



Published in final edited form as:

PM R. 2017 November ; 9(11): 1077–1084. doi:10.1016/j.pmrj.2017.03.012.

## Validation of a self-monitoring tool for use in exercise therapy

Camilla S. Powierza<sup>1</sup>, Michael D. Clark<sup>1,2,3</sup>, Jaime M. Hughes, MPH, MSW<sup>4,5</sup>, Kevin A. Carneiro, DO<sup>5</sup>, and Jason P. Mihalik, PhD, CAT(C), ATC<sup>2,3</sup>

<sup>1</sup>School of Medicine, The University of North Carolina at Chapel Hill, Chapel Hill, NC

<sup>2</sup>Matthew Gfeller Sport-Related Traumatic Brain Injury Research Center, Department of Exercise and Sport Science, The University of North Carolina at Chapel Hill, Chapel Hill, NC

<sup>3</sup>Human Movement Science Curriculum, Department of Allied Health Sciences, The University of North Carolina at Chapel Hill, Chapel Hill, NC

<sup>4</sup>Program on Integrative Medicine, The University of North Carolina at Chapel Hill, Chapel Hill, NC

<sup>5</sup>Department of Physical Medicine & Rehabilitation, The University of North Carolina at Chapel Hill, Chapel Hill, NC

### Abstract

**Background**—Aerobic exercise at a sub-symptom heart rate has been recommended as therapy for Post-Concussion Syndrome (PCS). Assessing adherence with an accurate heart rate monitoring instrument is difficult limiting the proliferation of large-scale randomized controlled trials.

**Objective**—To evaluate the validity of the Fitbit Charge HR against electrocardiogram (EKG) to monitor heart rate during a treadmill-based exercise protocol.

**Design**—A methods comparison study.

**Setting**—Sports medicine research center within a tertiary care institution.

**Participants**—A convenience sample of 22 healthy participants (12 female) aged 18–26 years (mean age:  $22 \pm 2$  years).

**Methods**—Fitbit Charge HR heart rate measurements were compared to EKG data concurrently collected while participants completed the Buffalo Concussion Treadmill Test (BCTT).

**Main Outcome Measures**—Agreement between Fitbit Charge HR and EKG was assessed by intraclass correlation coefficients ( $ICC_{3,1}$ ), Bland-Altman limits of agreement, and percent error.

**Results**—We observed a strong single-measure absolute agreement between Fitbit Charge HR and EKG ( $ICC_{3,1} = 0.83$ ; 95% CI: 0.67 – 0.90). Fitbit Charge HR underestimated heart rate compared to EKG (mean difference =  $-6.04$  beats per min (bpm); SD = 10.40 bpm; Bland-Altman

---

Corresponding Author: Jason P. Mihalik, 2201 Stallings-Evans Sports Medicine Center, Campus Box 8700, Chapel Hill, NC 27599, jmihalik@email.unc.edu, Phone: 919-962-2573, Fax: 919-962-7274.

**Note:** this work has not been presented at the AAPM&R Annual Assembly

**Publisher's Disclaimer:** This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

95% limits of agreement =  $-26.42$  to  $14.35$  bpm). 69.9% of Fitbit heart rate measurements were within 10% error compared to EKG, and 91.5% of all heart rate measurements were within 20% error.

**Conclusions**—While the mean bias in measuring heart rate was relatively small, the limits of agreement between the Fitbit Charge HR and EKG were broad. Thus, the Fitbit Charge HR would not be a suitable option for monitoring heart rate within a narrow range. For the purposes of post-concussion exercise therapy, the relatively inexpensive cost, easy implementation, and low maintenance make Fitbit Charge HR a viable option for assessing adherence to an exercise program when expensive clinical equipment is unavailable.

---

## Introduction

Behavioral and physiological self-monitoring devices may facilitate individualized medical care. By measuring and recording data on activity, heart rate, and sleep quality, such devices can provide objective information about a patient's day-to-day life. This information is useful when lifestyle modifications have been prescribed, particularly given the difficulty in assessing patient adherence to modifications critical to clinical improvement. Additionally, self-monitoring devices may measure meaningful clinical outcomes (e.g. average or maximum heart rate during exercise activities) and offer motivation for the patients themselves. This is only possible if the self-monitoring device provides accurate information for its clinical purposes. In the past two years, Fitbit-like devices' validity, reliability, and feasibility of the activity monitoring function have been published, and mostly have been found to be accurate [1–8]. However, only a few studies report these devices' accuracy in measuring heart rate [9–12] and sleep [13–16].

One potential self-monitoring device application is tracking prescribed home exercise protocols for patients with Post-Concussion Syndrome (PCS). These novel rehabilitation approaches demonstrate promising results in resolving symptoms refractory to cognitive and physical rest [17,18]. As many as 10% of concussed athletes will continue endorsing symptoms for more than the seven days typically required for recovery, with nearly one-quarter of these athletes continuing to experience symptoms 6–12 weeks after injury [19]. Patients with refractory concussion symptoms are clinically diagnosed with PCS, although no consensus has been reached on the symptom duration required for this diagnosis.

