR immediately after exercise rather than continuously during exercise. Therefore, the findings primarily reflect post-exercise HR assessment rather than real-time HR tracking. Future research should also investigate whether similar agreement patterns are observed during continuous treadmill running, in which motion artifacts may be more pronounced.

Early Version - Uncorrected

Future Research

Future studies should evaluate Apple Watch HR performance in more diverse populations, including females, older adults, and individuals with cardiovascular or metabolic conditions. In addition, comparisons with ECG during both continuous exercise and recovery phases would provide stronger evidence regarding device validity. Future work should also explore wearable HR accuracy under outdoor running conditions, where environmental factors such as temperature, sweat rate, and uneven terrain may further influence sensor stability. Finally, examining whether agreement differs according to strap tightness, wrist position, or dominant versus non-dominant wrist placement may help optimize wearable measurement protocols.

Conclusion

The Apple Watch Series 8 demonstrated high agreement compared with a fingertip pulse oximeter for measuring post-exercise heart rate following standardized 6-minute treadmill walking, brisk walking, and running in healthy young adult males. Agreement remained strong across all intensities, with small systematic bias and low absolute error. However, concordance showed a modest exercise intensity-dependent decline during running. Overall, these findings support the practical utility of wrist-based wearable devices for post-exercise heart rate assessment in controlled treadmill-based exercise settings, while emphasizing the importance of considering exercise intensity when interpreting wearable-derived HR outcomes.

Acknowledgment

The authors sincerely thank all the students who participated in this study.

Early Version - Uncorrected

Authorship Contribution

All authors were involved in the interpretation of the results and preparation of the manuscript. All authors met the criteria for authorship and approved the final version of the manuscript.

Funding

No external funding.

Ethics approval: The study was approved by the Ethics Committee of Khon Kaen University （HE682040）.

Statement: All authors confirm that the material is original, has not already been published, and has not been, nor will it be, submitted for publication elsewhere, so long as it is under consideration by JHSMR.

Conflict of Interest

The authors declare no competing interests.

References

1. Achten J, Jeukendrup AE. Heart rate monitoring: applications and limitations. Sports Med 2003;33(7):517-38.
2. Halson SL. Monitoring training load to understand fatigue in athletes. Sports Med 2014;44(Suppl 2):139-47.
3. Bent B, Goldstein BA, Kibbe WA, Dunn JP. Investigating sources of inaccuracy in wearable optical heart rate sensors. NPJ Digit Med 2020;3:18.
4. Fuller D, Colwell E, Low J, Orychock K, Tobin MA, Simango B, et al. Reliability and validity of commercially available wearable devices for measuring steps, energy expenditure, and heart rate: systematic review. JMIR Mhealth Uhealth 2020;8(9):e18694.
5. Chan ED, Chan MM, Chan MM. Pulse oximetry: understanding its basic principles facilitates appreciation of its limitations. Respir Med 2013;107(6):789-99.

