International Journal of EXERCISE SCIENCE

Original Research

Evaluation of Earbud and Wristwatch Heart Rate Monitors during Aerobic and Resistance Training

JENNIFER BUNN‡[1], ELIZABETH WELLS†[1], JOHN MANOR†[1], and MICHAEL WEBSTER‡[2]

[1]Department of Physical Therapy, Campbell University, Buies Creek, NC, USA; [2]Department of Kinesiology and Physical Education, Valdosta State University, Valdosta, GA, USA.

†Denotes graduate student author, ‡Denotes professional author

ABSTRACT

International Journal of Exercise Science 12(4): 374-384, 2019. Assessment of biometrics during exercise is evolving to create devices that are "all-inclusive", in an effort to decrease the number of devices required during exercise while providing comprehensive and accurate biometric measures. The purpose of this study was to determine the accuracy of two optical heart rate monitors, the Jabra earbud and the Mio Alpha wristwatch, during aerobic and anaerobic exercise. Twenty-two recreationally active participants (25.4 ± 6.9 years, 171 ± 11 cm, 73.9 ± 3.1 kg, and 25.2 ± 9.2% body fat) completed this study. Participants completed 30 minutes of treadmill activity, 25 minutes of high-intensity interval exercise (HIT), and 40 minutes of continuous outdoor activity of their choice, walking or running. Three heart rate (HR) monitors, (Polar chest strap, Mio Alpha, Jabra earbud) were worn during all exercises, with the Polar chest strap serving as the benchmark. HR was assessed in one-second intervals. Analyses included mean bias, mean absolute percent error (MAPE), and Lin's concordance coefficient. Overall, the Mio Alpha had a MAPE of 5.73 ± 10.19% and a moderate correlation with the benchmark, r(c) = 0.771, performing better in the treadmill and outdoor conditions. The Jabra earbud had a MAPE of 3.14 ± 6.13%, and a high correlation with the benchmark, r(c) = 0.939, performing well in all three conditions. Placing a HR monitor in an earbud is a viable option for obtaining an accurate HR assessment during different types of exercise. The accuracy of the Mio Alpha was likely affected by wrist movements during the HIT training.

KEY WORDS: Pulse, photoplethysmography, wearable technology, activity tracker, accelerometer

INTRODUCTION

Activity trackers are an evolving technology that allows the consumer to collect and monitor physical activity data such as steps, caloric expenditure, sleep, and heart rate (HR) (8). Popularity of these devices have risen exponentially in the past decade with advances in technology and health initiatives that promote physical activity (13, 18,20). The largest distinction in choosing activity tracker devices is in their purpose: is it being used for recreation or research? Indoor or outdoor conditions? Research employing HR assessment during activity is more likely to use a chest strap that transmits data to a specialized watch (10, 14, 26, 30). These chest strap units are accurate, as they measure cardiac electrical activity (12), but may be viewed as cumbersome because they include both a chest strap and watch, and the strap must be worn under clothing at the level of the inferior sternum. Recreational consumers are more apt to use lightweight and multi-functional devices during physical activity that are worn as a clip, placed in the shoe, as a wrist band, or in ear buds (30, 31). Improvements in technology have allowed devices to be worn in day-to-day life and the seamless nature of collecting biometric data from such devices is of value to both research (28) and consumer health applications. Van Remoortel et al, suggest caution when choosing activity monitors included in clinical trials on chronic disease populations due to lacking validation of activity monitors and the specific function capabilities (28). In consumer health, activity monitors are being utilized as a medium to bridge the gap between a sedentary lifestyle and a useful behavior change technique, which allows the consumer the ability to monitor personal health and encourage continued activity through the use of interactive data via computer or mobile app (17).