Specific recommendations for treating PCS patients are not well established. Typically, the same guidelines governing the acute period following concussion are used. This conservative management plan entails complete physical and cognitive rest with a gradual return-to-activity that is implemented when symptoms resolve [20,21]. The problem with this approach is that by the definition of PCS, symptoms are not resolving in a timely manner and further rest is unlikely to hasten symptom resolution.

Treadmill-based testing can determine both readiness of return-to-activity and aerobic exercise level in which the patient should engage. Preliminary studies have suggested using controlled and graded exercise, measured by a sub-symptom threshold target heart rate, as a PCS treatment to decrease symptoms and improve recovery time; however, the underlying mechanism by which this is achieved is not fully understood [17,18,20,22,23]. The Buffalo

Concussion Treadmill Test (BCTT) was developed as a way to estimate sub-symptom threshold heart rate by gradually raising heart rate in a controlled manner [18,20,24]. The inability to accurately measure heart rate in nonclinical settings (e.g. during home-based rehabilitation) and to quantify adherence to exercise therapy are major barriers to studying the efficacy of exercise as a PCS treatment in larger randomized controlled trials. The use of daily self-report logs when prescribed extensive at-home exercise regimens have been found to be largely inaccurate at capturing total amount of physical activity [25,26].

When exercise has been prescribed to PCS patients, the exercise dose intensity should be restricted to a heart rate whereby the patient does not experience symptom exacerbation. Leddy et al. suggest a heart rate that is 80% of the symptom-eliciting/exacerbating level determined through a treadmill-based exercise test [18]. An accurate self-monitoring device would be beneficial both to the patient, by providing real-time heart rate feedback, and to the provider, by providing detailed reports on their patient's activity frequency, duration, and intensity, as judged by heart rate. The purpose of this study was to evaluate the Fitbit Charge HR's ability to provide real-time, quantitative feedback on heart rate for use in PCS rehabilitation therapy. In this report, we delimit the scope of our investigation to the Fitbit Charge HR's agreement with electrocardiogram (EKG) in measuring heart rate in healthy, non-concussed individuals during the BCTT [20]. In consideration of recent studies outlining the accuracy of Fitbit devices in activity monitoring function [4,5,7,8,13], we hypothesize the Fitbit Charge HR will have good agreement with EKG in measuring heart rate during exercise. We report on the validity of Fitbit Charge HR device through intra-class correlation, Bland-Altman limits of agreement, and percent error.

## Methods

### Participants

A convenience sample of 22 participants (12 female) aged 18–26 years old (mean age:  $22 \pm 2$  years; mean mass:  $71.16 \pm 17.34$  kg; mean height:  $175.09 \pm 14.73$  cm) was recruited. Participants with an underlying heart condition, history of epilepsy, seizure, or balance disorder, recent musculoskeletal injury or concussion, reliance on an external device to ambulate, or who were pregnant, were excluded from the study. All participants provided informed consent. The study comprised one visit for exercise testing and a seven-day home trial in which the participants wore the Fitbit Charge HR device while sleeping; results from the sleep-related arm of study are not presented here. The study was approved by the Institutional Review Board at the authors' institution.

### Instrumentation

The Fitbit Charge HR was chosen because of its ability to provide continuous heart rate display and recording, as well as its ability to record sleep quality and quantity. Per the manufacturer's description, the device's PurePulse™ light-emitting diodes on the skin-facing surface monitor blood volume changes to continuously estimate heart rate [27]. The device stores heart rate data at 1 second intervals during exercise tracking (so-called "Activity Mode," per manufacturer), and at 5 second intervals at all other times. The Fitbit Charge HR was not used in Activity Mode, but rather kept in its default setting to emulate the subjects'

typical use, because subjects may not remember to switch the device to the activity mode. Seven Fitbit Charge HR devices were used in total and randomly selected for each participant.

EKG data was acquired at 500 samples/sec using the Biopac MP150 system. Heart rate was estimated by R–R interval using AcqKnowledge analysis software, which automatically identifies points of the ECG complex and calculates timing intervals to determine heart rate.

Settings were changed on the Fitbit Charge HR devices so that seconds were included in the time display. The time on the computer collecting EKG data was synced to match the time displayed on the Fitbit Charge HR devices to the second.

### **Exercise Protocol**

Participants were outfitted with the Fitbit Charge HR on their non-dominant hand, approximately 2 centimeters from the ulnar styloid process, and lead II of the wireless EKG. The EKG leads were placed at the right midclavicular line above the nipple and at the left midclavicular line below the nipple, with a grounding lead placed at the right midclavicular line below the nipple. A baseline heart rate was measured while each participant was sitting and restricting movement, before they completed the BCTT, as described previously [20]. Briefly, this protocol entails recording heart rate while the participant walks on a treadmill beginning with an incline of zero at a speed of 3.3–3.6 mph (1.48–1.61 m/s), varying only to accommodate walking comfort. After 2 minutes, the incline was raised by 1% and then 1% every subsequent minute, with speed held constant throughout the protocol. The protocol was terminated when either the maximum incline of 10% was reached, a heart rate exceeding 149 bpm was verified by EKG, or the participant asked to discontinue. The termination guidelines were created in attempt to mimic the estimated length of time and intensity of exercise that is typical in the BCTT protocol [20].

