COMPARISON OF THE ACCURACY OF THREE WRIST-WORN TRACKING DEVICES WHEN MEASURING HEART RATE DURING AEROBIC EXERCISE

Jialin Ding, Wuyue Hong, Hanjun Li,

Beijing Sport University, Beijing, China

The study aimed to assess the accuracy of the three wrist-worn heart rate monitors compared with a chest strap monitor at different aerobic exercise levels. Ten college students completed a three-level incremental load training on the elliptical machine with different brands of wrist-worn tracking devices including the Apple Iwatch7, Huawei Runner and Garmin 945. Polar H10 was used to measure HR as the gold standard. Huawei Runner showed the best agreement with Polar H10 (r = 0.981) followed by the Apple Iwatch (r = 0.973), the Garmin 945 was the most imprecise device (r= 0.866) when exercising on elliptical trainer. The paired relative error showed the best correlation between the Apple Iwatch7 and polar H10 both during exercise at 0.0 (± 0.5) and during resting at -0.1 (± 0.4). The mean output value ratio of Apple Iwatch7 was 78.10% (± 13.30%), while there wasn't the missing HR data in the Huawei and Garmin devices. There was no significant difference between left and right hands. The Apple and Huawei devices measured HR with acceptable accuracy during aerobic exercise. While the proportion of missing HR data of Apple Iwatch is more than 20% during exercise in order to ensure the accuracy of measurement by deleting potentially problematic data automatically. This problem needs to be considered for high-precision continuous monitoring scenarios.

KEYWORDS: wrist-worn heart rate monitor, accuracy, chest strap monitor

INTRODUCTION: Over the last decade, there has been a proliferation of commercially available HR monitors and wearable fitness devices (Gillinov et al., 2017). Consumer-based wrist-worn multisensor activity monitors have emerged as an increasingly popular way to track various physiological metrics such as heart rate (HR) and physical activity levels. Many consumers purchase these wearable fitness trackers to record their HR response to exercise intensity. It can help us strengthen cardiorespiratory endurance in the process of gradually increasing exercise intensity, and finally improve health (Case et al., 2015; Diaz et al., 2015). Therefore, the accuracy of HR monitoring is of great importance. Incorrect HR readings may prompt a consumer to increase or decrease exercise intensity, which can have health ramifications and negatively affect exercise performance (Shcherbina et al., 2019). Unfortunately, there appears to be an evidence-based research lag with respect to reporting the accuracy of commercially available devices. Wrist-worn heart rate monitors such as the Apple Watch are becoming more integrated in healthcare. However, validation studies of these consumer devices remain scarce (Falter et al., 2019). In some cases, by the time a study has gained approval, participants have been tested, data analyzed, and reports have been written and gone through the peer review process, a wearable technology device has been updated to the next model or has become obsolete (Huang et al., 2016). Broad and latest assessment of the monitors' accuracy is important both for the individuals who rely on these monitors to guide their athletic, physical, and rehabilitative activity and for the physicians to whom these individuals report their HR readings for the purpose of potentially guiding therapy (Gillinov et al., 2017).

Therefore, the objective of this study was to assess the accuracy of three commonly used, currently commercially available, optically based wearable HR monitors in an incremental powered study under aerobic exercise conditions. We hope that this paper can be a resource for both researchers as well as personal consumers wishing to utilize wearable technology devices for physical activity monitoring.

METHODS: Ten college students from Beijing Sports University took part in this study (5 males, 5 females, age 23.4 ± 0.8years, height 173 ± 7 cm, weight 62 ± 8 kg). Participants were recreationally active and none had tattoos on their wrists. All patients gave written informed consent to participate in the study.

Each participant underwent HR monitoring with an electrocardiographic chest strap monitor (Polar H10) and two different wrist-worn HR monitors every trial. The polar H10 (Polar Electro Inc., Bethpage, NY) was secured tightly on the subjects' chest to ensure contact with skin. Data from the monitor served as the reference standard for HR (beats per minute, bpm). And there were three different wrist-worn devices including the Apple Iwatch7, Huawei Runner and Garmin 945 in total. Of most importance was to ensure that the devices were properly worn and avoid simultaneously carrying multiple devices on one arm. It was usually 1 to 2 fingers (or 1.5-2.5cm) below the carpal bone and still kept it tightly on the wrist. They needed to wear these devices randomly and exchange the two monitors in left and right hands after each trial. And then just use the remaining device to measure the left and right hands respectively.

The HR data was measured when resting and performing on elliptical trainer with arm levers at three varying intensities. Because there were no standard workload settings for elliptical trainers, we identified three settings that were judged to represent light, moderate, and vigorous activity. The trial protocol was as follows:

Light for 3 min: resistance = 8, cadence =80 min-1; Moderate for 3 min: resistance = 8, cadence = 90min-1; Vigorous for 3 min: resistance = 8, cadence = 100 min-1.

