4-24-2023

Measuring Energy Expenditure and Heart Rate during Maximum Aerobic Testing with the Apple Watch Series 7

Flavia Rusterholz
Nova Southeastern University, fr433@mynsu.nova.edu

Corey A. Peacock
Nova Southeastern University, cpeacok@nova.edu

Andrew Rodriguez
Nova Southeastern University

Victoria Ortiz
Nova Southeastern University, vo167@mynsu.nova.edu

RESEARCH DIRECTS in HEALTH SCIENCES

2023, Volume 3 (Issue 1): 7 OPEN ACCESS

Measuring Energy Expenditure and Heart Rate during Maximum Aerobic Testing with the Apple Watch Series 7

Original Research

Flavia Rusterholz[1], Andrew Rodriguez[1], Victoria Ortiz[1], Corey A. Peacock[1]

[1]*Department of Health and Human Performance, Nova Southeastern University, Davie, FL, USA,*

Open Access

Published: April 24, 2023

Research Directs in Health Sciences: *2023, Volume 3 (Issue 1): 7*

ISSN: 2768-492X

Abstract

Introduction: Wrist-worn devices such as the Apple Watch have emerged as technology for tracking physical activity. The aim of this research study is to analyze the Apple Watch Series 7 (AW7) with measurements of the maximum heart rate (MHR) and maximum energy expenditure (MEE) during a maximal aerobic capacity test on the treadmill. AW7 measurements will be compared to the Polar Heart Rate Monitor (Polar) and the PARVO Metabolic Cart (PARVO).

Methods: 22 healthy and active subjects (mean ± SD: age 23.8 ± 4.0 years; BMI 23.0 ± 5.9 kg/m^2) volunteered for the study. The subjects confirmed their activity, health status, and were measured for body composition and aerobic capacity.

Results: No significant difference was found in MEE between PARVO (109.6 ± 41.7 kcal) and AW7 (98.7 ± 24.3 kcal) conditions; $t(21)=1.5$, $p = 0.153$. In addition, there was no significant difference in MHR between PARVO (186.2 ± 16.2 BPM) and AW7 (189.3 ± 8.5 BPM) conditions; $t(21)=-0.9$, $p = 0.379$.

Conclusions: The main findings of this study show that the MEE as well as the MHR between the AW7 compared to the PARVO are not different.

Key Words: Aerobic capacity testing, heart rate measurement.

Corresponding author: FLAVIA RUSTERHOLZ, fr433@mynsu.nova.edu

Introduction

The Apple Watch (AW) as well as the Samsung Watch are suggested the most prominent wrist-worn devices [1]. Statistics show that within the last three months of 2016, Apple shipped out more smartwatches than the industry of Swiss watches combined [2]. The popularity gain of the AW has been granted as a convenient health monitoring wellness tool [3]. It provides physiological features such as estimations of maximum energy expenditure (MEE), maximum heart rate (MHR) measurements, exercising minutes and total minutes of standing, which finally promotes a healthier lifestyle [4,5]. However, despite the popularity of wearing an AW for fitness tracking, their validity in measuring MHR as well as MEE must be confirmed. Previous research suggests that the MHR measurements of AW are generally reliable, whereas MEE is not [6-9].

Heart Rate Measurements of Apple Watches

The AW assesses heart rate (HR) using skin photoplethysmography, which allows measuring HR through green and infra-red lights as well as photodetectors to examine changes within the blood volume slightly below the surface of the