### **Data Retrieval**

Heart rate data from Fitbit Charge HR devices were extracted from the Fitbit website via a third-party program called Fitabase (Small Steps Labs). The Fitabase program allowed for automated time-stamped heart rate measurements collected by Fitbit Charge HR devices, data that are not normally as conveniently exported at the Fitbit consumer-level. Heart rate data were exported from Fitabase in an excel format, allowing for easier data analysis. Use of Fitabase was necessary for this validation study, but would not be necessary for patients to monitor heart rate during exercise therapy.

### **Data Reduction**

The Fitbit Charge HR recorded a heart rate measurement with variable intervals ranging between 2–15 seconds. Step interpolation was used to resolve the Fitbit Charge HR heart rate data to one measurement per second. This interpolation aligns with the real-time heart rate display of the device, which projects the last measurement recorded until there is a change in the heart rate measurement.

Heart rate from the EKG system was calculated using the R–R interval and was down-sampled and time-stamp matched to the one-second-resolved, interpolated Fitbit Charge HR data. We manually removed artifacts where the EKG waveform trace was improperly labelled (e.g. a motion artifact being labelled as an R wave resulting in a false R–R interval). Furthermore, we removed readings where there was a change of 20 bpm or greater over one second; such an artifact may result from a premature ventricular contraction. We also discarded data where the heart rate was recorded to be less than 40 bpm or greater than 170 bpm. Less than 2.3% of data were removed using these filtering rules.

### Statistical Analysis

Statistical analyses were performed using IBM SPSS 19 and SAS 9.3; figures were created in R. First, we compared the point estimates of heart rate at each second. Next, we created 5- and 10-second centered moving averages for each system and compared the resultant averages. To examine the agreement between the Fitbit Charge HR and the EKG system we computed intraclass correlation coefficients ( $ICC_{3,1}$ ), Bland-Altman limits of agreement, and percent error. The  $ICC_{3,1}$  were calculated using a 2-way, mixed model in which participants were treated as random effects and the two systems as fixed effects. Bland-Altman limits of agreement were calculated using a repeated-measures, matched replicates model, as described previously [28]. We report the 95% confidence intervals throughout. The percent error between each measurement from the devices was calculated by taking the absolute difference between the Fitbit Charge HR and EKG measurements and dividing by the EKG heart rate; we report the percent error as the percentage of error measurements that were less than 10% or less than 20%. Lastly, we employed a linear mixed model to determine if significant differences in measurement error were observed between the Fitbit Charge HR devices used in our study.

### Results

We collected a mean of 860 seconds (SD: 90s) across 22 participants (18,415 pairs of heartrate readings). On visual inspection, Fitbit Charge HR data did not track well with EKG for three subjects. A sample of individual time-series plots are given in Figure 1. Single-measure, absolute agreement  $ICC_{3,1}$  was 0.83 (95% CI: 0.67 – 0.90). A Bland-Altman plot was created for all 22 participants using the subject-specific mean difference and mean heart rate for each subject (Figure 2). The overall mean difference in heart rate between Fitbit Charge HR and EKG was –6.04 bpm (SD: 10.40 bpm, 95% limits of agreement: –26.42 to 14.35 bpm). Individual Bland-Altman plots are shown in Figure 3, with a summary of the mean difference and individual limits of agreement in Figure 4. 69.9% percent of Fitbit Charge HR heart rate readings were within 10% of the mean EKG heart rate readings, and 91.5% of Fitbit Charge HR heart rate readings were within 20% of the mean EKG heart rate.

Centered moving averages over five and ten seconds were calculated for each device and compared using  $ICC_{3,1}$  and Bland-Altman limits of agreement. These moving averages resulted in modest improvement of the ICCs and tighter 95% limits of agreement, but negligible change in the mean difference between the devices. These results are presented in

the Table 1. Mean difference in heart rate did not vary significantly by Fitbit Charge HR device ( $F_{6,12}=1.58$ ,  $P=.24$ ).

We ran the agreement analyses excluding the subjects with poor tracking between Fitbit and EKG (Table 2), resulting in some improvement in all measures. However, we have no compelling reason to believe the acquisition of data was compromised for these subjects.

## Discussion

The purpose of this study was to assess the validity of heart rate measurements from the Fitbit Charge HR in comparison to the gold standard of EKG. Our finding that the Fitbit Charge HR underestimates heart rate (mean underestimation =  $-6.04$  bpm) is unsurprising. Fitbit Charge HR uses a similar mechanism for heart rate estimation as standard pulse oximetry [27], which has been shown to underestimate heart rate in adults, especially during exercise [29,30]. Our results are consistent with a recent report by Jo et al., wherein the authors, comparing the Fitbit Charge HR to EKG, found a similar mean underestimate of 8.8 bpm [10]. For the purposes of monitoring exercise activity intensity, an underestimation of approximately 6 bpm is of little clinical importance. The variability of the agreement between the devices is of much greater concern.

