

Assessing the ability of the Fitbit Charge 2 to accurately predict VO_{2max}

Kaitlin A. Freeberg, Brett R. Baughman, Ted Vickey, Jeff A. Sullivan, Brandon J. Sawyer

Departments of Kinesiology and Biology, Point Loma Nazarene University, San Diego, CA, USA

Contributions: (I) Conception and design: All authors; (II) Administrative support: KA Freeberg, T Vickey, BJ Sawyer; (III) Provision of study materials or patients: KA Freeberg, BR Baughman; (IV) Collection and assembly of data: KA Freeberg, BR Baughman; (V) Data analysis and interpretation: KA Freeberg, BJ Sawyer; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

Correspondence to: Brandon J. Sawyer, PhD. Point Loma Nazarene University, Department of Kinesiology and Biology, 3900 Lomaland Drive, San Diego, CA 92106, USA. Email: bsawyer@pointloma.edu.

Background: The aim of this study was to assess the ability of the Fitbit Charge 2 (FBC2) to accurately estimate VO_{2max} in comparison to both the gold standard VO_{2max} test and a non-exercise VO_{2max} prediction equation.

Methods: Thirty healthy subjects (17 men, 13 women) between the ages of 18 and 35 (age = 21.7 ± 3.1 years) were given a FBC2 to wear for seven days and followed instructions on how to obtain a cardio fitness score (CFS). VO_{2max} was measured with an incremental test on the treadmill followed by a verification phase. VO_{2max} was predicted via a non-exercise prediction model (N-Ex) using self-reported physical activity level.

Results: Measured VO_{2max} was significantly lower than FBC2 predicted CFS ($VO_{2max} = 49.91 \pm 6.83$; CFS = 52.53 ± 8.43 , $P = 0.03$). N-Ex prediction was significantly lower than CFS but not significantly lower than measured VO_{2max} (N-Ex = 48.79 ± 6.32 ; CFS *vs.* N-Ex: $P = 0.01$; VO_{2max} *vs.* N-Ex: $P = 0.54$). Relationships between both VO_{2max} *vs.* CFS and VO_{2max} *vs.* N-Ex were good (ICC: VO_{2max} *vs.* CFS = 0.87, VO_{2max} *vs.* N-Ex = 0.87); Bland-Altman analysis indicated consistency of CFS measurement and lack of bias. The coefficient of variation (CV) and mean absolute percent error (MAPE) were greater with CFS than N-Ex (CV: CFS = $6.5\% \pm 4.1\%$, N-Ex = $5.6\% \pm 3.6\%$; MAPE: CFS = $10.2\% \pm 6.7\%$, N-Ex = $7.8\% \pm 5.0\%$). Heart rate (HR) estimated by the FBC2 was lower than estimated (Est) HR for pace based on HR extrapolation (FBC2 = 155 ± 18 bpm, Est = 183 ± 15 bpm, $P < 0.001$). The difference in CFS and VO_{2max} was inversely correlated with the difference in FBC2 HR and Estimated HR ($r = -0.45$, $P < 0.001$).

Conclusions: The FBC2 shows consistent, unbiased measurement of CFS while overestimating VO_{2max} in healthy men and women. The non-exercise VO_{2max} prediction equation provides a similar, slightly more accurate, VO_{2max} prediction than the CFS without the need for an exercise test or purchase of a Fitbit.

Keywords: Cardiorespiratory; VO_{2max} ; exercise; heart; rate; Fitbit; cardio; fitness; score

Received: 22 May 2019; Accepted: 23 August 2019; Published: 23 September 2019.

doi: 10.21037/mhealth.2019.09.07

View this article at: <http://dx.doi.org/10.21037/mhealth.2019.09.07>

Introduction

VO_{2max} testing is known as the gold standard for measuring cardiorespiratory fitness and is frequently used in research settings to determine the efficacy of training program interventions (1). Exercise physiology laboratories regularly use VO_{2max} testing to evaluate the cardiorespiratory health of individuals as well as develop exercise prescriptions (1).