Many wearables measuring HR utilize photoplethysmography (PPG), which is a light-emitting diode (LED) that illuminates skin and related capillaries, and measures the reflected light to a photodiode sensor. This enables measurement of alterations in blood volume associated with each cardiac systole (10) and calculates pulse via proprietary algorithms during rest and motion. Additionally, accelerometers and GPS are utilized separately or in combination for other metrics related to distance covered, steps taken, caloric expenditure, and exercise intensity. The ideal physical activity tracker provides valid and reliable data in both laboratory and free-living conditions. The desire for accurate biometric data from wearables is a need for both consumer (8, 9, 12) and research applications (28). Research and industry have suggested accuracy fall within 1% in laboratory conditions and up to 10% mean absolute percent error (MAPE) in free-living conditions for steps/day and HR (4, 5, 6, 9, 25). Examination of the accuracy of step counting of consumer devices (e.g. Fitbit, Apple, Garmin) have been mixed, with variations in correlations, mean bias, and MAPE depending on the devices under test (DUT), the benchmark used, placement of the device, and the protocol (1, 4, 7, 9, 11, 13, 19, 23). In a similar fashion to step accuracy, criteria researchers have proposed HR validity criteria includes $r > 0.90$, a mean bias of ≤ 3 bpm, and standard error ≤ 5 bpm (6, 7, 26). Comparison to either an electrocardiogram (ECG) or a previously validated chest strap device (e.g. Polar) are often viewed as acceptable benchmark comparisons (6, 10, 12, 14, 26, 30). Gillinov et al, and Terbizan et al, compared wearable HR monitors during aerobic conditions and demonstrated high correlation of Polar chest strap to ECG (10, 26). Gilinov et al, demonstrated that the Polar chest strap had the best agreement with ECG [r(c) = 0.996] compared to the Apple Watch [r(c) = 0.92], TomTom Spark [r(c) = 0.83], Garmin Forerunner [r(c)=0.81], Scosche Rhythm+ [r(c) = 0.75], and Fitbit Blaze [r(c) = 0.76] during treadmill, stationary bike and elliptical conditions for low to moderate activity (10). Terbizan et al, found the Polar Accurex II and Polar Vantage XL demonstrated high accuracy at rest and during moderate activity ($r > 0.90$; standard error of estimate (SEE) 5 beats/min) (26). Performance of such devices have been validated, and previous literature has demonstrated that, during low intensity cycling, walking, jogging, running, arm-raises, lunges, and isometric plank, the Fitbit Charge HR and Basis Peak had moderate to strong correlation, respectively, to HR measured by ECG (12). Gillinov et al. (10), determined that among popular commercial activity trackers the Apple iWatch demonstrated agreement with ECG during aerobic exercises, including: treadmill, elliptical with and without arms, and cycling. However, this study removed poorly recorded data from the DUTs, which may have provided a more positive result for the devices. Specifically, most of the data removed was during exercises that involved significant arm swing actions, similar to that found during running.

The popularity of activity trackers is evident from the prevalent use in today's society and the continual progression of tracker models to improve accuracy and convenience for the end user (8, 12, 20). In addition to individual consumer use, several health care companies have incorporated wearable use as encouragement for increased physical activity in hopes of reducing healthcare costs and the burden of chronic disease. Health conscious users as well as individuals with chronic disease are able to monitor health and exercise using activity trackers, leading to an increased exercise adherence and stronger commitment to one's health (12, 19). The continuous demand and utilization of activity trackers, among both researchers and recreational users, indicates the need for validation and reliability studies. Additionally, correlation of these data during various modes of exercise (e.g. aerobic with and without use of arms and resistance training) under a variety of conditions (e.g. environmental and physical characteristics of participants) is warranted due to the multi-purpose use of activity trackers in all facets of exercise (8, 10).

The purpose of this study was to determine the accuracy of the Jabra ear bud HR monitor and the Mio Alpha wrist watch, during treadmill, strength training, and outdoor aerobic exercises. These devices utilize "all inclusive" technology designed to optimize biometric measurements during exercise while minimizing hardware. The Mio Alpha (Vancouver, BC Canada) is a PPG device worn on the wrist that functions as a wrist watch with a built-in activity tracker. The watch has a snug, comfortable fit with a 5.3"-7.5" wrist band, to ensure accurate biometric measurements and is designed to be sweat-proof and water proof up to 30 meters. Previous literature evaluating the Mio Alpha suggests variation in accuracy (r = .55-.93) depending upon the type of exercise performed and the frequency of data collected (21, 24, 25, 29,30). The Jabra Pulse (Copenhagen, Denmark) earbud is a PPG device featuring multi-functional capabilities of a hands-free activity tracker and music player. The earbuds have a tight, comfortable fit inside of the ear canal, with sizing available from small to large for earwings, eargels, and foamtips, which decreases the rate of erroneous accessory motions recorded during activity. Earbuds have previously been evaluated for accuracy in HR measurement (15,22), but at the time of this writing, this is the first study to evaluate the Jabra earbuds specifically. The designs of both activity trackers were to improve accuracy of physical activity tracking in a single convenient, hands-free Bluetooth device.