skin [10]. Through previously conducted research it can be identified that the AW does in fact provide precise MHR readings in terms of one's performance (i.e. running at different velocities, walking, resistance exercise, cycling, and elliptical) [6-9,11,12]. There were multiple studies comparing the accuracy of the MHR to either a 3- or a 12-lead electrocardiogram, whereas all of them confirmed the AW to be a valid instrument of measuring MHR during different activities [6,7,13-15]. Other studies confirmed the accuracy of the AW MHR measurements through comparison with a Polar heart rate monitor (Polar) chest strap [8,9,11,12] and/or electrocardiographic limb leads [11,12]. Additional studies have shown that the MHR was most accurate during sedentary behavior and walking, however, as running speed and intensity increased, measurement errors did as well [11,12,15,16]. Furthermore, most studies included multiple wrist-worn devices to compare the MHR accuracy among the different brands, whereas the AW measurements consistently provided the lowest error [6-9,11,12,14,15,17-19]. Since the AW is the most accurate wrist-worn device for measuring MHR, even across different series (series 1-6), it is suggested utilizing the MHR measurement of those watches for cardiac rehabilitation [6]. One previous study assessed the first three stages of the Bruce Protocol, however to date, no studies have compared MHR measurements of the AW series 7 (AW7) during a maximal oxygen consumption test (VO_2 max test) [14]. Overall, previously performed studies conclude that MHR measurements from the AW across different series provide accurate results. Moreover, when comparing the AW to other popular wrist-worn devices, it consistently shows the lowest error rate for MHR measurements as well as MEE estimations [7,8].

Energy Expenditure Estimations of Apple Watches

The AW possesses the function of estimating the calories burnt during various activities as well as the total daily MEE [5]. The MEE displayed on the AW gets predicted through the individual's metabolic rate [8,20]. The metabolic rate is another prediction based on a person's gender, weight, and height, which can be modified on an iPhone app that connects to the AW [5]. However, the outcomes of previous studies show that the reading accuracies of MEE on AW varies between the series as well as the intensity level of exercise performed when comparing to an indirect calorimetry [6-9,13,14,17-19,21-25]. Most studies included multiple wrist-worn wearables to determine which device consistently produced the lowest error rate for measuring MEE during various activities [7-9,14,17-19,22-24]. The results showed that the AW provided the most accurate results across multiple studies compared to all other wearables tested [7-9,18,19,23,24]. However, the error rate was still too high to consider the AW a valid instrument for measuring one's MEE during exercise [6-9,13,17,19,23]. On the contrary, studies indicated that the AW series 1 and 2 are valid measurements of MEE during running on the treadmill when compared to the indirect calorimetry [25,26]. It is well supported by previous findings that the MEE estimations of AW have poor accuracy, however, it is important to note that there has only been research done with Apple Watches up to series 6.

As an overview of the literature, the results show that the MHR assessment of AW during different physical activities is accurate, whereas the MEE measurements tend to be inaccurate. The purpose of this study is to compare the AW7 with the metabolic cart (PARVO) during a Bruce Protocol VO_2 max test to provide a better understanding of the measurements of MHR and MEE. The literature shows that the AW measurements of MHR and MEE have not been compared to a criterion device during a maximal exertion test. It is hypothesized that the MHR assessment results will be indifferent as past AW models have shown precise measurements whereas the MEE will be different when compared to a PARVO during a maximal exertion test.

Scientific Methods

Participants

25 participants originally enrolled in this study, whereas only 22 participants were able to complete the study. All the 22 subjects (mean ± SD: age 23.8 ± 4.0 years; height 175.2 ± 10.4 cm; weight 70.4 ± 6.1 kg; BMI 23.0 ± 5.9 kg/m^2) volunteered to be a part of this one-time interaction study. After the subjects entered the lab, they were asked to read and sign the informed consent form, while researchers were there to answer any upcoming questions about the procedure. To be excluded from the study, the subject must have had or still has heart issues, dizziness, prescribed medication for a chronic medical condition, or any injuries stopping the participant of running. The study was approved by the Nova Southeastern University Institutional Review Board and subjects provided informed consent to participate in the study.

Protocol

Physical Activity Readiness Questionnaire for Everyone

Participants filled out the Physical Activity Readiness Questionnaire for everyone (PAR-Q+) to evaluate their health status. It is confirmed that the PAR-Q+ serves as an international standard screening tool to assess the risks in pre-

physical activity participation [27]. The requirements to complete the study were for the subjects to be healthy, active, and aged between 18-45 to participate. To be considered active, the subject must be regularly trained, meaning at least work out twice a week for the past consecutive three months. If any questions of the first part within the PAR-Q+ are answered with "yes", the subject was excluded from the study, as the health status of the individual was not considered "cleared" to complete the testing.