Furthermore, VO_{2max} is a strong predictor of cardiovascular disease (CVD) risk and overall CVD mortality (2). Maximal exercise testing has become the standard for measuring functional capacity, evaluating therapy, estimating risk, and organizing transplantation candidacy in patients with heart failure (3). Maximal exercise testing is also important in diagnosing and assessing coronary artery disease, peripheral

Table 1 Descriptive statistics of all participants (n=30)

Subject characteristic	Outcome, mean \pm SD
Age (yr)	21.7 \pm 3.1
BMI (kg/m ²)	23.5 \pm 2.6
Body fat (%)	20.5 \pm 7.1

arterial disease, heart failure, valvular heart disease, and unexplained exertional dyspnea (3). Use of exercise testing by physicians and non-physicians has grown extensively, resulting in the administration of millions of tests (4).

Despite the accuracy and proliferation of maximal testing, there are difficulties involved that make this type of testing less accessible to the general population. $\text{VO}_{2\text{max}}$ testing requires maximal effort and thus puts tremendous strain on the body. Furthermore, maximal testing requires access to a lab and specific equipment necessary for assessing oxygen uptake, single tests of which can be expensive for the general population. Fitbit has released the Charge 2 watch [Fitbit Charge 2 (FBC2)], which is advertised to predict $\text{VO}_{2\text{max}}$ by displaying a user friendly “cardio fitness score” (CFS). Using the relationship between running pace and heart rate (HR), the watch calculates a score comparable to one’s $\text{VO}_{2\text{max}}$ in mL/kg/min. To our knowledge, Fitbit has not released research on how the FBC2 specifically predicts $\text{VO}_{2\text{max}}$, thus the level of prediction accuracy is unclear. The accurate prediction of $\text{VO}_{2\text{max}}$ by a wrist worn device is appealing due to the lower cost, less strenuous testing methodology, and potential for more widespread awareness of cardiovascular health.

Other companies, such as Garmin, have created wearable personal fitness devices to estimate $\text{VO}_{2\text{peak}}$. One study sought to validate the use of the Garmin Forerunner 920XT watch in $\text{VO}_{2\text{peak}}$ estimation (5). Sixteen subjects were instructed to jog or run for ten minutes around a football field wearing HR monitors and the GPS Garmin watch and perform a treadmill $\text{VO}_{2\text{max}}$ test 2–5 days later (5). Results showed no significant differences between the mean $\text{VO}_{2\text{peak}}$ from the Garmin watch and the treadmill test as well as a high Pearson correlation coefficient ($r=0.84$), suggesting the Garmin Forerunner 920XT provides a relatively accurate prediction of $\text{VO}_{2\text{peak}}$ (5). However, to our knowledge, no studies have been performed to evaluate the accuracy of the FBC2 in $\text{VO}_{2\text{max}}$ estimation.

This study aimed to assess the ability of the FBC2 to

accurately estimate $\text{VO}_{2\text{max}}$ in comparison to both the gold standard $\text{VO}_{2\text{max}}$ test and a non-exercise $\text{VO}_{2\text{max}}$ prediction equation. We hypothesized that the FBC2 would overestimate $\text{VO}_{2\text{max}}$ due to its reduced HR monitoring accuracy at increased exercise intensities (6–9).

Methods

Experimental design

Thirty subjects (17 men, 13 women) were given the FBC2 to wear for seven days and followed instructions on how to obtain a CFS. Subjects came into the laboratory on two separate occasions. $\text{VO}_{2\text{max}}$ was predicted on their first visit via a non-exercise prediction model (N-Ex) using self-reported physical activity level (10) and subjects performed submaximal exercise to become familiar with the maximal exercise equipment. $\text{VO}_{2\text{max}}$ was measured at their second visit via an incremental test on the treadmill followed by a verification phase. Body composition was also assessed to determine accurate subject characteristics. Participants were advised to perform their individual runs at least 48 hours apart and abstain from physical activity 48 hours prior to their measured $\text{VO}_{2\text{max}}$ test.