METHODS

Participants

The subject population consisted of a convenience sample of 25 subjects, ranging in age from 18-65 years, and considered to be recreationally active. All subjects were considered low risk for cardiovascular disease according to the American College of Sports Medicine standards. Participants were excluded if they placed in the moderate to high risk for cardiovascular disease or if they were unable to complete the test. Prior to data collection, subjects completed IRB university-approved documents. This study was conducted in accordance with the Declaration of Helsinki (32).

Height and weight were measured using standard procedures on a Health-O-Meter 402KLS (Sunbeam-Oster, Boca Raton, FL). Body composition was measured via bioelectrical impedance analysis (BIA; InBody 570, Biospace, Cerritos, CA). A total of three HR monitors were placed on each subject during data collection. The monitors included the Polar RS800CX (Polar Electro Inc., Lake Success, NY), the Mio Alpha, and the Jabra Pulse earbud. The Polar PS800CX was placed on the chest below the sternum and was used as the standard of measure. The Mio Alpha was placed on the left wrist, and the Jabra Pulse was placed in both ears. The Mio Alpha and the Jabra Pulse both transmitted information via Bluetooth low energy to an Apple iTouch 5 (Apple, Cupertino, CA), utilizing the Wahoo Fitness (Atlanta, GA) application. Second-by-second HR data was obtained for each device through the Wahoo Fitness app. The exercise was comprised of three separate protocols: 1) treadmill, 2) high-intensity intermittent (HIT) exercise, and 3) outdoor exercise. The devices were all started simultaneously, and HR was collected each second during the exercise protocols from all three devices.

Table 1. 30-minute treadmill protocol

Time (min)	Activity
0-0:30	Stand
0:30-1:15	3.4 mph Walk
1:15-2:00	2.2 mph Walk
2:00-3:30	Self-selected run speed (generally between 6 and 9 mph)
3:30-5:00	3.0 mph Walk
5:00-6:00	Self-selected run speed (generally between 6 and 9 mph, but greater than the 2:00-3:30 run)
6:00-6:20	2.2 mph Walk
6:20-8:00	Stand
8:00-18:00	64-74% age-predicted maximal heart rate
18:00-23:00	75-95% age-predicted maximal heart rate
23:00-28:00	2.5 mph
28:00-30:00	Passive cool down

Protocol

The treadmill exercise protocol consisted of an 8-minute dynamic exercise, followed by two ten-minute steady-state segments, followed by a 2-minute passive cool down equaling 30 minutes of activity. The details of the treadmill protocol are listed in Table 1. The HIT exercise protocol, shown in Table 2, consisted of 25 minutes total with two rounds of two minutes of each the following activities: elliptical trainer, burpees, jump rope, dumbbell thrusters, and sit-ups. Participants were given 30 seconds of recovery between each activity. The outdoor exercise protocol consisted of approximately 40 minutes of continuous activity of their choice, walking or running. Some participants completed all three exercise protocols in a single testing session, while others completed the exercise protocols in two testing sessions. The accuracy and correlation of HR between devices was not affected by participants who completed the protocols over two testing session because the research study monitored correlation between devices, not the impact of exercise intensity to HR rise to steady-state.

Table 2. 25-minute HIT protocol

Time (min) – Round 1	Time (min) – Round 2	Activity
0-2:00	12:30-14:30	Elliptical trainer (self-paced)
2:00-2:30	14:30-15:00	Rest
2:30-4:30	15:00-17:00	Burpees
4:30-5:00	17:30-18:00	Rest
5:00-7:00	18:30-20:30	Jump rope
7:00-7:30	20:30-21:00	Rest
7:30-9:30	21:00-23:00	Dumbbell thrusters
9:30-10:00	23:00-23:30	Rest
10:00-12:00	23:30-24:30	Sit-ups
12:00-12:30	24:30-25:00	Rest