InBody Assessment

The subject's height was measured and an InBody assessment was taken, to assess their body composition. The body composition evaluation includes measuring weight, lean body mass, fat mass, total body water and percent body fat. It is suggested that the InBody scan is a reliable tool to use for body composition assessments [28].

Rate of Perceived Exertion Scale

The rate of perceived exertion (RPE) scale is a way of measuring self-reported physical activity intensity, whereas it is suggested that the CR-10 Borg Scale is a well-accepted tool in monitoring the subject's load during testing [29]. The scale ranges from 0 (nothing at all) to 10 (extremely strong, almost max). Thus, whenever an RPE of 10 was indicated by the subject, the test was concluded soon thereafter as the maximal aerobic capacity (VO_2 max) of the individual was reached. Furthermore, the subjects were instructed to push the red button on the treadmill saying "STOP" in case of emergency.

Maximal Oxygen Consumption Testing

Subjects were prepared for the laboratory exercise testing. The researchers adjusted the gender, age, weight, and height on the AW activity app as well as on the PARVO. The PARVO analyzes the subject's oxygen consumption, which is the air expired of the participant during running on the treadmill to measure the MEE. It is concluded that metabolic cart testing is suitable for both maximal as well as submaximal aerobic capacity testing [30]. The VO_2 max test is performed using the Bruce protocol [31], whereas it is determined that the Bruce protocol is a traditionally used measurement to assess the maximal aerobic capacity for healthy subjects [32]. The subjects were informed that whenever they give up, meaning they have reached their VO_2 max, they can step off the treadmill and the test will be stopped immediately both on the PARVO and the AW7. The subjects were asked to refer from any talking or grunting during the test, as this can alter the MEE measured by the PARVO. Prior to testing, participants were assisted in the positioning of the AW7, Polar chest strap model H10 for the MHR measurement as well as the facial mask covering the nose and mouth for the PARVO to obtain the MEE measurement. The Polar assesses MHR through the PARVO; thus, this study compares the measurement of MHR of a wrist-worn device (AW), and a chest strap (Polar). The Polar chest strap's validity has been approved [33]. Then, the subject was finally ready to step onto the treadmill. The AW7 and the PARVO were started at the exact same time. The RPE and MHR was assessed after each stage of the Bruce Protocol (every 3 minutes), whereas the subject pointed at the number indicating the RPE level and the MHR was read from the AW7. As soon as the subject reached its maximal effort, the test was ended on the watch and the metabolic cart at the same time. Then the mask was then removed, and the subject walked around the laboratory for 2 minutes. After those 2 minutes, the MHR was read again from both the AW7 as well as the Polar for any potential differences.

Statistical Analysis

Descriptive statistics (means and standard deviations) were calculated for age, height, weight, and BMI. A paired samples t-test was utilized to compare MEE and MHR between PARVO and AW7. Pearson's Correlation was utilized to analyze relationships between multiple variables; MEE (PARVO and AW), MHR (PARVO and AW) height, weight, VO_2 max, and body fat percentage. Significance was set at $P \leq 0.05$ and all analysis was conducted via SPSS version 27.

Results

All descriptive data including age, height, weight, and BMI was calculated for the subjects. A paired samples t-test was conducted to compare the MEE and MHR using the PARVO and the AW7 during a VO_2 max test. Below shows that there was no significant difference between the PARVO and AW7 when comparing the MEE and MHR during the maximal aerobic test with the conditions (Table 1).

A Pearson correlation analysis was run, to find significant relationships between variables. Below shows the Pearson correlation analysis of MEE (from PARVO and AW7) and MHR (from PARVO and AW7) with multiple other measured variables (Table 2).