Participants

On the basis of previously published data (11), we calculated that completing 27 subjects in our study would yield 95% power to detect a 2% difference in $\text{VO}_{2\text{max}}$ between CFS and measured $\text{VO}_{2\text{max}}$ (at a two-tailed alpha level of 0.05). Planning for subject attrition, we enrolled 34 subjects. Two subjects dropped out due to time constraints and two subjects were excluded from data analysis due to failure to adhere to instructions on how to obtain a CFS, resulting in a final sample size of 30. Physical characteristics of the participants who completed the study are shown in *Table 1*. Inclusion criteria were healthy, non-sedentary individuals aged 18–35 years old. Non-sedentary individuals were those who answered above a zero on the self-reported physical activity questionnaire (12). The study was approved by the university institutional review board; all subjects provided written informed consent and completed a Physical Activity Readiness Questionnaire (PAR-Q) before initiating the study to determine if the subject was healthy enough to exercise. Answering “yes” to any questions on the PAR-Q would immediately disqualify anyone from participation in the study.

Assessment of body composition

Body composition measurement was performed using air displacement plethysmography (Bod Pod Cosmed, Rome, Italy) (13). Subjects were fasted and refrained from exercise 12 hours prior to testing. Wearing minimal clothing (spandex shorts or swimsuit) and a swim cap, subjects were weighed on a calibrated digital scale and height was recorded from a wall-mounted stadiometer (Seca, Birmingham, UK). The subject was then instructed to sit quietly within the BOD POD chamber for two measurements of body volume, each lasting about 45 seconds. If these two measurements agreed within 150 mL, they were averaged. If the two measurements did not agree within 150 mL, a third measurement was taken and the two values that were the closest and met criteria for agreement were averaged. Using the data collected for body mass and body volume as well as the predicted thoracic lung volume, body density and percent body fat were calculated using the Siri equation (14).

Assessment of submaximal HR and equipment familiarization

Height and weight measurements were inserted into the N-Ex using self-reported physical activity level to predict VO_{2max} (10). Subjects were equipped with an oronasal mask connected to a standard nonrebreathing valve (Hans Rudolph, Shawnee, KS, USA) for continuous measurement of ventilation and respiratory gas exchange data using a previously validated (15) metabolic measurement system (Parvo Medics TrueOne 2400; Parvo medics, Sandy, UT, USA). A standard 3-point calibration was performed before each test or every four hours per manufacturer recommendations. While measuring gas exchange and HR data, subjects performed a submaximal treadmill run at 60, 70, 80, and 90% of their estimated VO_{2max} (10) to become familiar with the equipment. Subjects ran for three minutes at each intensity. Using steady state HR from each running pace, linear regression equations were created for each subject using running pace to estimate HR. These equations were subsequently used to estimate HR from the GPS measured running pace during the independent runs while wearing the FBC2. The estimated (Est) HR was then compared to the FBC2 measured HR.

Assessment of CFS

Subjects were assigned a FBC2 to wear for seven days. The

FBC2s were updated with the latest firmware at the time of the study which was version 22.55.2. During the seven days, subjects were asked to complete two independent runs on flat terrain with the FBC2. Acceptable locations for running were recommended and GPS tracking from the watch confirmed participants ran on flat terrain. Each of these runs consisted of a 5-minute warm up at a self-selected speed. With GPS and Bluetooth on and paired with their Fitbit account on their smart phone, subjects then performed a 10-minute run. Based on the instructions from the manufacturer on how to obtain a CFS, subjects were instructed to run at as high of an intensity as could be continuously sustained for the full 10 minutes. Subjects then synced watch data to their phone application and a CFS was calculated. Screenshots of the CFS, average pace, time, and average HR were sent to the primary investigator after each of the two runs.

Assessment of VO_{2max}

Subjects were set up with the same metabolic cart and procedures as during the familiarization visit. The incremental test protocol was chosen using an estimated VO_{2max} and estimated speed and grade that were designed to elicit exhaustion in approximately 10 minutes (12). After collecting 2 minutes of resting data, subjects warmed up for five minutes at a speed of 3.5–4.0 mph and 0% grade on the treadmill (Trackmaster, Carrollton, TX, USA). After the warm-up phase, the speed increased to a constant based on the individualized protocol (4–7 mph) and treadmill grade increased continuously by 1% every minute until volitional exhaustion. After exhaustion was reached, the treadmill speed and grade were immediately reduced to 2.5 mph and 0% grade for a 10-minute recovery period. The verification phase was then performed at 110% of peak work rate reached during the initial bout (16). VO_{2max} was confirmed if the verification phase attained a VO_{2max} value within 3% of the incremental test (17). If the verification phase yielded a VO_{2max} which was more than 3% below the VO_{2max} value from the incremental test, subjects were required to come back and repeat their verification phase at the same intensity. If the verification phase was more than 3% above the incremental test VO_{2max} , VO_{2max} value from subjects were required to do another VO_{2max} test with both incremental and verification phases until 3% criterion was achieved. VO_{2max} from each test was determined by taking the average of the two highest consecutive 15 sec VO_2 values. Verbal