Early Version - Uncorrected

6. Dehghanojamahalleh S, Kaya M. Sex-related differences in photoplethysmography signals measured from finger and toe. IEEE J Transl Eng Health Med 2019;7:1-7.
7. Orini M, Guvensen G, Jamieson A, Chaturvedi N, Hughes AD. Movement, sweating, and contact pressure as sources of heart rate inaccuracy in wearable devices. In: 2022 Computing in Cardiology (CinC). IEEE; 2022.
8. Shcherbina A, Mattsson CM, Waggott D, Salisbury H, Christle JW, Hastie T, et al. Accuracy in wrist-worn, sensor-based measurements of heart rate and energy expenditure in a diverse cohort. J Pers Med 2017;7(2):3.
9. Hough P, Glaister M, Pledger A. The accuracy of wrist-worn heart rate monitors across a range of exercise intensities. J Phys Act Res 2017;2(2):112-6.
10. Gillinov S, Etiwy M, Wang R, Blackburn G, Phelan D, Gillinov AM, et al. Variable accuracy of wearable heart rate monitors during aerobic exercise. Med Sci Sports Exerc 2017;49(8):1697-703.
11. Ho WT, Yang YJ, Li TC. Accuracy of wrist-worn wearable devices for determining exercise intensity. Digit Health 2022;8:20552076221124393.
12. Zhang Y, Weaver RG, Armstrong B, Burkart S, Zhang S, Beets MW. Validity of wrist-worn photoplethysmography devices to measure heart rate: a systematic review and meta-analysis. J Sports Sci 2020;38(17):2021-34.
13. ATS Committee on Proficiency Standards for Clinical Pulmonary Function Laboratories. ATS statement: guidelines for the six-minute walk test. Am J Respir Crit Care Med 2002;166(1):111-7.
14. Du H, Wonggom P, Tongpeth J, Clark RA. Six-minute walk test for assessing physical functional capacity in chronic heart failure. Curr Heart Fail Rep 2017;14(3):158-66.
15. Pasadyn SR, Soudan M, Gillinov M, Houghtaling P, Phelan D, Gillinov N, et al. Accuracy of commercially available heart rate monitors in athletes: a prospective study. Cardiovasc Diagn Ther 2019;9(4):379-85.
16. Craig CL, Marshall AL, Sjostrom M, Bauman AE, Booth ML, Ainsworth BE, et al. International physical activity questionnaire: 12-country reliability and validity. Med Sci Sports Exerc 2003;35(8):1381-95.

Early Version - Uncorrected

17. Rogers DW, Himariotis AT, Sherriff TJ, Proulx QJ, Duong MT, Noel SE, et al. Test-retest reliability and concurrent validity of photoplethysmography finger sensor to collect measures of heart rate variability. Sports (Basel) 2025;13(2):29.

18. Cozgarea A, Cozma D, Teodoru M, Lazar-Hocher AI, Cirin L, Faur-Grigori AA, et al. Heart rate recovery: up to date in heart failure-a literature review. J Clin Med 2024;13(11):3328.

19. Ozemek C, Bonikowske A, Christle J, Gallo P. ACSM's guidelines for exercise testing and prescription. Lippincott Williams & Wilkins; 2025.

Early Version - Uncorrected

Table 1. Agreement between wrist-worn and fingertip heart rate measurements after 6-minute exercise across exercise intensities

Intensity	n	Pearson r	Lin's CCC	95% CI (CCC)	Bias (Wrist − Fingertip)	95% CI (Bias)	LOA Lower (bpm)	95% CI (Lower LOA)	LOA Upper (bpm)	95% CI (Upper LOA)
Walking	15	0.981	0.979	0.9412 - 0.992	0.333	−1.41 - 2.08	−5.85	−8.90 - −2.79	6.52	3.46 - 9.57
Brisk Walking	15	0.993	0.987	0.966- 0.995	-1.000	−1.69 - −0.31	−3.46	−4.67 - −2.24	1.46	0.24- 2.67
Running	15	0.983	0.947	0.873- 0.978	-1.533	−2.15 - −0.91	−3.74	−4.83 - −2.65	0.67	−0.42 - 1.76

Note: Bias represents the mean difference between methods (wrist - fingertip). Limits of agreement (LoA) were calculated as mean difference ± 1.96 SD.

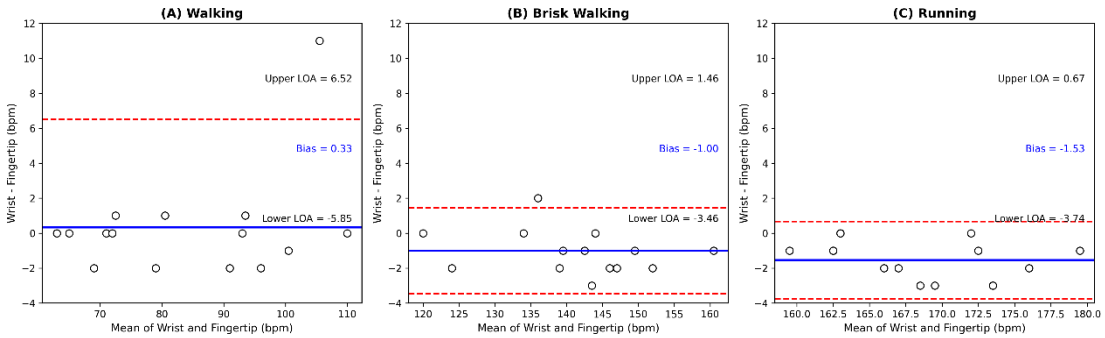
Abbreviations: CCC, concordance correlation coefficient; LoA, limits of agreement; SD, standard deviation.