While the ICCs were generally high ( $>0.8$ ) indicating a high overall consistency between the EKG and Fitbit Charge HR, the range of disagreement between the devices was rather large. The 95% limits of agreement were  $-26$  and  $+14$  bpm. This range represents the degree of uncertainty between the devices at the 95% confidence level. These limits did not appear to vary by mean heart rate. While these broad limits of agreement suggest the Fitbit Charge HR is not as accurate as the EKG, the percent error was generally low. Less than 10% of heart rate readings exceeded 20% error, and 30% of readings differed by more than 10%. The ICCs, mean difference, and limits of agreement were only marginally improved by using centered moving averages. Heart rate accuracy was not majorly improved by taking an average of the Fitbit Charge HR over five or ten seconds.

One possible cause of disagreement between devices may be due to differences in the devices' method of estimating heart rate. EKG measures heart rate at a high sampling rate, estimating from a single R-R interval, which is inherently unstable. Fitbit Charge HR measures heart rate at a much lower sampling rate compared to EKG and thus would not reflect normal second-to-second heart rate variability. However, this inaccuracy may be clinically negligible, since we compared five- and ten-second centered moving averages, with only minimal improvement in the limits of agreement.

Although EKG is considered the gold standard for heart rate measurement, its use remains limited outside the clinical/medical setting. The EKG is designed for stationary patients, and leads should be placed by a trained professional to maintain maximal accuracy. In this study, a wireless EKG device was chosen to accommodate the motion inherent in the exercise protocol we employed. However, wireless EKG devices are inclined to have increased signaling noise as activity level increases, with increased errors seen during accelerated

movements [31]. Although our EKG data were fairly noise-free, a small percentage of data were filtered out to decrease any motion-related artifacts.

Three participants had poor quality Fitbit Charge HR data. We lack the data to determine why these participants had poor agreement between devices, but this may be suggestive of subsets of participants for which the Fitbit Charge HR is ill-suited. Poor performance of the devices could be a function of peripheral circulation in distal extremities, but since peripheral pulses were not assessed during our study, this remains speculative. Previous studies have suggested that changes in contact force between photoplethysmographic sensors, such as the one used by Fitbit Charge HR devices, and skin can affect the quality of blood flow, causing artifacts in heart rate measurement [32]. Improper positioning of the device, such as within a skin fold, or the presence of hair or sweat can interrupt signal transduction [33]. Even microstructures such as a participant's skin pigmentation and vascular network have been suggested to cause variations in readings, although this remains uncertain [10,33]. The fit of the Fitbit Charge HR was assessed for each participant in our study and did not appear to be problematic per manufacturer guidelines. Our analysis did not provide any conclusions for why these three participants had poor Fitbit Charge HR data. The reasons for poor quality Fitbit Charge HR data on some participants should be examined in future validation studies.

A limitation to these results is that our study population consisted of young, healthy, and physically-active adults. It is unknown if the agreement would change in those with more sedentary lifestyles, circulatory problems, or even those in our eventual clinical population of concussed athletes. Fitbit Charge HR would likely be susceptible to more errors with any high-intensity activity involving excessive arm movement (e.g., boxing, rowing, or P90X) or in individuals whose wrist size would prevent an adequate fit of the device. The walking treadmill protocol used was purposefully chosen to emulate previously proposed PCS exercise therapy [17,18,20,22,23]. As such, our results may not be generalizable to all intensities of exercise. A recent study by Jo et al. found that the Fitbit Charge HR was less accurate as the heart rate increased [10]. We did not observe a significant variation in mean differences between Fitbit Charge HR and EKG across heart rate, but given our protocol did not include vigorous exercise or a period of rest, this analysis is beyond the scope of our investigation. The agreement between devices during a higher intensity exercise protocol needs to be further studied.

## Conclusion

With less than 10% of heart rate measurements exceeding 20% error, the Fitbit Charge HR provides a moderately accurate method to monitor heart rate during exercise in nonclinical settings. For individuals free from a cardiovascular disease that might require more advanced cardiac monitoring, the Fitbit Charge HR is a viable option for exercise tracking due to its relatively inexpensive cost, easy implementation, and minimal maintenance requirements. The Fitbit Charge HR may become a useful tool in PCS therapy in that it provides clinicians with an adequately accurate method to monitor adherence to exercise therapy. However, due to the broad limits of agreement between the Fitbit Charge HR and EKG, the device is a

poor choice for monitoring heart rate in real-time, particularly in the context of staying within a narrow range of heart rates or below a target heart rate.