Table 2 Average VO_{2max} values from each testing method (n=30)

Subject pool	Measured VO_{2max} (mL/kg/min)	Cardio fitness score	Non-exercise prediction
Combined (mean \pm SD)	49.91 \pm 6.83	52.53 ^{*,†} \pm 8.43	48.79 \pm 6.32
Women (n=13)	45.48 \pm 5.90	46.38 \pm 5.97	44.56 \pm 3.87
Men (n=17)	53.31 \pm 5.50	57.24 \pm 6.92	52.02 \pm 5.98

*, significantly higher than measured VO_{2max} , $P=0.03$; †, significantly higher than Non-Ex, $P<0.01$.

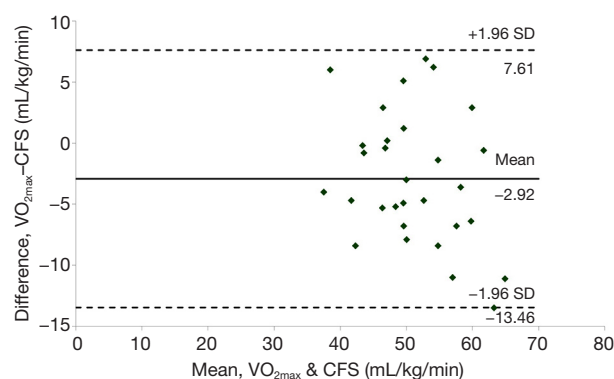


Figure 1 Bland-Altman plot of mean and difference between measured VO_{2max} and FBC2 CFS. The solid line represents the mean difference of -2.92 mL/kg/min and the dashed lines are the 95% limits of agreement. FBC2, Fitbit Charge 2; CFS, cardio fitness score.

encouragement was given throughout all laboratory VO_{2max} tests.

Data analysis

All data were analyzed using SPSS Software (SPSS 21.0; IBM Corp., Armonk, NY, USA). All data in text, tables, and figures are presented as means and standard deviations (SD) and significance was set at $P<0.05$. We tested the outcome variables for normality with the Shapiro-Wilk test to assure all variables met the assumptions of the statistical tests used. A repeated measures analysis of variance (RMANOVA) with a Bonferonni post-hoc test was used to test for differences between the three methods of VO_{2max} measurement (VO_{2max} , CFS, and N-Ex). The assumption of sphericity was tested before interpreting the results of the RMANOVA. Coefficients of variation (CVs) and mean absolute percent error (MAPE) were calculated to determine prediction accuracy of the CFS and N-Ex. Bland-Altman plots and intraclass correlation coefficients (ICCs) were used to test

for bias and consistency in VO_{2max} estimation by CFS, N-Ex, and measured VO_{2max} . Pearson correlations were used to examine the relationship between the difference in VO_{2max} measures (measured VO_{2max} – CFS) and the difference in HR measured by the FBC2 and estimated by the linear regression equations.

Results

VO_{2max} differences

There was a significant main effect for a difference in VO_{2max} across the three tests ($P<0.01$). Measured VO_{2max} was significantly lower than CFS ($VO_{2max}=49.91\pm6.83$ mL/kg/min; CFS $=52.53\pm8.43$ mL/kg/min, $P=0.03$) (Table 2). The N-Ex prediction was significantly lower than the CFS but not significantly lower than measured VO_{2max} (N-Ex $=48.79\pm6.32$ mL/kg/min; CFS *vs.* N-Ex: $P<0.01$; VO_{2max} *vs.* N-Ex: $P=0.54$). CVs were similar with CFS and N-Ex when compared to the gold standard measured VO_{2max} value (CFS $=6.5\%\pm4.1\%$; N-Ex $=5.6\%\pm3.6\%$). MAPE was larger for CFS than N-Ex when compared to VO_{2max} (CFS $=10.2\%\pm6.7\%$; N-Ex $=7.8\%\pm5.0\%$). Bland-Altman analysis indicated consistent, unbiased measurement of CFS (Figure 1). ICCs between both VO_{2max} *vs.* CFS and VO_{2max} *vs.* N-Ex were good (VO_{2max} *vs.* CFS $=0.87$, VO_{2max} *vs.* N-Ex $=0.87$).