Statistical Analysis

Heart rate was assessed each second during all three exercise protocols. It was anticipated that there will be data lost due to poor connection between the benchmark device and the smart device, as well as technical errors associated with the devices. Data related to these issues were removed from the data set and not included in the analyses. Data were analyzed using Microsoft Excel (Los Angeles, CA). The Jabra Pulse ear bud and the Mio Alpha were only compared with the benchmark for the separate treadmill, HIT, and outdoor protocols, as well as in combination which is consistent with previous literature (10, 25, 31). Each analysis between the DUTs and the benchmark included mean bias, Lin's concordance, Bland-Altman plots, MAPE, and data distribution assessment. Limits of agreement (LoA) were calculated (mean bias ± SD * 1.96) for each Bland-Altman assessment. Lin's concordance was chosen because it assesses concordance between continuous data (15), and was consistent with previous literature. Determinants of agreement for Lin's Concordance was as follows: >0.99 almost perfect, 0.95-0.99 substantial, 0.90-0.95 moderate, and < 0.90 poor (15).

RESULTS

Participant demographics were 25.4 ± 6.9 years, 171 ± 11 cm, 73.9 ± 3.1 kg, and 25.2 ± 9.2% body fat. Of the 25 participants that completed the study, data were only analyzed for 22 (16 females, 9 males). Three participants' data were removed from analysis due to technical errors with the ear buds and/or Mio Alpha watch. These errors included improper programming, technician errors, and complete Bluetooth disconnection. Post-processing of all of the data revealed some incidences of incomplete Bluetooth connectivity issues and general artifact in the benchmark device. Approximately 5% of the data was removed to provide clean and usable benchmark data. There were approximately 125,400 total data points in this study, and this process removed roughly 5% of the data, totaling approximately 6,270 data points removed.

Table 3 shows a comparison between the Mio Alpha watch and the Jabra ear bud with the benchmark for all three exercise types individually and in combination. The Mio Alpha tended to underestimate HR for each of the exercise conditions, whereas the Jabra Pulse underestimated HR only during the HIT exercise. Both devices tested well in the treadmill and outdoor running scenarios, but agreement with the benchmark device was lower during the HIT exercise for both devices. Specifically, the Mio Alpha performed poorly in the HIT exercise with an absolute percent difference greater than 10%, and a low correlation with the benchmark. The Jabra Pulse earbud showed poor concordance during HIT exercise with r(c) = 0.861, and the HIT activity produced worse performance in the earbud than the other two activity conditions. For the combined data, the Jabra earbud showed substantial concordance, r(c) = 0.939, and the Mio Alpha showed poor concordance, r(c) = 0.771.

Table 3. Comparison of Mio Alpha and Jabra Pulse HR assessment to Polar chest strap for all three exercise protocols individually, and with all data combined. * indicates a MAPE > 10%, mean bias > 3 bpm, or r(c) < 0.90.

		Bias (bpm)	% of Data within ± 5 bpm	MAPE (SD)	Agreement r(c) (95% CI)
Treadmill	Mio	-0.3 (10.9)	72	4.80 (8.33)	.911 (.909, .912)
	Jabra	0.8 (7.5)	85	2.48 (4.90)	.943 (.942, .944)
HIT	Mio	-17.7 (27.6)*	44	12.50 (15.32)*	.329 (.320, .337)*
	Jabra	-3.6 (11.7)*	80	3.53 (7.00)	.861 (.858, .864)*
Outdoor	Mio	-1.9 (8.7)	82	2.83 (5.25)	.924 (.923, .925)
	Jabra	0.8 (8.5)	79	3.64 (6.69)	.953 (.952, .954)
Combined	Mio	-4.9 (17.2)*	70	5.73 (10.19)	.771 (.769, .774)*
	Jabra	-0.3 (9.2)	82	3.14 (6.13)	.939 (.938, .940)

Figure 1 shows the Bland-Altman plots for each of the three exercise protocols for the Mio and Jabra HR monitors. Figure 1b shows the large variation in data obtained from the Mio Alpha during the HIT exercise. The performance of the Jabra earbud was consistent throughout all three exercise protocols, with slighter larger LoA during the HIT exercise session.