This protocol was for female, while the resistance for male was set 10 and the others remained the same. Each subject spent 9 min at one trial and exchanged devices left and right for the next group of trial after a 2-min rest. And then continued to test the remaining device in the same process. For rest, the three devices were recorded the heart rate in quiet state for 3 min and it still needed to collect data on left and right hands. HR signals for all devices were checked at the beginning of each trial segment to ensure device function.

Sample size was based on the use of Pearson Correlation Coefficient (r) to compare HR measurements with wearable, optically based HR monitors to those obtained with the electrocardiographic chest strap monitor (polar H10). On the basis of previous work (Derrick, 1994), we deemed a $r > 0.8$ to represent acceptable accuracy in HR measurement. Each of the HR monitoring devices was assessed for accuracy by calculating the difference between the measured value and the comparative value. The paired absolute error was calculated as the mean of the device measurement - reference standard × 100/reference standard, while the paired relative error was calculated as the mean of the absolute value of device measurement-reference standard × 100/reference standard. We also selected the indicators of 5bmp and 10bmp accuracy, which was calculated as the number of the absolute value of the device measurement - reference standard divided by the total number when the absolute value was greater than 5 and 10 respectively. The proportion of missing HR data was calculated as the number of missing values divided by the number of total values, and the output value ratio was calculated as 1- the proportion of missing data. The paired sample T test was used to verify the difference between left and right hands.

RESULTS: Measured HR ranged from 53 to 190bpm. The output value ratio of Huawei and Garmin device was basically consistent with polar H10. There were 12 pairs of values in one minute. However, the Apple Iwatch7 had three missing values and the mean output value ratio was 78.10% (± 13.30%), while this was not the case with the other two devices whose mean output value ratio is both up to 100%. There was a gap of less than 10bmp between polar and the Huawei as well as Apple HR monitors under exercise conditions but extended to nearly 20bpm for some values of Garmin 945.

The correlation between the HR values on the Huawei Runner and our gold standard chest band sensor was the best during exercise and rest (exercise: 0.981, rest: 0.888), whereas the paired relative error between the Apple Iwatch7 and polar H10 showed to be the significantly lowest at 0.0 (± 0.5) during exercise and at -0.1 (± 0.4) during resting. Mean HR accuracy across the activity and the rest was analyzed and compared with the reference standard. These values were shown in Table 1 and Table 2 respectively.

There was no significant difference between left and right hands. The P values of the left and right hands of Apple, Huawei and Garmin devices were 0.346, 0.656 and 0.191 when exercising, while at rest the P values were 0.476, 0.584 and 0.366 respectively.

Table 1: HR monitor differences from Polar H10 when exercising on elliptical trainer

Brand	HR (bpm) Differences from Polar								
	Paired Relative Error		Paired Absolute Error		r	5bmp Accuracy	10bmp Accuracy	the output value ratio	
	Mean	SD	Mean	SD				Mean	SD
Apple Iwatch7	0.0	0.5	1.8	1.4	0.973	89.4%	97.9%	78.10%	13.30%
Huawei Runner	-2.0	2.0	2.8	1.9	0.981	85.4%	93.1%	100.0%	0.00%
Garmin 945	-3.2	10.9	7.0	10.4	0.866	71.7%	80.7%	100.0%	0.00%

Table 2: HR monitor differences from Polar H10 when resting

Brand	HR (bpm) Differences from Polar								
	Paired Relative Error		Paired Absolute Error		r	5bmp Accuracy	10bmp Accuracy	the output value ratio	
	Mean	SD	Mean	SD				Mean	SD
Apple Iwatch7	-0.1	0.4	1.5	0.8	0.828	92.0%	98.8%	85.70%	9.30%
Huawei Runner	-0.1	0.9	1.6	1.0	0.888	94.4%	98.9%	100.0%	0.00%
Garmin 945	-0.4	1.2	2.5	1.0	0.607	84.4%	94.7%	100.0%	0.00%

DISCUSSION: The results of this study demonstrated that optically based wearable HR monitors have different accuracy than electrode containing chest strap monitors (Polar H10). Similar to previous studies (Bunn, 2018), most wrist-worn HR monitors underestimated the HR compared with a polar chest strap. The Apple device showed the best perform in heart rate accuracy during exercise, while the Garmin945 performed the worst. The Huawei Runner and Apple Iwatch showed close paired relative and absolute error when resting. The majority of error was from devices failing to track during dynamic activities.