HR differences

HR estimated by the FBC2 was lower than Est based on HR extrapolation (FBC2 $=155\pm18$ bpm, Est $=183\pm15$ bpm, $P<0.001$) (Figure 2). The difference in CFS and VO_{2max} (measured VO_{2max} – CFS) was inversely correlated with the difference in FBC2 HR and Est HR (Est HR – FBC2 HR) ($r=-0.45$, $P<0.01$) (Figure 3).

Discussion

Our study found that the FBC2 produces a consistent, unbiased estimate of VO_{2max} (CFS) while significantly overestimating VO_{2max} when compared to the gold-standard value obtained from the incremental test with verification. Interestingly, the value predicted by the N-Ex model is not significantly different from the measured VO_{2max} and therefore slightly more accurate than the FBC2 CFS in predicting VO_{2max} . This suggests that an individual who does not want to perform a maximal exercise test or purchase a FBC2 may still benefit from completing a non-

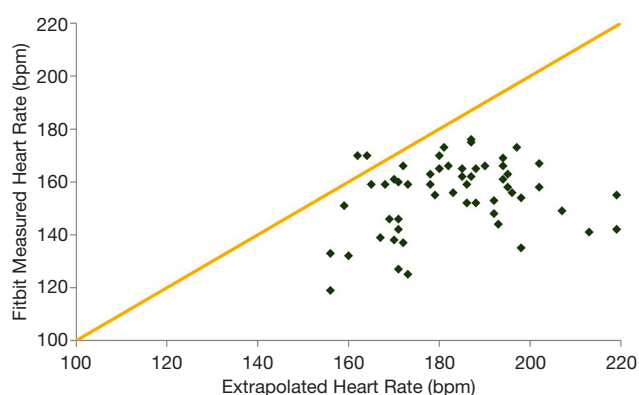


Figure 2 HR values from FBC2 and extrapolation. The line is the line of identity. FBC2, Fitbit Charge 2.

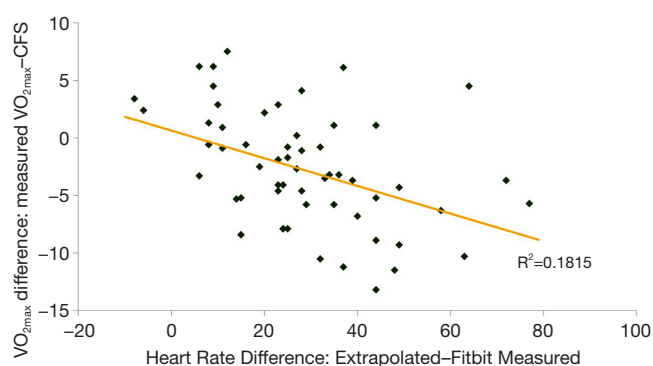


Figure 3 Differences between extrapolated HR and FBC2 Measured HR in relation to differences between measured VO_{2max} and CFS. FBC2, Fitbit Charge 2; CFS, cardio fitness score.

exercise self-reported physical activity questionnaire, which predicts VO_{2max} with good accuracy.

A similar study was performed on the Garmin Forerunner 920XT and found that the Garmin watch was highly correlated to aerobic capacity measurements obtained via open-circuit spirometry (Garmin: $r=0.84$) (5). Unlike our study, however, the Garmin watch was not significantly different from the measured aerobic capacity (5). This difference in significance between studies could be attributed to the different software and prediction equations within the watches, as Garmin uses a company called FirstBeat Technologies and Fitbit does not (18). Furthermore, the use of a HR monitor strap during the Garmin watch run may have provided more accurate HR data than was obtained from the Fitbit wrist worn HR sensor.