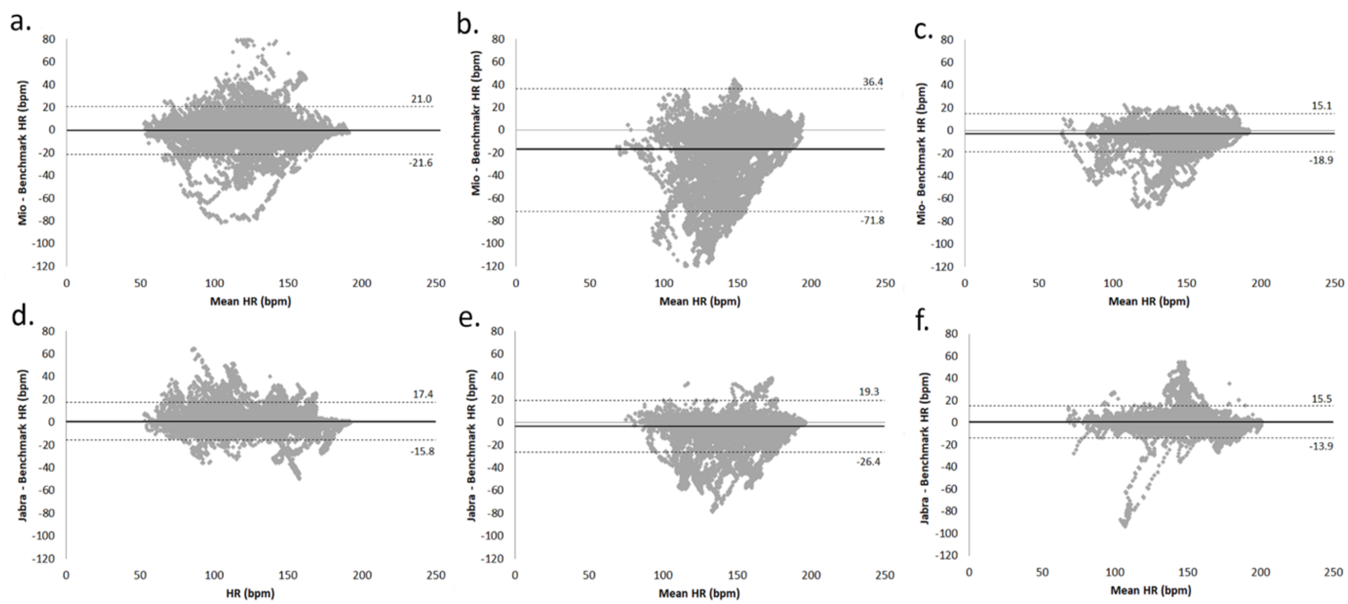

Figure 1. Bland-Altman plots of the Mio Alpha during the a) treadmill, b) HIT, and C) outdoor exercise sessions. Bland-Altman plots of the Jabra Pulse during the d) treadmill, e) HIT, and f) outdoor exercise sessions. Mean bias is represented by the dark line and limits of agreement are shown by the dashed lines.

DISCUSSION

When addressing accuracy of wearable devices, it is important to address standards for correlation, mean bias, and MAPE. The technology industry and research indicate that correlation between a DUT and the benchmark should be ≥0.90, the MAPE should be ≤10%, and mean bias should be ≤ 3 bpm (5, 6, 27). The Mio Alpha met each of these accuracy standards for the treadmill and outdoor protocols, but not for the HIT protocol. The combined Mio Alpha data met the MAPE and mean bias standard, but not the correlation standard. The Jabra Pulse earbud met each of the three standard for the treadmill and outdoor protocols, as well as with the combined protocols. But, like the Mio Alpha, the Jabra Pulse earbud did not meet the accuracy standards with the HIT exercise protocol.

The results of the present study are similar to previous research, showing good accuracy of the Mio Alpha during treadmill and running/walking exercise (24, 30) and poor accuracy during weight lifting (24). The Mio Alpha performed better during the walking and running trials in the present study than it did in the study conducted by Stahl et al. (25). When comparing these studies, it is important to note that Stahl recorded HR every minute, Wang recorded HR every three minutes, and Spierer collected HR every five seconds. Additionally, Stahl's participants wore several wrist-based HR monitors on the same arm simultaneously. While Stahl did note that the fit of the DUTs were still in accordance with the manual specifications, this may have affected the performance of the HR monitors. The differences in timing of HR collection are substantial, especially with changes in exercise intensity, which cause HR to change quickly and then steady off. If a DUT is slow to detect this change in real-time, then this error is unassessed by the methods used by Wang and Stahl. The present study allowed for capture of these oscillations in real time, and indicates if a DUT lags behind or under/overestimates these changes. Stahl's and Wang's studies did not capture these alterations in real time, but rather in stages after HR reached steady state. This effect is illustrated in Figure 2. This figure is an example of one participant's data during the treadmill exercise. This figure clearly shows differences in HR that occurred between the benchmark and the Mio Alpha within the first minute of exercise, and also between the benchmark and the Jabra Pulse between 1153 and 1201 seconds. When capturing data in 1-minute or 3-minute increments, both of these discrepancies would have been missed, likely resulting in a better performance for the DUT.