## Acknowledgments

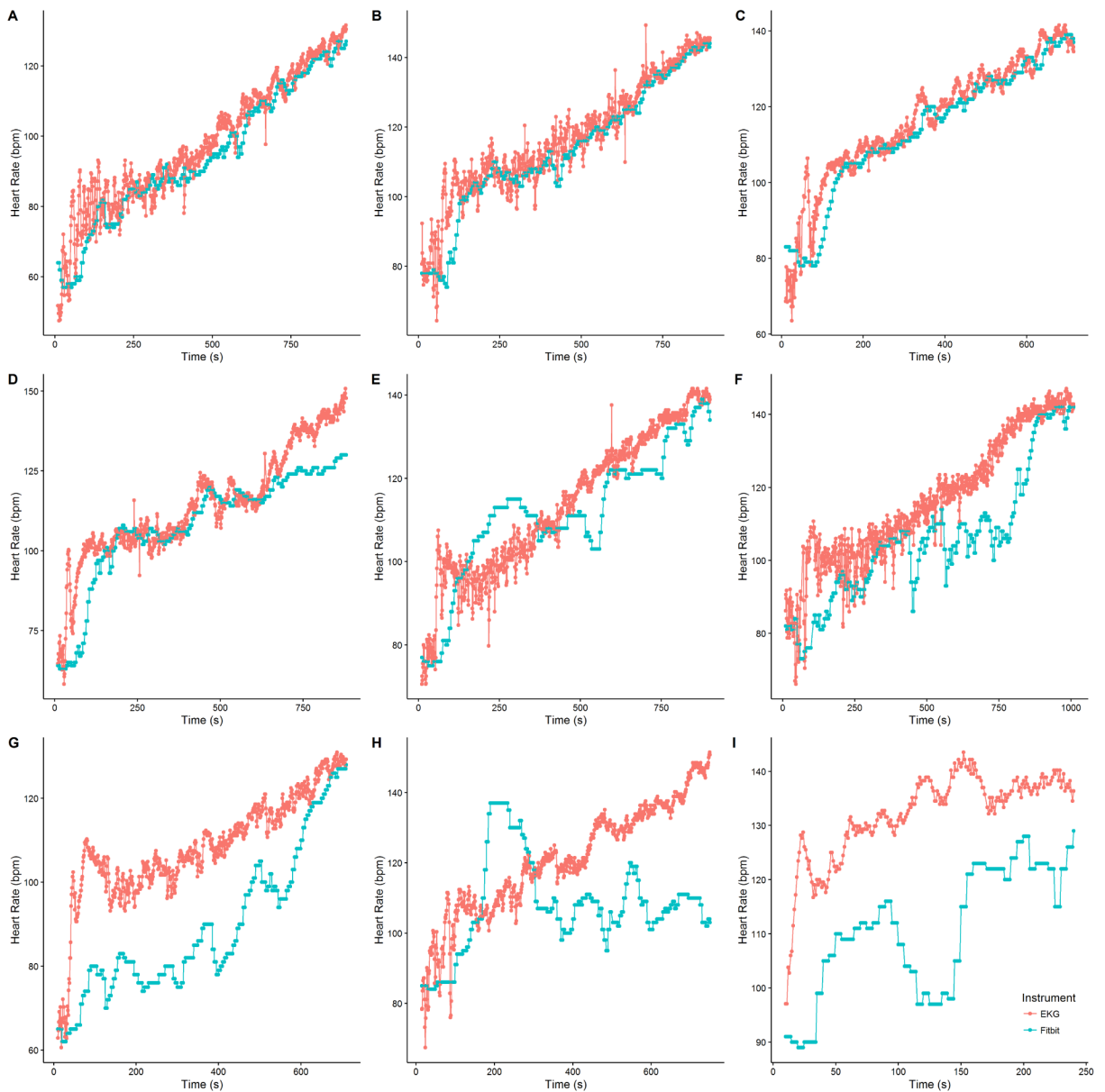
**Funding:** This work was supported by National Institutes of Health grants F30NS090816, T35DK007386, and T32ATT003378, and internal funding from the Matthew Gfeller Sport-Related Traumatic Brain Injury Research Center.