A recent study tested the accuracy of the Polar RS300X

fitness watch against a laboratory test of aerobic capacity (19). Eighteen college-age students completed a VO_{2max} test on the treadmill and performed a Polar fitness test (19). The Polar fitness test required that the subject report their physical activity level from the last three months based on descriptions provided by Polar, then lie supine for five minutes while the Polar HR strap recorded data (19). At the end of the test, the watch would display a VO_{2max} value based on the subject's age, height, weight, sex, activity level, maximum HR, and seated HR (19). The paired samples T-test showed no significant differences between the Polar VO_{2max} value and the metabolic cart value (Polar: 47.67 mL/kg/min *vs.* Metabolic Cart: 44.09 mL/kg/min, $P=0.111$) (19). Both the Garmin and Polar wrist worn fitness devices were not significantly different from metabolic cart values, suggesting they may be appropriate means of measuring aerobic capacity for those not requiring the accuracy of laboratory equipment.

Difficulty with the Fitbit measuring an accurate HR during runs may play a crucial role in the accuracy of the CFS (8,9). Wallen *et al.* found that among the Apple Watch, Fitbit Charge HR, Samsung Gear S and Mio Alpha, all devices underestimated HR in comparison to electrocardiography (8). However, it is important to note that these underestimations are not always clinically significant and may only reach significance under certain situations. For example, studies show that as exercise intensity increases, there is greater underestimation of HR (6,9). Our study discovered an inverse relationship between the difference in CFS and VO_{2max} and the difference in FBC2 HR and individual subject extrapolated HR. In other words, the more the FBC2 underestimated HR, the more it overestimated VO_{2max} . Thus, if the FBC2 underestimates HR during a run then it will most likely overestimate VO_{2max} , assuming the lower measured HR for a given pace is evidence of higher fitness level.

One strength of this study was that VO_{2max} testing was performed with a verification phase, the current gold standard methodology for verifying if subjects reach a "true" VO_{2max} (20). All subjects in this study verified their maximal values within 3% and were required to repeat their tests if values were not confirmed. Also, subjects performed two individual runs and an average CFS was used for statistical analyses to assess intraclass reliability and assure the subjects first run did not skew results. Subjects wore the same FBC2 for all seven days of the study and the watch was worn for at least 2 nights before subjects performed their runs in order to allow the FBC2 to get accustomed to the individual's resting HR. Although instruction was given,

subjects were not supervised and no verbal encouragement was given during their individual runs. As such, some subjects had difficulties obtaining a CFS and may have performed better if given encouragement similar to that given during the VO_{2max} test. However, these user errors are a better depiction of the general population, as the average individual would likely not have a personal trainer encouraging them and confirming proper use of the FBC2. To improve our study, measurement of HR during the 10-minute runs with a chest strap HR monitor would have been more accurate than extrapolating the data. Despite all subjects running on flat terrain, it would have been more controlled if individual runs were all recorded at a single location. Furthermore, darker skin tones and larger wrist circumferences have been associated with decreased accuracy of wearable devices (21,22), however, these data were not collected. The current study looked at the accuracy of VO_{2max} estimation by the FBC2 in a group of healthy young men and women; subject race and ethnicity were not reported. Therefore, future studies should determine the accuracy of the FBC2 for predicting VO_{2max} in adults varying in age, race, and ethnicity to enhance the generalizability of our results.

The results of our study suggest the FBC2 provides a consistent, unbiased prediction while overestimating VO_{2max} in young, healthy men and women. A non-exercise prediction equation provides a similar, slightly more accurate, VO_{2max} prediction than the CFS without the need to perform an exercise test or purchase a wearable device. The accuracy of the FBC2 CFS may be limited by its ability to correctly detect exercise HR at increased submaximal intensities.

Acknowledgments

None.

Footnote

Conflicts of Interest: The authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study was approved by the university institutional review board, all subjects provided written informed consent.