The Jabra Pulse met research and industry standards for the treadmill and outdoor sessions, as well as when all data were combined. This performance is in line with that of previous literature exploring the potential of a HR monitor placed in an earbud (14, 22). The Jabra Pulse did not perform as well during the HIT exercise session as the other exercise sessions, but was not prone to the same level of error as the Mio Alpha. This is likely because exercises involving the arms, such as the elliptical trainer, burpees, and dumbbell thrusters, cause more movement in the wrist and forearm, which may result in more movement of, and error in, wrist-based HR devices. Studies have shown this same problem to occur in the Mio Alpha (24), as well as other devices (10, 12). As such, the Jabra Pulse may have an optimal location to assess HR in these types of activities.

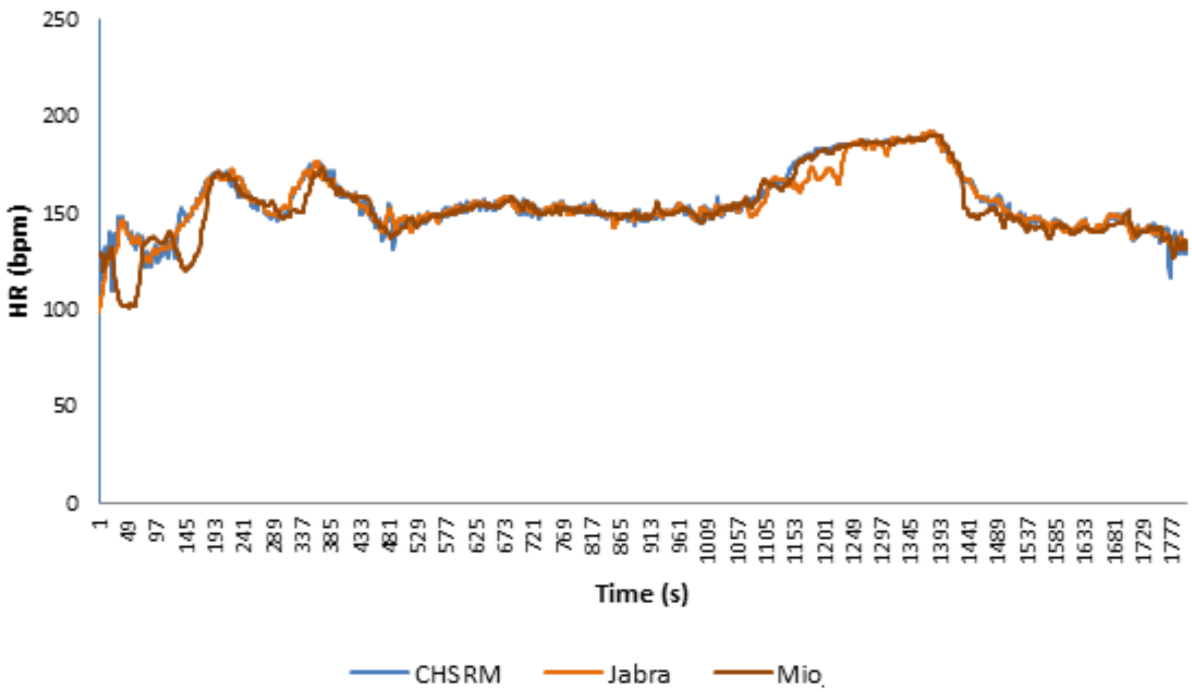

Figure 2. Second-by-second HR data for one participant during the treadmill exercise. Data shown include the benchmark chest strap (CHSRM), the Jabra Pulse, and the Mio Alpha.