Table 2. Absolute error metrics between wrist-worn and fingertip heart rate measurements after 6-minute exercise

Intensity	n	MAE (bpm)	RMSE (bpm)	MAPE (%)
Walking	15	1.53	3.07	1.69
Brisk Walking	15	1.27	1.57	0.89
Running	15	1.53	1.88	0.90

Note: MAE, mean absolute error; RMSE, root mean square error; MAPE, mean absolute percentage error. Error metrics were calculated using fingertip measurements as the reference method.

Early Version - Uncorrected



(A) (B) (C)

Figure 1 Bland-Altman plots showing the agreement between wrist-worn and fingertip HR measurements across exercise intensities. The solid line represents the mean bias, and the dashed lines represent the 95% limits of agreement (n = 15 for each condition).

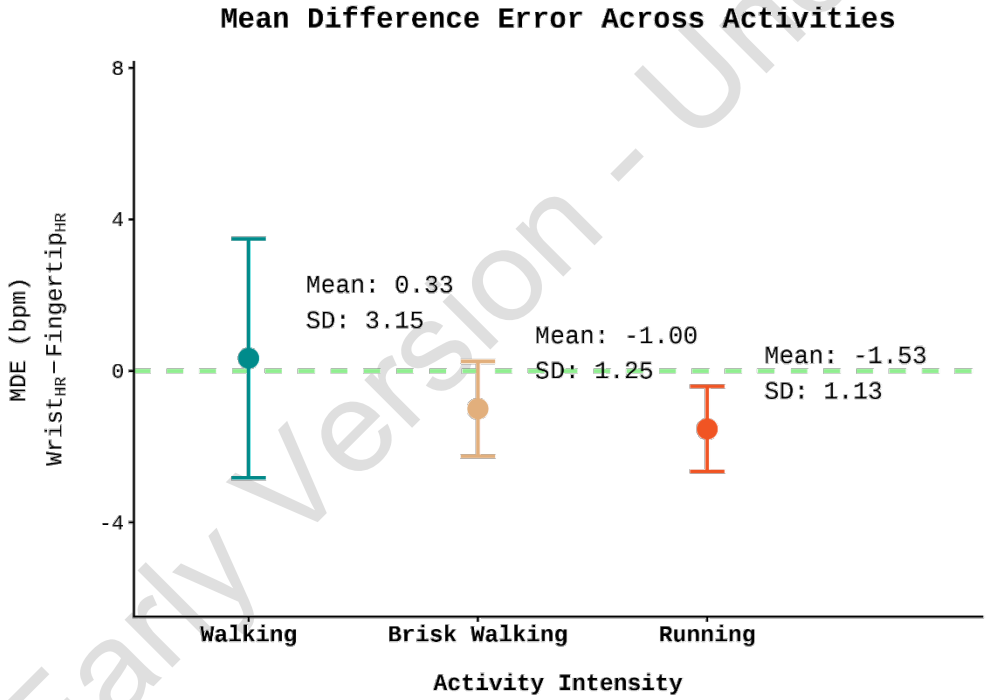


Figure 2 Comparison of measurement bias and stability across activity intensities. The bar chart illustrates the mean difference (wrist - fingertip), and error bars indicate the standard deviation (n = 15 per condition).

Early Version - Uncorrected