## References

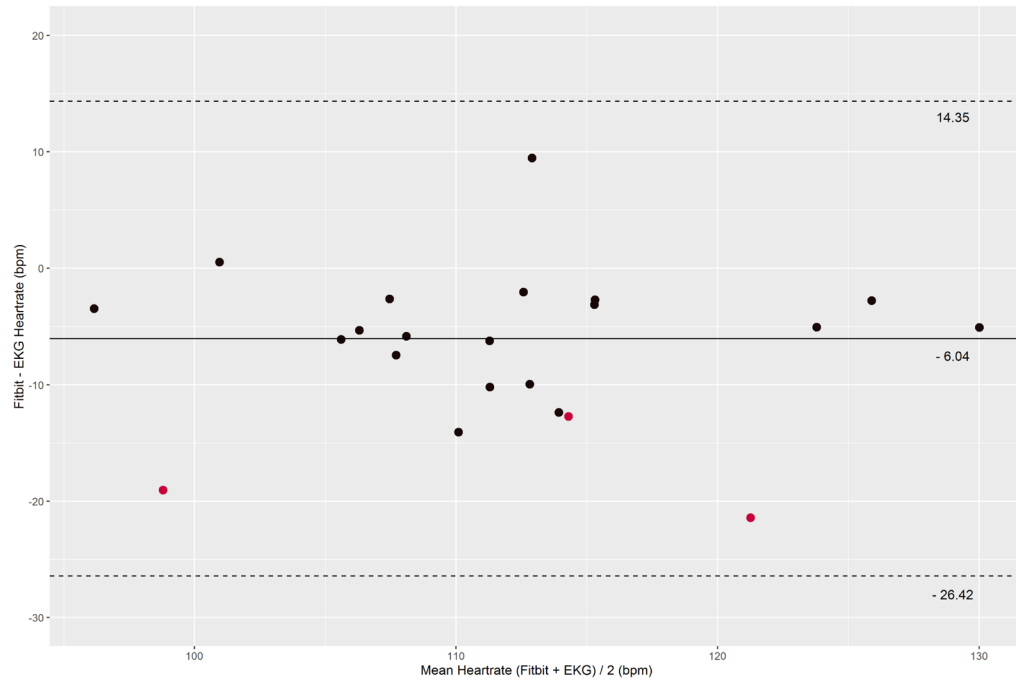
1. Bai Y, Welk GJ, Nam YH, Lee JA, Lee J-M, Kim Y, et al. Comparison of Consumer and Research Monitors under Semistructured Settings. *Med Sci Sports Exerc.* 2015
2. Cadmus-Bertram LA, Marcus BH, Patterson RE, Parker BA, Morey BL. Randomized Trial of a Fitbit-Based Physical Activity Intervention for Women. *Am J Prev Med.* 2015; 49:414–8. [PubMed: 26071863]
3. Dannecker KL, Sazonova NA, Melanson EL, Sazonov ES, Browning RC. A comparison of energy expenditure estimation of several physical activity monitors. *Med Sci Sports Exerc.* 2013; 45:2105–12. [PubMed: 23669877]
4. Diaz KM, Krupka DJ, Chang MJ, Peacock J, Ma Y, Goldsmith J, et al. Fitbit®: An accurate and reliable device for wireless physical activity tracking. *Int J Cardiol.* 2015; 185:138–40. [PubMed: 25795203]
5. Dontje ML, de Groot M, Lengton RR, van der Schans CP, Krijnen WP. Measuring steps with the Fitbit activity tracker: an inter-device reliability study. *J Med Eng Technol.* 2015; 39:286–90. [PubMed: 26017748]
6. Sasaki JE, Hickey A, Mavilia M, Tedesco J, John D, Kozey Keadle S, et al. Validation of the Fitbit wireless activity tracker for prediction of energy expenditure. *J Phys Act Health.* 2015; 12:149–54. [PubMed: 24770438]
7. Takacs J, Pollock CL, Guenther JR, Bahar M, Napier C, Hunt MA. Validation of the Fitbit One activity monitor device during treadmill walking. *J Sci Med Sport Sports Med Aust.* 2014; 17:496–500.
8. Vooijs M, Alpay LL, Snoeck-Stroband JB, Beerthuizen T, Siemonsma PC, Abbink JJ, et al. Validity and usability of low-cost accelerometers for internet-based self-monitoring of physical activity in patients with chronic obstructive pulmonary disease. *Interact J Med Res.* 2014; 3:e14. [PubMed: 25347989]
9. Åkerberg A, Koshmak G, Johansson A, Lindén M. Heart rate measurement as a tool to quantify sedentary behavior. *Stud Health Technol Inform.* 2015; 211:105–10. [PubMed: 25980855]
10. Jo E, Lewis K, Directo D, Kim MJ, Dolezal BA. Validation of Biofeedback Wearables for Photoplethysmographic Heart Rate Tracking. *J Sports Sci Med.* 2016; 15:540. [PubMed: 27803634]
11. Leth S, Hansen J, Nielsen OW, Dinesen B. Evaluation of Commercial Self-Monitoring Devices for Clinical Purposes: Results from the Future Patient Trial, Phase I. *Sensors.* 2017:17.
12. Stahl SE, An H-S, Dinkel DM, Noble JM, Lee J-M. How accurate are the wrist-based heart rate monitors during walking and running activities? Are they accurate enough? *BMJ Open Sport Exerc Med.* 2016; 2:e000106.
13. Ferguson T, Rowlands AV, Olds T, Maher C. The validity of consumer-level, activity monitors in healthy adults worn in free-living conditions: a cross-sectional study. *Int J Behav Nutr Phys Act.* 2015; 12:42. [PubMed: 25890168]
14. Lee J, Finkelstein J. Consumer sleep tracking devices: a critical review. *Stud Health Technol Inform.* 2015; 210:458–60. [PubMed: 25991187]
15. Meltzer LJ, Hiruma LS, Avis K, Montgomery-Downs H, Valentin J. Comparison of a Commercial Accelerometer with Polysomnography and Actigraphy in Children and Adolescents. *Sleep.* 2015; 38:1323–30. [PubMed: 26118555]