Limitations found during the study included: fit of the ear bud device, Bluetooth connectivity, and reliable chest strap recordings. The fit of the ear bud was a potential issue with the Jabra device. The Jabra Pulse earbud is accompanied with several options for both the gel and the ear tip in order to achieve a good fit. We used the 3-mm gel with a small fin and the small tip with most of the participants, and made adjustments to the ear bud gels as necessary. However, in some cases it wasn't discovered that an adjustment was necessary until testing had already begun, or perhaps after the participant had been exercising for a while and their physiology had changed (sweaty skin). Adjustments were made as necessary to ensure a good fit. Fit was the most problematic during the HIT exercise protocol. This is likely due to the ballistic nature of the movements chosen.

Another issue observed during data collection was the loss of Bluetooth connectivity that occurred between all of the HR monitors and the iTouch recording device using the Wahoo Fitness application. These disconnects occurred most frequently during the outdoor running trial, but were also experienced during both indoor trials. Music was not used during any of the testing sessions, but it would be useful to assess if Bluetooth connectivity is affected, either positively or negatively, by a constant stream of music in the Jabra Pulse.

Post-processing of heart rate for all three devices was necessary when contact loss was apparent or the chest strap reported heart rates outside of the expected physiological range. Issues related to the chest strap use included general artifact that occurs during exercise. The Mio Alpha and Jabra Pulse were similarly affected, as three participants' data were removed from analyses in all three exercise conditions due to technical errors. Outside of these participants, approximately

5% of the data was lost to other technical issues related to poor connectivity and general device artifact. These issues are not unusual with wearable technology assessment and have also been noted in previous literature (7, 10, 22, 30).

Both the Jabra Pulse and Mio Alpha performed well during the lab-controlled and ambulatory running and walking trials, but exhibited poorer performance during resistance training work, and exercise using substantial arm motion. Further research is necessary to explore why differences exist with both devices during these exercise conditions, particularly with the use of wrist-based devices, as that is the most common location for consumer-devices. Future research is also warranted using similar protocols with oscillating HR outputs, different environments and activities, and second-by-second measurements. This type of research will benefit the efficacy of optical heart rate monitors to provide accurate data utilized from activity monitors in controlled research settings (17). Furthermore, research on activity monitors will improve the ability and accuracy of the lay person to monitor personal cardiovascular health during recreational activity (9, 12, 17).

ACKNOWLEDGEMENTS

This study was funded by GN Netcom.

REFERENCES

1. An HS, Jones GC, Kang SK, Welk GJ, Lee JM. How valid are wearable physical activity trackers for measuring steps? Eur J Sport Sci 17(3):360-8, 2017.

2. Boudreaux BD, Hebert EP, 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 50(3):624-33, 2018.

3. Bunn JA, Jones C, Oliviera A, Webster MJ. Assessment of step accuracy using the Consumer Technology Association standard, J Sports Sci 37(3):244-248, 2019. DOI:10.1080/02640414.2018.1491941

4. Chen MD, Kuo CC, Pellegrini CA, Hsu MJ. Accuracy of wristband activity monitors during ambulation and activities. Med Sci Sports Exerc 48(10):1942-9, 2016.

5. Consumer Technology Association. Physical activity monitoring for fitness wearables: Step counting. in. ANSI/CTA-20562016.

6. Consumer Technology Association. Physical activity monitoring for heart rate. In. CTTA-20652018.

7. Dolezal BA, Boland DM, Carney J, Abrazado M, Smith DL, Cooper CB. Validation of heart rate derived from a physiological status monitor-embedded compression shirt against criterion ECG. J Occup Environ Hyg 11(12):833-9, 2014.

8. Evenson KR, Goto MM, Furberg RD. Systematic review of the validity and reliability of consumer-wearable activity trackers. Int J Behav Nutr Phys Act 12:159, 2015.

9. Fokkema T, Kooiman TJ, Krijnen WP, CP VDS, M DEG. Reliability and validity of ten consumer activity trackers depend on walking speed. Med Sci Sports Exerc 49(4):793-800, 2017.

10. Gillinov S, Etiwy M, Wang R, Blackburn G, Phelan D, Gillinov AM, Houghtaling P, Javadikasgari H, Desai MY. Variable accuracy of wearable heart rate monitors during aerobic exercise. Med Sci Sports Exerc 49(8):1697-703, 2017.

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

12. Jo E, Lewis K, Directo D, Kim MJ, Dolezal BA. Validation of biofeedback wearables for photoplethysmographic heart rate tracking. J Sports Sci Med 15(3):540-7, 2016.

13. Kooiman TJ, Dontje ML, Sprenger SR, Krijnen WP, van der Schans CP, de Groot M. Reliability and validity of ten consumer activity trackers. BMC Sports Sci Med Rehabil 7:24, 2015.

14. Laukkanen RMT, Virtanen PK. Heart rate monitors: State of the art. J Sports Sci 16(suppl):S3-S7, 1998.

15. Leboeuf SF, Aumer ME, Kraus WE, Johnson JL, Duscha B. Earbud-based sensor for the assessment of energy expenditure, HR, and VO2max. Med Sci Sports Exerc 46(5):1046-52, 2014.

16. Lin L. A note on the concordance correlation coefficient. Biometrics 56:324-5, 2000.

17.Lyons EJ, Lewis ZH, Mayrsohn BG, Rowland JL. Behavior change techniques implemented in electronic lifestyle activity monitors: a systematic content analysis. J Med Internet Res. 16(8):e192, 2014.

18. Nelson EC, Verhagen T, Noordzij ML. Health empowerment through activity trackers: an empirical smart wristband study. Computers in Human Behavior 62:364-74, 2016.

19. Nelson MB, Kaminsky LA, Dickin DC, Montoye AH. Validity of consumer-based physical activity monitors for specific activity types. Med Sci Sports Exerc 48(8):1619-28, 2016.

20. O'Connell S, G OL, Kelly L, Murphy E, Beirne S, Burke N, Kilgannon O, Quinlan LR. These shoes are made for walking: Sensitivity performance evaluation of commercial activity monitors under the expected conditions and circumstances required to achieve the international daily step goal of 10,000 steps. PLoS One 11(5):e0154956, 2016.

21. Parak J, Korhenen I. Evaluation of wearable consumer heart rate monitors based on photoplethysmography. In. 2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. Chicago, IL2014.

22. Poh M, Kim K, Goessling A, Swenson N, Picard R. Cardiovascular monitoring using earphones and a mobile device. IEEE Pervasive Computing 11(4):18-26, 2012.

23. Sears T, Avalos E, Lawson S, McAlister I, Eschbach LC, Bunn J. Wrist-worn physical activity trackers tend to underestimate steps during walking. International Journal of Exercise Science 10(5):764-73, 2017.

24. Spierer DK, Rosen Z, Litman LL, Fujii K. Validation of photoplethysmography as a method to detect heart rate during rest and exercise. J Med Eng Technol 39(5):264-71, 2015.

25. Stahl SE, An HS, Dinkel DM, Noble JM, Lee JM. How accurate are the wrist-based heart rate monitors during walking and running activities? Are they accurate enough? BMJ Open Sport Exerc Med 2(1):e000106, 2016.

26. Terbizan D, Dolezal BA, Albano C. Validity of seven commercially available heart rate monitors. Measurement in Physical Education and Exercise Science 6:243-7, 2002.

27. Tudor-Locke C, Sisson SB, Lee SM, Craig CL, Plotnikoff RC, Bauman A. Evaluation of quality of commercial pedometers. Can J Public Health 97 Suppl 1:S10-5, S-6, 2006.

28. Van Remoortel H, Giavedoni S, Raste Y, Burtin C, Louvaris Z, Gimeno-Santos E, Langer D, Glendenning A, Hopkinson NS, Vogiatzis I, Peterson BT, Wilson F, Mann B, Rabinovich R, Puhan MA, Troosters T. Validity of activity monitors in health and chronic disease: A systematic review. International journal of behavioral nutrition and physical activity 9:84-84, 2012.

29. Wallen MP, Gomersall SR, Keating SE, Wisloff U, Coombes JS. Accuracy of heart rate watches: Implications for weight management. PLoS One 11(5):e0154420, 2016.

30. Wang R, Blackburn G, Desai M, Phelan D, Gillinov L, Houghtaling P, Gillinov M. Accuracy of wrist-worn heart rate monitors. JAMA Cardiol 2(1):104-6, 2017.

31. Woodman JA, Crouter SE, Bassett Jr DR, Fitzhugh EC, Boyer WR. Accuracy of consumer monitors for estimating energy expenditure and activity type. Med Sci Sports Exerc 49(2):371-377, 2017.

32. World Medical Association. World Medical Association Declaration of Helsinki: Ethical Principles for Medical Research Involving Human Subjects. JAMA 310:2191-2194, 2013.