References

1. Bassett DR Jr, Howley ET. Limiting factors for maximum oxygen uptake and determinants of endurance performance. *Med Sci Sports Exerc* 2000;32:70-84.
2. Katzmarzyk PT, Church TS, Blair SN. Cardiorespiratory fitness attenuates the effects of the metabolic syndrome on all-cause and cardiovascular disease mortality in men. *Arch Intern Med* 2004;164:1092-7.
3. Arena R, Myers J, Williams MA, et al. Assessment of functional capacity in clinical and research settings: a scientific statement from the American Heart Association Committee on Exercise, Rehabilitation, and Prevention of the Council on Clinical Cardiology and the Council on Cardiovascular Nursing. *Circulation* 2007;116:329-43.
4. Gibbons L, Blair SN, Kohl HW, et al. The safety of maximal exercise testing. *Circulation* 1989;80:846-52.
5. Kraft GL, Roberts RA. Validation of the Garmin Forerunner 920 XT Fitness Watch VO 2 peak Test. *Int J Innov Educ Res* 2017;5:61-7.
6. Boudreaux BD, Hebert EP, Hollander DB, et al. Validity of Wearable Activity Monitors during Cycling and Resistance Exercise. *Med Sci Sports Exerc* 2018;50:624-33.
7. 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* 2019;37:1411-9.
8. Wallen MP, Gomersall SR, Keating SE, et al. Accuracy of Heart Rate Watches: Implications for Weight Management. *PLoS One* 2016;11:e0154420.
9. Wang R, Blackburn G, Desai M, et al. Accuracy of Wrist-Worn Heart Rate Monitors. *JAMA Cardiol* 2017;2:104-6.
10. Jackson AS, Blair SN, Mahar MT, et al. Prediction of functional aerobic capacity without exercise testing. *Med Sci Sports Exerc* 1990;22:863-70.
11. Weiglein L, Herrick J, Kirk S, et al. The 1-mile walk test is a valid predictor of VO_{2max} and is a reliable alternative fitness test to the 1.5-mile run in U.S. Air Force males. *Mil Med* 2011;176:669-73.
12. Sedgeman D, Dalleck L, Clark IE, et al. Analysis of square-wave bouts to verify VO_{2max} . *Int J Sports Med* 2013;34:1058-62.
13. McCrory MA, Gomez TD, Bernauer EM, et al. Evaluation of a new air displacement plethysmograph for measuring human body composition. *Med Sci Sports Exerc* 1995;27:1686-91.
14. Siri WE. Body composition from fluid spaces and density:

- analysis of methods 1961. *Nutrition* 1993;9:480-91; discussion 480, 492.
15. Bassett DR Jr, Howley ET, Thompson DL, et al. Validity of inspiratory and expiratory methods of measuring gas exchange with a computerized system. *J Appl Physiol* (1985) 2001;91:218-24.
 16. Midgley AW, Carroll S. Emergence of the verification phase procedure for confirming “true” VO₂max. *Scand J Med Sci Sports* 2009;19:313-22.
 17. Midgley AW, McNaughton LR, Carroll S. Verification phase as a useful tool in the determination of the maximal oxygen uptake of distance runners. *Appl Physiol Nutr Metab* 2006;31:541-8.
 18. Automated Fitness Level (VO₂max) Estimation with Heart Rate and Speed Data. [cited 2019 Aug 12]. Available online: https://assets.firstbeat.com/firstbeat/uploads/2017/06/white_paper_VO2max_30.6.2017.pdf
 19. Kraft GL, Dow M. Validation of the Polar Fitness Test. *Int J Innov Educ Res* 2018;6:27-34.
 20. Poole DC, Jones AM. Measurement of the maximum oxygen uptake Vo₂max: Vo₂peak is no longer acceptable. *J Appl Physiol* 1985 2017;122:997-1002.
 21. Menghini L, Gianfranchi E, Cellini N, et al. Stressing the accuracy: Wrist-worn wearable sensor validation over different conditions. *Psychophysiology* 2019;e13441.
 22. 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* 2017. doi: 10.3390/jpm7020003.

doi: 10.21037/mhealth.2019.09.07

Cite this article as: Freeberg KA, Baughman BR, Vickey T, Sullivan JA, Sawyer BJ. Assessing the ability of the Fitbit Charge 2 to accurately predict VO₂max. *mHealth* 2019;5:39.