The new wrist-worn HR monitors do not measure cardiac electrical activity, rather, they rely on photoplethysmography. The monitor illuminates the skin with an LED and then measures the amount of light reflected back to a photodiode sensor, this enables detect the variations of blood volume associated with the pulse of blood caused by each cardiac contraction. Potential sources of error with optically based monitors include motion artifact from physical movement, misalignment between the skin and the optical sensor, variations in skin color, ambient light, and poor tissue perfusion (Alzahrani et al., 2015). The accuracy of such monitor during exercise is controversial, some studies suggesting that wrist-worn HR monitors perform the best at rest or slow walking, and others asserting good accuracy even during vigorous exercise (Spierer et al., 2015). Extending that work, we assessed the performance of wearable HR monitors using aerobic exercise modalities (elliptical trainer with arms) at different levels of intensity.

The accuracy could be accepted when the activity mode was set. Prior research had suggested that photoplethysmography sensors used to measure the HR were liable to poor accuracy when there was repetitive wrist motion and any activity was at a high intensity. However, we found reasonable accuracy in HR estimations for these three devices under the elliptical machine when the device was in the activity mode setting. It appeared that the monitors used different HR measurement algorithms depending on the activity mode selected. It might be that the activity mode algorithms implement less smoothing than the nonactivity mode algorithm and frequency to collect data was higher, thereby designed to respond faster to rapid HR changes.

We found the Apple Iwatch had some missing data which were attributable to failure to record HR sometimes. Therefore, raw data intercepted by time period were extracted for every participant in order to check the proportion of missing HR data. Three participants had serious missing data in the Apple Iwatch in their left hands, and the output value ratio was less than 80%. However, we speculated that it falsely increased the accuracy of the devices by removing some potentially problematic data. The other two devices Apple and Garmin with the 100% output value ratio may have resulted in a relatively poor accuracy.

The three monitors were chosen because of their apparent popularity with the public, and each monitor was the manufacturer's most recent offering at the time of the study. However, the devices were merely assessed in young, healthy volunteers exercising in a laboratory setting. Therefore, the results might vary for different characteristic of individuals. In addition, these results might not be representative of those obtained during more vigorous exercises or different kinds of activities (e.g., running on pavement, swimming, or other sports participation). It would be important to do further investigations in more different settings to corroborate our results. Another limitation was that the HR measurements from the wearable devices were not compared against a true gold standard such as ECG.

CONCLUSION: This study demonstrates that optically based wrist-worn HR monitors vary in their accuracy whether at rest or during aerobic exercise. The Apple and Huawei devices were within an acceptable error range in measuring HR except for the Garmin. Individuals who use such monitors should be aware of the possibility of inaccurate measurements and chose suitable products for themselves according to different needs. The proportion of missing HR data of Apple Iwatch is more than 20% during exercise in order to ensure the accuracy of measurement by deleting potentially problematic data automatically. This problem needs to be considered for high-precision continuous monitoring scenarios.

REFERENCES

Gillinov S, Etiwy M, Wang R et al., 2017. Variable Accuracy of Wearable Heart Rate Monitors during Aerobic exercise [J]. Med Sci Sports Exerc. 49(8):1697-1703.

Case MA, Burwick HA, Volpp KG, Patel MS., 2015. Accuracy of smartphone applications and wearable devices for tracking physical activity data. JAMA. 313(6):625 – 6.

Diaz KM, Krupka DJ, Chang MJ, et al., 2015. Fitbit: an accurate and reliable device for wireless physical activity tracking. Int J Cardiol. 185:138 – 40.

Shcherbina, Anna, Mattsson, C., Mikael, & Waggott, et al., 2017. Accuracy in wrist-worn, sensor-based measurements of heart rate and energy expenditure in a diverse cohort. Journal of Personalized Medicine.

Falter Maarten, Budts Werner, Goetschalckx Kaatje, Cornelissen Véronique, Buys Roselien., 2019. Accuracy of Apple Watch Measurements for Heart Rate and Energy Expenditure in Patients with Cardiovascular Disease: Cross-Sectional Study. [J]. 7(3):

Huang Y, Xu J, Yu B, Shull PB. Validity of FitBit., 2016. Jawbone UP, Nike+ and other wearable devices for level and stair walking. Gait Posture 48:36-41

Derrick, TR, Bates, BT & Dufek, JS., 1994. Evaluation of time-series data sets using the Pearson product-moment correlation coefficient. Medicine & Science in Sports & Exercise 26 (7): 919–28.

Bunn, J. A, Navalta, J. W, Fountaine, C. J, & Reece, J. D., 2018. Current state of commercial wearable technology in physical activity monitoring 2015-2017. International Journal of Exercise Science, 11(7), 503-515.

Alzahrani A, Hu S, Azorin-Peris V, et al., 2015. A multi-channel optoelectronic sensor to accurately monitor heart rate against motion artefact during exercise. Sensors (Basel). 15(10):25681 – 702.

Spierer DK, Rosen Z, Litman LL, Fujii K., 2015. Validation of photoplethysmography as a method to detect heart rate during rest and exercise. J Med Eng Technol. 39(5):264-271.