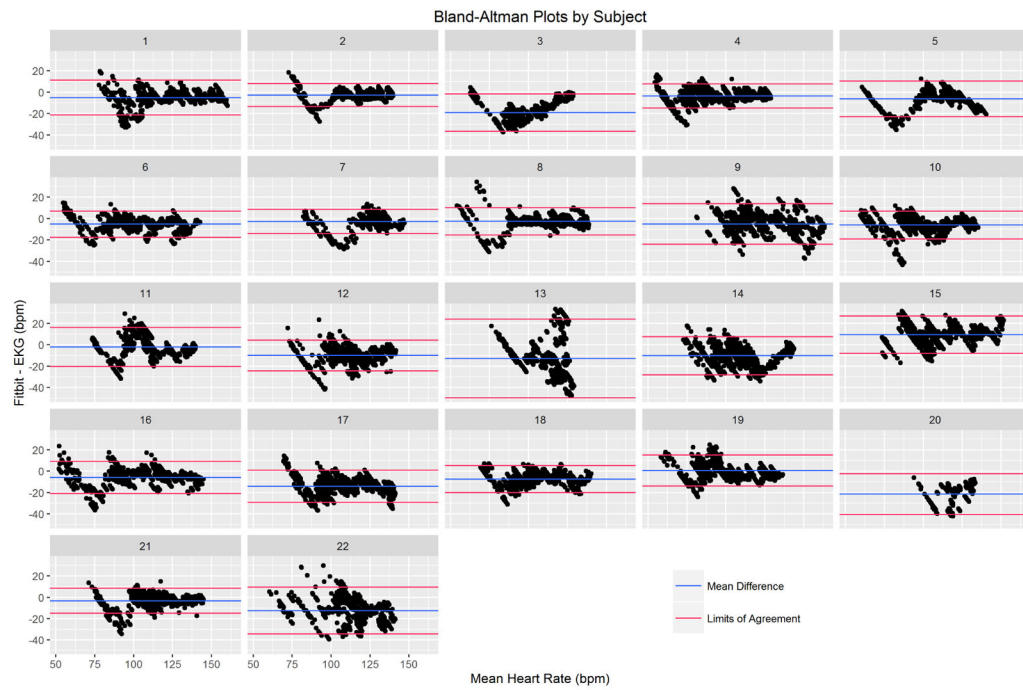
16. Perez-Macias, JM., Jimison, H., Korhonen, I., Pavel, M. Comparative assessment of sleep quality estimates using home monitoring technology. *Eng. Med. Biol. Soc. EMBC 2014 36th Annu. Int. Conf. IEEE, IEEE*; 2014; p. 4979-82.
17. Gagnon I, Galli C, Friedman D, Grilli L, Iverson GL. Active rehabilitation for children who are slow to recover following sport-related concussion. *Brain Inj.* 2009; 23:956–64. [PubMed: 19831492]
18. Leddy JJ, Kozlowski K, Donnelly JP, Pendergast DR, Epstein LH, Willer B. A preliminary study of subsymptom threshold exercise training for refractory post-concussion syndrome. *Clin J Sport Med.* 2010; 20:21–7. [PubMed: 20051730]
19. McCrear M, Guskiewicz K, Randolph C, Barr WB, Hammeke TA, Marshall SW, et al. Incidence, Clinical Course, and Predictors of Prolonged Recovery Time Following Sport-Related Concussion in High School and College Athletes. *J Int Neuropsychol Soc.* 2013; 19:22–33. [PubMed: 23058235]
20. Leddy JJ, Willer B. Use of Graded Exercise Testing in Concussion and Return-to-Activity Management. *Curr Sports Med Rep.* 2013; 12:370–6. [PubMed: 24225521]
21. McCrory P, Meeuwisse W, Aubry M, Cantu B, Dvorák J, Echemendia R, et al. Consensus statement on Concussion in Sport - The 4th International Conference on Concussion in Sport held in Zurich, November 2012. *Phys Ther Sport.* 2013; 14:e1–13. [PubMed: 23664041]
22. Baker JG, Freitas MS, Leddy JJ, Kozlowski KF, Willer BS. Return to Full Functioning after Graded Exercise Assessment and Progressive Exercise Treatment of Postconcussion Syndrome. *Rehabil Res Pract.* 2012; 2012:1–7.
23. Darling SR, Leddy JJ, Baker JG, Williams AJ, Surace A, Miecznikowski JC, et al. Evaluation of the Zurich Guidelines and Exercise Testing for Return to Play in Adolescents Following Concussion. *Clin J Sport Med.* 2014; 24:128–33. [PubMed: 24184849]
24. Leddy JJ, Baker JG, Kozlowski K, Bisson L, Willer B. Reliability of a graded exercise test for assessing recovery from concussion. *Clin J Sport Med Off J Can Acad Sport Med.* 2011; 21:89–94.
25. Prince SA, Adamo KB, Hamel ME, Hardt J, Connor Gorber S, Tremblay M. A comparison of direct versus self-report measures for assessing physical activity in adults: a systematic review. *Int J Behav Nutr Phys Act.* 2008; 5:56. [PubMed: 18990237]
26. Sallis JF, Saelens BE. Assessment of physical activity by self-report: status, limitations, and future directions. *Res Q Exerc Sport.* 2000; 71:S1–14. [PubMed: 10925819]
27. Fitbit Charge HR Product Manual n.d.
28. Bland JM, Altman DG. Agreement between methods of measurement with multiple observations per individual. *J Biopharm Stat.* 2007; 17:571–82. [PubMed: 17613642]
29. Iyriboz Y, Powers S, Morrow J, Ayers D, Landry G. Accuracy of pulse oximeters in estimating heart rate at rest and during exercise. *Br J Sports Med.* 1991; 25:162–4. [PubMed: 1777787]
30. Norton LH, Squires B, Craig NP, McLeay G, McGrath P, Norton KI. Accuracy of pulse oximetry during exercise stress testing. *Int J Sports Med.* 1992; 13:523–7. [PubMed: 1459747]
31. Takalokastari T, Alasaarela E, Kinnunen M, Jamsa T. Quality of the Wireless Electrocardiogram Signal During Physical Exercise in Different Age Groups. *IEEE J Biomed Health Inform.* 2014; 18:1058–64. [PubMed: 24132026]
32. Rafolt D, Gallasch E. Influence of contact forces on wrist photoplethysmography--prestudy for a wearable patient monitor. *Biomed Tech (Berl).* 2004; 49:22–6. [PubMed: 15032494]
33. Butler MJ, Crowe JA, Hayes-Gill BR, Rodmell PI. Motion limitations of non-contact photoplethysmography due to the optical and topological properties of skin. *Physiol Meas.* 2016; 37:N27–37. [PubMed: 27100666]