Table 1. Comparisons between AW7 and PARVO of MEE and MHR

	PARVO (n = 22)	**AW7** (n = 22)	**p-Value** (n = 22)
MEE (kcal)	109.6 ± 41.7	98.7 ± 24.3	p = 0.153
MHR (BPM)	186.2 ± 16.2	189.3 ± 8.5	p = 0.379

*Significant Value p ≤.005; ** Significant Value p = <0.001

Table 2. Correlations between multiple variables of MEE and MHR

	MEE PARVO	**MEE AW7**	**MHR PARVO**	**MHR AW7**
Height (cm)	r = 0.597 P = 0.003*	r = 0.186 P = 0.406	r = 0.206 P = 0.358	r = 0.105 P = 0.643
Weight (kg)	r = 0.541 P = 0.009*	r = -0.034 P = 0.879	r = -0.098 P = 0.664	r = -0.144 P = 0.522
VO_2 max (mL/kg/min)	r = 0.772 P = <0.001**	r = 0.729* P = <0.001	r = -0.248 P = 0.265	r = -0.200 P = 0.372
Body fat (%)	r = -0.572 P = 0.005*	r = -0.335 P = 0.121	r = 0.122 P = 0.590	r = 0.210 P = 0.349

*Significant Value p ≤.005; ** Significant Value p = <0.001

Discussion

The outcomes of this study show that there was no significant difference between the MHR measurement of the Polar when compared with either the PARVO or AW7. This aligns with previous research, where the MHR measurement of the AW had great validity, when compared to a 12-lead ECG, a 3-lead ECG, or a Polar H7 chest strap [6-9]. Contrary to the results of another investigation, where they found that the MHR was slightly underestimated, the mean of the MHR of the AW7 in the present study was slightly overestimated compared to the PARVO [14].

Like the MHR outcomes, when comparing MEE on both the AW7 and the PARVO, it also shows that there is no difference between the variables, which is in agreement with previously performed studies [21,22,25]. Contrary to the results MEE found, most of the studies (performed with past models) do not suggest the MEE of AW to be valid as the measurement errors are too large [6-9,13,19,23,34].

Moreover, the MEE PARVO had several positive correlations with other variables, such as height (r= 0.597, P = 0.003) and weight (r = 0.541, P = 0.009), whereas the MEE AW7 did not have a significant correlation with any of these mentioned variables. This suggests that the AW7 is unable to take height and weight into consideration when tracking physical activity. Nevertheless, the present study is in alignment with studies previously performed where BMI, age, and sex did not influence the accuracy of the MHR AW7 and neither with the PARVO [11,12].

Conclusions

In conclusion, contrary to the hypothesis the results of the study suggest that not only the MHR of the AW7, but also its MEE measurement were indifferent from the PARVO assessment, coinciding with findings of previous studies. Thus, it suggests that the AW7 is a suitable instrument to measure both variables MEE and MHR when compared to the PARVO and Polar during a VO_2 max test performed on a treadmill.

Acknowledgements

None

Conflict of Interest

No conflict of interest.

References

1. Tien NH DN, Chi DT. Product policy in international marketing comparative analysis between Samsung and Apple. *IJAFR*. 2019;1(2):129-133.
2. Massoomi MR HE. Increasing and Evolving Role of Smart Devices in Modern Medicine. *Eur Cardiol*. 2019;14(3)doi:10.15420/ecr.2019.02
3. Lui GY, Loughnane D, Polley C, Jayarathna T, Breen PP. The Apple Watch for Monitoring Mental Health–Related Physiological Symptoms: Literature Review. *JMIR Mental Health*. 2022;9(9):e37354.
4. Jo A, Coronel BD, Coakes CE, Mainous AG, 3rd. Is There a Benefit to Patients Using Wearable Devices Such as Fitbit or Health Apps on Mobiles? A Systematic Review. *Am J Med*. Dec 2019;132(12):1394-1400.e1. doi:10.1016/j.amjmed.2019.06.018
5. Apple. *Empowering people to live a healthier day*. 2022. https://www.apple.com/newsroom/pdfs/Health-Report-July-2022.pdf
6. Falter M, Budts W, Goetschalckx K, Cornelissen V, Buys R. Accuracy of Apple Watch Measurements for Heart Rate and Energy Expenditure in Patients With Cardiovascular Disease: Cross-Sectional Study. *JMIR Mhealth Uhealth*. Mar 19 2019;7(3):e11889. doi:10.2196/11889
7. Shcherbina A, Mattsson CM, Waggott D, et al. Accuracy in Wrist-Worn, Sensor-Based Measurements of Heart Rate and Energy Expenditure in a Diverse Cohort. *J Pers Med*. May 24 2017;7(2)doi:10.3390/jpm7020003
8. Duking P, Giessing L, Frenkel MO, Koehler K, Holmberg HC, Sperlich B. Wrist-Worn Wearables for Monitoring Heart Rate and Energy Expenditure While Sitting or Performing Light-to-Vigorous Physical Activity: Validation Study. *JMIR Mhealth Uhealth*. May 6 2020;8(5):e16716. doi:10.2196/16716
9. Hajj-Boutros G, Landry-Duval MA, Comtois AS, Gouspillou G, Karelis AD. Wrist-worn devices for the measurement of heart rate and energy expenditure: A validation study for the Apple Watch 6, Polar Vantage V and Fitbit Sense. *Eur J Sport Sci*. Jan 31 2022:1-13. doi:10.1080/17461391.2021.2023656
10. Kamal AA, Harness JB, Irving G, Mearns AJ. Skin photoplethysmography--a review. *Comput Methods Programs Biomed*. Apr 1989;28(4):257-69. doi:10.1016/0169-2607(89)90159-4
11. Wang R, Blackburn G, Desai M, et al. Accuracy of Wrist-Worn Heart Rate Monitors. *JAMA Cardiol*. Jan 1 2017;2(1):104-106. doi:10.1001/jamacardio.2016.3340
12. Gillinov S, Etiwy M, Wang R, et al. Variable Accuracy of Wearable Heart Rate Monitors during Aerobic Exercise. *Med Sci Sports Exerc*. Aug 2017;49(8):1697-1703. doi:10.1249/MSS.0000000000001284
13. Nuss KJ SJ, Archambault LJ, Schlemer EJ, Blake S, Courtney JB, Hulett NA, Li K. Accuracy of heart rate and energy expenditure estimations of wrist-worn and arm-worn Apple Watches. *Journal for the Measurement of Physical Behaviour*. 2018;2(3)doi:10.1123/jmpb.2018-0061
14. Wallen MP, Gomersall SR, Keating SE, Wisloff U, Coombes JS. Accuracy of Heart Rate Watches: Implications for Weight Management. *PLoS One*. 2016;11(5):e0154420. doi:10.1371/journal.pone.0154420
15. Thomson EA, Nuss K, Comstock A, et al. Heart rate measures from the Apple Watch, Fitbit Charge HR 2, and electrocardiogram across different exercise intensities. *J Sports Sci*. Jun 2019;37(12):1411-1419. doi:10.1080/02640414.2018.1560644
16. Kushhal A NS, Evans W, Gleadall-Siddall DO, Page R, O'Doherty AF, Carrol S, Ingle L, Abt G. Validity and Reliability of the Apple Watch for Measuring Heart Rate During Exercise. *Int J Sports Med*. 2017;38(14)doi:10.1055/s-0043-123585
17. Dooley EE DE, Golaszewski NM, Bartholomew JB. Estimating Accuracy at Exercise Intensities: A Comparative Study of Self-Monitoring Heart Rate and Physical Activity Wearable Devices. *JMIR Mhealth Uhealth* 2017;5(3)doi:10.2196/mhealth.7043
18. Bai Y HP, Mantis C, Welk GJ. . Comparative evaluation of heart rate-based monitors: Apple Watch vs Fitbit Charge HR. *J Sports Sci*. 2019;36(15)doi:10.1080/02640414.2017.1412235
19. Boudreaux BD HE, Hollander DB, Williams BM, Cormier CL, Naquin MR, Gillan WW, Gusew EE, Kraemer RR. Validity of Wearable Activity Monitors during Cycling and Resistance Exercise. *Med Sci Sports Exerc*. 2018;50(3)doi:10.1249/MSS.0000000000001471
20. Argent R, Hetherington-Rauth M, Stang J, et al. Recommendations for Determining the Validity of Consumer Wearables and Smartphones for the Estimation of Energy Expenditure: Expert Statement and Checklist of the INTERLIVE Network. *Sports Medicine*. 2022/08/01 2022;52(8):1817-1832. doi:10.1007/s40279-022-01665-4
21. Kwon S, Kim Y, Bai Y, Burns RD, Brusseau TA, Byun W. Validation of the Apple Watch for Estimating Moderate-to-Vigorous Physical Activity and Activity Energy Expenditure in School-Aged Children. *Sensors (Basel)*. Sep 25 2021;21(19)doi:10.3390/s21196413

22. Fuller D AJ, Simango B, Dorani F, Bozorgi A, Luan H, Basset F. Using machine learning methods to predict physical activity types with Apple Watch and Fitbit data using indirect calorimetry as the criterion. *Research Square*. 2020;doi:10.21203/rs.3.rs-17022/v
23. Pope ZC ZN, Li X, Liu W, Gao Z. Accuracy of commercially available smartwatches in assessing energy expenditure during rest and exercise. *JMPB*. 2016;2(2)doi:10.1123/jmpb.2018-0037
24. Nuss KJ, Thomson EA, Courtney JB, et al. Assessment of Accuracy of Overall Energy Expenditure Measurements for the Fitbit Charge HR 2 and Apple Watch. *Am J Health Behav*. May 1 2019;43(3):498-505. doi:10.5993/AJHB.43.3.5
25. Zhang P, Burns RD, Fu Y, Godin S, Byun W. Agreement between the Apple Series 1, LifeTrak Core C200, and Fitbit Charge HR with Indirect Calorimetry for Assessing Treadmill Energy Expenditure. *Int J Environ Res Public Health*. Oct 10 2019;16(20)doi:10.3390/ijerph16203812
26. Fuller D, Anaraki JR, Simango B, et al. Using machine learning methods to predict physical activity types with Apple Watch and Fitbit data using indirect calorimetry as the criterion. Research Square; 2020.
27. Schwartz J OP, Takito MY, Saunders B, Dolan E, Franchini E, Rhodes RE, Bredin SSD, Coelho JP, Dos Santos P, Mazzuco M, Warburton DER. Translation, Cultural Adaptation, and Reproducibility of the Physical Activity Readiness Questionnaire for Everyone (PAR-Q+): The Brazilian Portuguese Version. *Front Cardiovasc Med*. 2021;26(8)doi:10.3389/fcvm.2021.712696
28. McLester CN, Nickerson BS, Kliszczewicz BM, McLester JR. Reliability and Agreement of Various InBody Body Composition Analyzers as Compared to Dual-Energy X-Ray Absorptiometry in Healthy Men and Women. *J Clin Densitom*. Jul - Sep 2020;23(3):443-450. doi:10.1016/j.jocd.2018.10.008
29. Zhao H, Nishioka T, Okada J. Validity of using perceived exertion to assess muscle fatigue during resistance exercises. *PeerJ*. 2022;10:e13019. doi:10.7717/peerj.13019
30. Jensen K, Jorgensen S, Johansen L. A metabolic cart for measurement of oxygen uptake during human exercise using inspiratory flow rate. *Eur J Appl Physiol*. Jul 2002;87(3):202-6. doi:10.1007/s00421-002-0616-2
31. Bruce RA, Kusumi, F., & Hosmer, D. Maximal oxygen intake and nomographic assessment of functional aerobic impairment in cardiovascular disease. *Am Heart J*. 1973;85(4)
32. Hanson NJ, Scheadler CM, Lee TL, Neuenfeldt NC, Michael TJ, Miller MG. Modality determines VO2max achieved in self-paced exercise tests: validation with the Bruce protocol. *Eur J Appl Physiol*. Jul 2016;116(7):1313-9. doi:10.1007/s00421-016-3384-0
33. Gilgen-Ammann R, Schweizer T, Wyss T. RR interval signal quality of a heart rate monitor and an ECG Holter at rest and during exercise. *Eur J Appl Physiol*. Jul 2019;119(7):1525-1532. doi:10.1007/s00421-019-04142-5
34. Germini F, Noronha N, Borg Debono V, et al. Accuracy and Acceptability of Wrist-Wearable Activity-Tracking Devices: Systematic Review of the Literature. *J Med Internet Res*. Jan 21 2022;24(1):e30791. doi:10.2196/30791