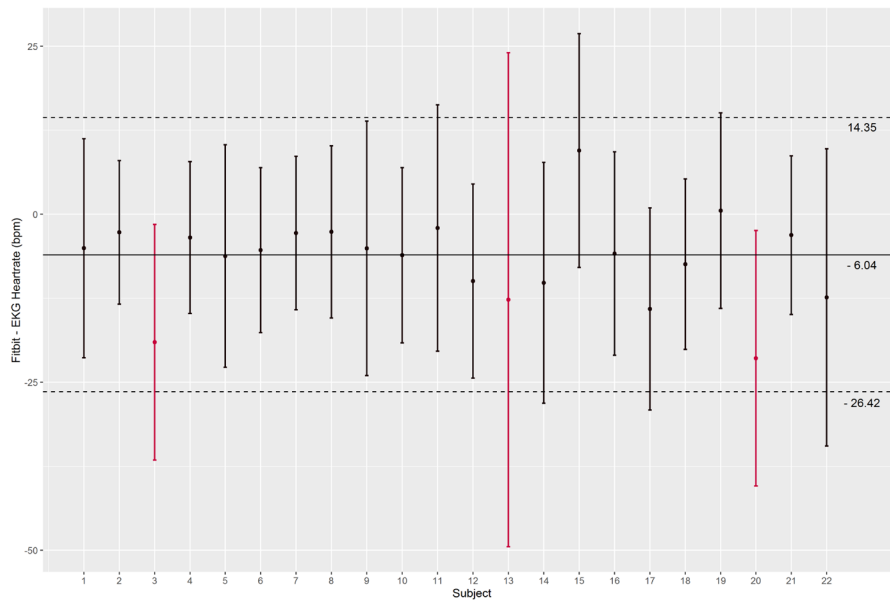
**Figure 1.** Time-series data for nine subjects; A–F show representative subjects (4, 21, 2, 5, 11, and 14, respectively) with heart rate data closely aligned for both Fitbit Charge HR and EKG. In contrast, G–I are subjects (3, 13, and 20, respectively) for whom there was large disagreement between the two devices.



**Figure 2.** Bland-Altman plot of the mean difference in heart rate recordings by device for each subject. The overall mean difference and 95% limits of agreement are shown. The red data points reflect subjects for whom there was poor tracking between the EKG and Fitbit.



**Figure 3.** Individual Bland-Altman plots are shown for each subject. The within-subject mean differences and 95% limits of agreement are shown as blue and red lines, respectively.



**Figure 4.**

The subject-specific mean difference between Fitbit and EKG are given as point estimates with 95% limits of agreement represented as error bars. The overlaid solid line is the overall mean difference between Fitbit and EKG heart rates and the overlaid dashed lines are the overall limits of agreement for the entire dataset.

**Table 1**

Assessment of agreement between Fitbit and EKG (N=22). Mean Difference is computed by subtracting EKG-derived heart rate from the Fitbit-measured heart rate, and reported in beats per minute

Comparison	ICC (95% CI)	Mean Difference (95% Limits of Agreement)
Single measures	0.83 (0.67, 0.90)	-6.04 (-26.42, 14.35)
5 sec moving average	0.83 (0.66, 0.90)	-6.07 (-25.73, 13.59)
10 sec moving average	0.84 (0.65, 0.91)	-6.07 (-25.08, 12.95)

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

**Table 2**

Assessment of agreement between Fitbit and EKG after removing three subjects with poorly tracking Fitbit and EKG data (N=19). Mean Difference is computed by subtracting EKG-derived heart rate from the Fitbit-measured heart rate, and reported in beats per minute

Comparison	ICC (95% CI)	Mean Difference (95% Limits of Agreement)
Single measures	0.87 (0.76,0.92)	-4.96 (-22.74,12.81)
5 sec moving average	0.88 (0.75,0.93)	-5.00 (-21.95,11.95)
10 sec moving average	0.88 (0.75,0.93)	-5.01 (-21.26,11.24)

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript