



**AMERICAN COLLEGE
of SPORTS MEDICINE®**
LEADING THE WAY

... Published ahead of Print

Development and Calibration of a PATCH Device for Monitoring Children's Heart Rate and Acceleration

Bridget Armstrong¹, R. Glenn Weaver¹, Jonas McAninch², Michal T. Smith¹, Hannah Parker¹,
Abbi D. Lane¹, Yuan Wang³, Russ Pate¹, Mafruda Rahman², David Matolak²,
and M. V. S. Chandrashekhar²

¹Department of Exercise Science, University of South Carolina, Columbia, SC; ²Department of Electrical Engineering, University of South Carolina, Columbia, SC; ³Epidemiology and Biostatistics at the University of South Carolina, Columbia, SC

Accepted for Publication: 16 January 2024

Medicine & Science in Sports & Exercise®. Published ahead of Print contains articles in unedited manuscript form that have been peer reviewed and accepted for publication. This manuscript will undergo copyediting, page composition, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered that could affect the content.

Copyright © 2024 American College of Sports Medicine

Development and Calibration of a PATCH Device for Monitoring Children's Heart Rate and Acceleration

Bridget Armstrong¹, R. Glenn Weaver¹, Jonas McAninch², Michal T. Smith¹, Hannah Parker¹,
Abbi D. Lane¹, Yuan Wang³, Russ Pate¹, Mafruda Rahman², David Matolak², and M. V. S.
Chandrashekhhar²

¹Department of Exercise Science, University of South Carolina, Columbia, SC; ²Department of
Electrical Engineering, University of South Carolina, Columbia, SC; ³Epidemiology and
Biostatistics at the University of South Carolina, Columbia, SC

Address for correspondence: Bridget Armstrong, Department of Exercise Science, Arnold
School of Public Health, University of South Carolina, 921 Assembly St. RM 132, Columbia, SC
29208; Phone: 803-576-8418; E-mail: BA12@mailbox.sc.edu

Conflict of Interest and Funding Source: Funding supported by the National Institute of
General Medical Sciences of the National Institutes of Health P20GM130420 and
R21DK131387. All authors declared no conflicts of interest.

ABSTRACT

Introduction: Current wearables that collect heart rate and acceleration were not designed for children and/or do not allow access to raw signals, making them fundamentally unverifiable. This study describes the creation and calibration of an open-source multichannel platform (PATCH) designed to measure heart rate and acceleration in children ages 3-8 years. **Methods:** Children (N = 63; Mean age = 6.3) participated in a 45-minute protocol ranging in intensities from sedentary to vigorous activity. Actiheart-5 was used as a comparison measure. We calculated mean bias, mean absolute error (MAE) mean absolute percent error (MA%E), Pearson correlations and Lin's concordance correlation coefficient (CCC). **Results:** Mean bias between PATCH and Actiheart heart rate was 2.26 BPM, MAE was 6.67 BPM and M%E was 5.99%. The correlation between PATCH and Actiheart HR was .89 and CCC was .88. For acceleration, mean bias was 1.16mg and MAE was 12.24mg. The correlation between PATCH and Actiheart was .96 and CCC was .95. **Conclusions:** The PATCH demonstrated clinically acceptable accuracies to measure heart rate and acceleration compared to a research grade device.

Key Words: CHILDREN, MEASUREMENT, HEART RATE, ACCELEROMETRY, PHYSICAL ACTIVITY, WEARABLE DEVICES, PHOTOPLETHYSMOGRAPHY

INTRODUCTION

Combining heart rate and accelerometry provides greater precision in estimating health indicators such as physical activity energy expenditure (PAEE),(1-10) compared to either method in isolation. Commercial wearables (e.g., Apple Watch, Fitbits) and scientific grade sensors (e.g., Actiheart, Empatica E4, Biovotion Everion) have the capability to collect heart rate and acceleration data simultaneously. However, the relationships between heart rate, acceleration, and PAEE are dependent on growth and maturation(11), meaning that age-specific equations for PAEE prediction for children across development stages are needed. However, commercial devices such as Fitbit use proprietary algorithms to determine physical activity level and heart rate and often prevent researchers from easily accessing the raw signals.(12) As a result, any equations researchers might try to develop to estimate children's PAEE using data from such devices must be based on secondary data output (i.e. step counts and pre-processed heart rate). Equations that estimate PAEE based on this type of secondary output may become invalid because of manufacturer updates to hardware or software.

A further limitation of existing wearable devices is that both commercial wearables and research devices were designed for adults and have the potential to be distracting and uncomfortable for small children. For example, while the Actiheart demonstrates high levels of validity for energy expenditure, it is especially fragile, difficult to mend in the field and requires frequent maintenance - all of which limit its utility with children.(13) Similarly, the limited battery life of the Empatica E4 or Biovotion Everion prevents measurements over standard free-living timeframes of 7+ days. Therefore, we need an open source, unobtrusive device to accurately measure heart rate and physical activity among children in free-living conditions. This

study describes the creation and calibration of a small, open-source multichannel device designed to measure heart rate and physical activity among children ages 3-8 years.

METHODS

This study was guided by the Keadle et al. (2019)(14) framework outlining the development, calibration and validation process of wearable devices. Based on this framework, the current study is a combination of Phase 0 and Phase 1. Specifically, we focused on testing and refining the technical reliability of the PATCH monitor to measure underlying signals of heart rate and acceleration compared to existing research grade monitors. Phase 0 consisted of orbital shaker table protocols, while Phase I included controlled laboratory-based testing of selected activities using fixed start/stop times.

General system design description

The PATCH (which stands for Platform for Accurate Tracking of Children's Health) is a single device that integrates multiple passive sensor components including accelerometer and photoplethysmography (PPG). The PATCH is designed so that additional sensors (such as pulse oximetry or temperature sensing) can be included in future iterations. The PATCH is attached to the skin using a polyester spunlace adhesive designed for children.(15) PATCH is designed to be affixed to the chest to be less obtrusive than commercial wrist-based wearables in an effort to ultimately improve acceptability(16) and wear-time(17) among children. Another benefit of the chest placement is that it may reduce signal distortion from motion artifacts that typically impact PPG measurements taken from extremities like the wrist.(18)

The PATCH simultaneously records PPG and triaxial motion when mounted to the chest.⁽¹⁹⁾ Unlike commercial wearables, there is no interactive screen/watch face so that individuals will not be biased by data feedback. This decision was made to reduce observation bias and maximize ecological validity.

A key philosophy driving PATCH development is to make both hardware and software “open-source”. This includes using commercial off-the-shelf components with detailed documentation for obsolescence-tolerant hardware. For example, our PPG sensor comes with detailed CAD designs that can be replicated.⁽²⁰⁾ Our data analysis and algorithm development and documentation are performed on standard research tools (i.e., Matlab, R etc.), for transparency and ease of reproducibility across research groups to encourage uptake by the field. This open-source implementation enables a modular platform that will allow incorporation of more sensor channels and biometric data collection in future iterations of the PATCH.

Importantly, raw signal data from the PATCH are processed using transparent algorithms, alleviating the problem of proprietary algorithms with consumer wearables. Along with advances in microcontroller technology that enable multiple data channels to be collected in a single PATCH, advances in processing algorithms and advanced statistical methods allow for raw data to be processed efficiently off unit. This can result in improved battery life and allow for extended wear periods in free living environments.⁽²¹⁾

PPG is currently the standard heart rate measurement technique in low-cost consumer health trackers due to its compact footprint, a critical requirement for use with children. PPG’s

advantages over electrocardiography (ECG) are (1) lower required sample rate leading to longer battery life, (2) ease and robustness of mounting (PPG can be used wherever there is detectable blood flow), and (3) future compatibility for additional simultaneous measurements (i.e., oxygen saturation (SpO₂)).

PATCH Version 2.3.7

The components present in the PATCH device (version 2.3.7) are listed in Supplemental Table A (Supplemental Digital Content, <http://links.lww.com/MSS/C954>) along with web links where the parts were purchased. Briefly, the components include an Arduino Micro evaluation board, microSD card and triple-axis accelerometer breakout boards, and a PPG heart-rate sensor. The Arduino Micro uses a ATmega32U4 microcontroller with 32 kB flash memory, 2.5 kB SRAM, and is underclocked from a maximum of 16 MHz to 2 MHz. Underclocking, the process of reducing a microcontroller's clock frequency below its default or maximum setting, allows for more energy-efficient and controlled operation which is particularly beneficial in applications requiring power savings, precise timing, and noise reduction. The accelerometer breakout board utilizes the analog ADXL326 chip to measure triple-axis acceleration at up to ± 16 G. The PPG sensor is a Pulse Sensor Amped PPG sensor made by World Famous Electronics (https://github.com/WorldFamousElectronics/PulseSensorAmped_Hardware)(20). These components interface to measure PPG heart rate and triple-axis acceleration data, which is written to a microSD card. The system is powered by a 3.7 V li-po battery that is regulated down to 3.3 V using Micro's onboard voltage regulator. These components are placed in a 3D-printed PETG case and attached to the subject's chest during trials.

Acceleration was measured using a triple axis analog accelerometer ADXL326 with +/- 16g range powered with a regulated voltage supply $V_s = 3.3V$ supply from an Arduino micro set at a nominal 50Hz bandwidth (22). 0 g acceleration was specified at $V_s/2 = 1.65V$ with a sensitivity of 62.7mV/g giving an output range of 0.6468V for -16g to 2.6532V for +16g. This analog voltage was read by a 10-bit analog-to-digital converter (ADC) on the Arduino micro with an internal reference voltage of 2.5V, so that the ADC could code accelerations ranging from -16g to +13.56g, well beyond the +/-8g range for existing research grade accelerometers (23). The 10-bit ADC converted this analog accelerometry signal to a digital signal with 2^{10} discrete steps or 1024 levels of 2.44mV each. Given a specified accelerometer sensitivity of 62.7mV/g, we estimated a nominal resolution $2.44mV/62.7mV/g = 38.9mg$, in excellent agreement with our observations of ~40mg resolution steps during sedentary activity for a single axis (see Supplemental Figure A, Supplemental Digital Content, <http://links.lww.com/MSS/C954>, Typical variation of measured vector magnitude of acceleration during sedentary activity). Per manufacturer recommendation, the analog rms noise floor was estimated from the specified noise density of $300\mu g/\sqrt{Hz}$ and 50Hz bandwidth as $noise\ density \times \sqrt{bandwidth} * 1.6 = 3mg$, lower than the resolution of the ADC. Temperature drift was specified ~1mg/°C, giving variation as much as 17mg given body temperature 37°C over a nominal room temperature of 20°C, although the accelerometer in the version 2.7.3 casing was not in intimate contact with the skin. Our nominal designed sampling rate was the same as for PPG and was set at 86.8Hz, although for each trial, the actual sample rate was estimated by taking the total number of samples divided by the total time of the trial. For most trials, the sample rate was indeed at the designed 86.8Hz, although for 6 trials, the effective sample rate

was ~81-85Hz, which we attribute to variability in the SD card breakout board used for data acquisition.

Heart rate processing

Heart rate was estimated in both time and frequency domains and then compared for self-consistency in Matlab (24). In the time domain, clear peaks corresponding to systolic and diastolic blood pressure waves are seen. Systolic amplitudes were normalized and counted in the time domain using systole-systole time difference. In the frequency domain, we detected the peak power frequency using a short time Fourier transform (STFT) visualized using a spectrogram.(25)

Preliminary motion artifact removal. We first removed slow non-periodic motion artifacts from breathing, sweating, adhesive tension changes etc. While breathing itself is indeed a periodic physiological event, it is considered non-periodic in the context of PPG signal analysis because its variable nature can introduce non-periodic variations that interfere with the primary objective of extracting heart rate information accurately. Therefore, it is necessary to preprocess PPG signals to remove or mitigate these slow motion artifacts to enhance the accuracy and reliability of heart rate estimation. This slow baseline drift was isolated by performing a moving mean over 0.5 seconds to smooth out, and suppress the systolic and diastolic peaks, typically <0.3 seconds in duration, while preserving the other motion artifacts. Artifacts include not only cardiac-related peaks but also motion-induced noise, baseline drift, or other sources of interference, all of which are stochastic in nature and never repeating. To isolate this baseline drift, the desired HR signal must be suppressed. Once the stochastic baseline is identified, it is then subtracted from the raw

signal leaving only the HR signal. In other words, this is a time-domain high-pass filter. Mathematically, this is equivalent to creating a sequence y_n via convolution of PPG samples x with a 0.5 second (length $L \sim 45$ sample) moving average (MA) filter with impulse response (IR) h ($=\text{ones}(1,L)/L$): $y=x*h$. The MA filter has the Dirichlet frequency response $H(f)=\sin(pLf/f_s)/\sin(pf/f_s)$, which for our case has the first-null bandwidth of $f_s/45 \sim 1.9$ Hz. This baseline drift was then subtracted from the raw signal to only show the heartbeats, which are now flat with respect to time, although their relative amplitude is not constant. The subtraction is the creation of sequence $v=y*q$, where $q=d(n)-h$. Hence v is a high-pass filtered version of the original data x , $V(f)=X(f)[1-H(f)]$. This signal was then low-pass filtered at 3.5 Hz, i.e. 210 BPM to remove high frequency noise (e.g., from power lines), and smooth out the traces, while preserving the heart rate signal. The lowpass filter is a custom infinite IR (IIR) filter designed using Matlab's "lowpass" function, with rolloff parameter specified by the steepness value of 0.8. At this point in the processing, we have sequence $z=v*hlp$, with hlp the IIR filter's response.

Time Domain Heart Rate. Systolic peaks were tracked using a moving maximum function, producing sequence $w=\text{movingmax}(v,0.5)$, where the functional operation employs a 0.5-second rectangular window. Systolic peak amplitudes were then normalized, $r=w/\max(w)$, with the maximum taken over each 0.5-second interval to simplify beat counting in the time domain. Normalization produced asymmetric PPG traces typically varying between about -0.5 and $+1$, although some negative excursions are larger. The systolic peak heights as heartbeat markers were identified using a 0.5 peak amplitude threshold. For the subjects in our study this did not cause spurious diastolic detection, as the diastolic peaks were typically close to 0, or negative. From the difference between subsequent beats, the instantaneous heart rate was determined:

specifically, we compute a sequence of periods $\{T_{0,k}\}$, where k is an index, and the k^{th} period is $T_{0,k} = \text{peak}(\mathbf{r}_k) - \text{peak}(\mathbf{r}_{k-1})$, with the peak function selecting the maximum-valued sample within the k^{th} 0.5-second window. This yields a sequence of heart rates $\{HR_k\}$ in beats per minute (BPM), with $HR_k = 60/T_{0,k}$.

Outliers were removed from the heart rate sequence and replaced using Matlab's "filloutliers" method, with a 40-beat moving median window, which removed points more than 3 local scaled MADs away from the local median. Outliers occurred through sharp jolts to the sensor due to poor mounting and are described in the signal quality section below. These heart rates were then averaged over 40 beats, or ~30s. The approximately 30s window was chosen to be consistent with the 30s-time window used in the frequency domain determination of heart rate. Any outliers that were flagged are not used in the statistical calculations, with an estimated <0.1% of heart rate values being discarded. The discarded values are replaced with the value from the previous sample i.e., 11.5 milliseconds before, too short a time for heart rate to change. Furthermore, for discarded values that were replaced, correlation with the frequency domain calculation below provides an additional check on accuracy.

Frequency Domain Heart Rate. A STFT was used to identify the fundamental frequency domain content of the heartbeat from PPG data. The spectrogram is computed for the sequence $\mathbf{v}$, yielding $V(f,t)$ over 20-second windows yielding a STFT every 4.8 seconds for estimation of peak power at the heart rate frequency over time. The peak fundamental frequency powers were normalized to 1 to remove the slow changes in the heart rate amplitude over the course of the trial. This peak power frequency and peak line width were determined at each time in the

spectrogram and converted to a heart rate in BPM by multiplying by 60. To prevent spurious assignment of higher order harmonics due to “streaking” from motion artifacts,(24) or during periods of large change in heart rate such as between rest and running, we include a check to ensure that jumps in frequency $>0.8\text{Hz}$ i.e. 48BPM between successive time steps of 4.8 seconds were flagged as physiologically implausible,(26) and the previous time’s heart rate was carried forward.

The heart rate in the time domain vs frequency domain was then interpolated back to the systolic time stamps determined in the time domain above for ease of comparison between the two domains. Figure 1 shows an example of a heart rate spectrogram with good quality signal, along with the corresponding ECG heart rate and extracted frequency domain PPG heart rate overlaid with the time-domain PPG heart rate described above.

Signal Quality. We used two continuous metrics of signal quality (24) to determine the usability of each recording, self-consistency signal quality index and STD line width. Descriptions of the signal quality calculations can be found elsewhere (24) and in the Supplemental Material (Supplemental Digital Content, <http://links.lww.com/MSS/C954>).

Accelerometer processing

The tri-axial accelerometer signal was converted into a vector magnitude (VM) using the square root of the sum of x, y and z axis accelerations squared. After measurement of each axis, X, Y, and Z independently, with acceleration a_x, a_y, a_z the measured bit value from the ADC,

read back from the SD card storage, was converted back to the corresponding voltage by multiplying by 2.5V/1024bits in post-processing. The original acceleration in g's was obtained by subtracting 1.65V (the 0g offset) and then dividing by the specified sensitivity of 62.7mV/g. The acceleration VM e was then calculated by:

$$|a| = \sqrt{a_x^2 + a_y^2 + a_z^2}$$

From this value, we subtract 1g and replaced negative values with zero to obtain Euclidean Norm Minus One (ENMO) (27). Similar ENMO calculations are performed for Actiheart.

There are unavoidable small drifts in supply voltage for a single device as well as device to device variations from part variations during manufacturing. Such variations are inevitable in battery powered mobile and wearable electronics. The tolerance corresponding to this variation is specified at 1% (28) or ~33mV for $V_s=3.3V$ and leads to error as large as $33mV/(sensitivity=62.7mV/g) = 0.5g$ or ~250mg for the 0g condition at the nominal $V_s/2=1.65V$ range. This is the critical range where sedentary activity can be easily misinterpreted as light or vice versa, rendering the data unusable at low activity levels. Fortunately, the sensitivity of the ADXL326 is also proportional to the supply voltage, so that the sensitivity only changes by ~1%, giving $|a|$ and ENMO *variations* that are insensitive to V_s . In other words, the accelerometer sensitivity is the key robust specification we use to obtain reliable data.

To eliminate the impact of V_S drift and device-to-device variability, we took the numerical derivative of the raw measured $|a|$ giving the ‘jerk’, or rate of variation of $|a|$.⁽²⁹⁾ The derivative of a constant number is 0, thus eliminating the baseline drift in the electronics while measuring earth’s gravity of 1g during sedentary activity levels. We then took the numerical integral of this derivative to recover the original signal with the baseline drift in the measured gravity of earth removed i.e., the ENMO with the negative values still present. The negative values are then set to zero for subsequent analysis, in line with current accelerometer processing.⁽³⁰⁾ The same derivative/integral approach followed by setting negative values to zero was performed for the criterion Actiheart. Finally, data were smoothed using centered moving average with a five second window that progressed by one reading each step. Data were then aggregated at the 5 second level, consistent with previous research. ⁽³⁰⁾ We converted ENMO from G’s to mg (by multiplying ENMO values by 1,000) for ease of interpretation.

Research Grade Comparison Measure – Actiheart 5

The PATCH device was compared to a research grade device, the Actiheart 5 (Actiheart, CamNTECH Ltd.). Actiheart 5 is a single-lead ECG device that also includes a tri-axial accelerometer. Actiheart was worn on the chest using an Actiheart brand elasticated soft electrode strap. Electrodes were wet before the start of each protocol. Actiheart ECG sampling frequency ranges from 128-1024 Hz. The ECG sampling rate in the current study was set to 256 Hz. Actiheart 5 software was used to process the raw ECG signal and produce heart rate estimates at the one second level. The Actiheart accelerometer has dynamic range between +/- 8 g and a sampling frequency between 30–100 Hz. For the current study, the accelerometer sampling frequency was set to 50 Hz.

Data Alignment

For initial coarse alignment, PATCH starting timestamps were manually recorded by the protocol coordinator using an online atomic clock when the device was turned on for the trial. The data from PATCH activation to protocol start was ignored for subsequent analysis (see Matlab scripts https://github.com/ACOI-UofSC/PATCH_2.0/tree/MSSE-2023), as was done for Actiheart. The time difference between Actiheart turn-on and PATCH turn-on is specified in the Matlab script. Thus, the Actiheart and PATCH data are coarsely aligned within the uncertainties of (i) time delay between PATCH turn-on and recording timestamp ~ 2 seconds, and (ii) drift between PATCH and Actiheart clocks that is variable across devices. Visual inspection of large jumps in heart rate readings between physical activity levels in the protocol was used to verify this coarse alignment. For heart rate, this delay is reasonable, particularly given that frequency domain STFT are computed every 4.8s. However, this apparently small error is unacceptable for accelerometry as spikes in acceleration from steps are < 1 second. In fact, without further refined alignment for accelerometry, an incorrect conclusion of poor correlation in ENMO between the 2 devices can arise due to a small $< 0.1\%$ drift over the course of the trial, even if visual inspection of the raw signals shows identical signal shape.

We used an additional fine alignment procedure to correct for small drifts between PATCH and Actiheart signals. A cross-correlation between the 2 accelerometry signals was performed on the jerk (i.e., derivative of the VM as described above). Here, the Actiheart accelerometer sampling rate was fixed at the nominal 50Hz, while the PATCH sample rate (fs_{PPG} in Matlab script) was changed from its measured value in steps of 0.01Hz while a cross-correlation between $EN_{Actiheart}$ and EN_{PATCH} was performed. To directly compare the two signals

at different sample rates, we interpolated the higher sample rate PATCH signal down to the lower sample rate Actiheart accelerometer signal at its nominal 50Hz rate giving EN_{PATCH} (interp). A linear regression between 5 second moving window RMS $EN_{Actiheart}$ and EN_{PATCH} (interp) was performed to maximize the regression R^2 value, while the cross-correlation time delay between signals was allowed to freely float. Typical refinements were changes in relative sample rates of $\sim 0.1\text{Hz}$ i.e. $\sim 0.1\%$ and time delays ~ 10 seconds leading to improvements in regression R^2 to >0.9 . This refinement in time alignment did not significantly impact heart rate alignment and corresponding heart rate quality metrics (as heart rate changes over seconds, not milliseconds), underscoring the robustness of our time alignment protocols during the trials. Once fine alignment was performed, the protocol data was broken up by target physical activity level using second level times recorded during the protocol with Actiheart and PATCH timestamps corresponding to the closest match to the recorded time. The error from this “closest match” is ~ 1 sample, or ~ 20 milliseconds for Actiheart and ~ 11.5 milliseconds for PATCH, more than sufficient for heart rate comparisons.

Shaker Table Study (Phase 0 validation)

Shaker Table Protocol. Intra-device reliability and validity of raw accelerations for all PATCH devices were tested, with the speed of a mechanical shaker table (Scientific Industries, Bohemia, NY, USA; Mini-300 Orbital-Genie, Model 1500) as the criterion. Each device was securely mounted directly to the twin ratcheting clamps of a mechanical shaker table that produces controlled oscillations at frequencies between approximately $f_{shaker}=0.6$ and 5 Hz. We converted f_{shaker} in Hz to acceleration using the expression for centripetal acceleration, $a_{orbital} = v^2/r_{orbital}$, (31) where $r_{orbital}$ is the radius of rotation for the orbital shaker $r_{orbital}$.

From the manual for this particular shaker,(32) the specified diameter of the orbit is $2r_{orbital}=1.9\text{cm}$ and the rotational speed is given by $v = 2\pi r_{orbital}f_{shaker}$, since for each complete cycle of 2π radians, the table traverses a distance of circumference $2\pi r_{orbital}$ in time $1/f_{shaker}$. In other words

$$a_{orbital}(\text{cm}/\text{s}^2) = 4\pi^2 r_{orbital} f_{shaker}^2$$

To convert this acceleration to units of earth's gravity in g's, divide $a_{orbital}$ by $9.81\text{m}/\text{s}^2$.

For reliability and validity testing, five identical devices were mounted. Five 2-minute trials were conducted. There was a one-minute rest period between each oscillation. Consistent with past studies,(33, 34) each trial began with a 60 second rest period (i.e., no movement) followed by a standardized series of oscillations at three f_{shaker} (2.4 Hz, 2.8 Hz, and 3.2 Hz) lasting two minutes each. These frequencies were chosen because they are consistent with human movement ranging from 1.5 to 16 mph.(35) The start and stop time were noted at each frequency. Following all testing, raw acceleration data from the PATCH devices was downloaded. Data were aligned using cross-correlation and visual inspection. Only the middle minute was used to avoid trial contamination or edge effects.(36)

Semi-Structured Study with Children (Phase 1 validation)

Participants. We recruited a total sample of 63 children between the ages of 3-5 (mean age = 6.3 years, ± 1.7 , demographics in Table 1). This study used a convenience sampling approach. Children between the ages of 3-8 years were recruited from the local community through a

combination of fliers, university newsletters, word of mouth and targeted social media advertising. Inclusion criteria for children in the study were being between the ages of 3-8 at enrollment and being able to ambulate without assistive devices (i.e., wheelchair). Children were excluded from the study if a doctor said they should refrain from exercise, which was assessed using parent report prior to study enrollment. The study was conducted in accordance with the Helsinki Declaration and approved by the University of South Carolina IRB in September 2021 (Pro00114456). Interested parents completed an online eligibility survey and consent form. Parents who completed the online consent were then contacted by phone to verbally confirm their desire to participate and schedule an in-person session. Verbal assent was obtained from children prior to any data collection. Protocols were conducted at the university of South Carolina either indoors or outdoors depending on availability and weather. Parents were instructed to have their children avoid caffeine on the day of the assessment.

Semi-Structured Child Protocol. Children completed a semi-structured series of activities ranging in intensity from sedentary (i.e., lying down, sitting, and standing) to light (i.e., walking at a self-selected pace) to vigorous (i.e., running) (see Figure 1 and Supplemental Table B, Supplemental Digital Content, <http://links.lww.com/MSS/C954>, Protocol tasks).(37) A table of protocol activities is presented in Supplemental Table B (Supplemental Digital Content, <http://links.lww.com/MSS/C954>). The entire protocol was designed to last no longer than ~45 minutes. A 5-minute seated session between the light/moderate and the moderate/vigorous activities was used to ensure heart rate returned to a resting state prior to engaging in the subsequent higher intensity physical activity. The protocol was designed such that no more than 5 continuous minutes would be spent running semi-structured activities. This shorter length was

selected given concerns about children's ability to sustain these intensities for longer time periods.

Statistical Analysis

Shaker Table Study (Phase 0 validation). For the shaker table portion of the study, we calculated the ENMO value for each sample. In line with current accelerometer processing,(38, 39) we then calculated the average ENMO across each minute for each trial and speed. Prior to statistical analyses, descriptive means and standard deviations for the ENMO mean were calculated across devices at each oscillation frequency. To test reliability, single, absolute intraclass correlation coefficient (ICC) was calculated across all 5 patch devices. ICC values less than 0.50 were defined as poor reliability, between 0.50 and 0.75 as moderate reliability, between 0.75 and 0.90 as good reliability, and greater than 0.90 as excellent reliability.(40) For the validity testing, Pearson product moment (r), and Lin's CCC were calculated to assess correlation and agreement between the PATCH compared to the criterion (i.e., speed of the shaker table). We used CCC because it measures precision and accuracy. While the Pearson correlation coefficient measures the linear relationship between two variables, it does not directly capture differences in the intercept and slope. Lin's concordance coefficient, on the other hand, considers both the precision (how closely the data points cluster around the line of perfect agreement) and accuracy (how well the line fits the data in terms of slope and intercept). Correlation coefficients were interpreted based on recommendations from Altman (1991), with coefficients less than 0.20 as poor and greater than 0.80 as excellent.(41) We calculated both mean error and mean absolute error (MAE). Bland-Altman plots were used to visually examine the trend and directionality of the mean bias between PATCH against the speed of the shaker

table. The Formula for MAE is presented below where x_i and y_i are the respective PATCH and shaker table estimated ENMO at the i th aligned time point.

$$MAE = \frac{1}{n} \sum_{i=1}^n |x_i - y_i|$$

Semi-structured Child Protocol (Phase I validation). For heart rate, in line with similar studies,(42) we only used the middle minute (seconds 240 to 360) of the 5 minute sessions to retain steady state readings from the devices. For the 10-minute supine resting portion, the 6th minute was selected. To examine heart rate accuracy between the Actiheart and PATCH measures, we calculated mean bias and MAE between PATCH calculated heart rate and Actiheart ECG heart rate. We also calculated the MA%E defined as the MAE divided by the Actiheart value using the formula below:

$$MAE_{perc} = \frac{1}{n} \sum_{i=1}^n |x_i - y_i| / x_i$$

where x_i and y_i are the Patch and Actiheart estimated heart rate at the i th aligned time point. The agreement between PATCH and Actiheart heart rate and acceleration was quantified using mean bias, upper and lower limits of agreement with 95% confidence intervals using Bland-Altman plots.

As a metric of feasibility, we calculated data coverage for the PATCH and Actiheart for both acceleration and heart rate estimation. Data coverage refers to the extent to which a given

device has data points recorded and available at each epoch or time interval, ensuring a comprehensive representation of information throughout the entire duration of data collection or monitoring. For HR, data were aggregated at the second level and were counted as missing if they did not have a recorded HR value for that second. Data coverage for accelerometry was conducted at the 50hz level, and were counted as missing if they did not have a value recorded. Data coverage across all trials was calculated as the percentage of epochs that contained a value (either HR or acceleration). We present the Actiheart coverage to provide a comparison to PATCH on the typical level of signal drop and data loss seen in currently available research grade devices that measure both heart rate and acceleration. For acceleration, we calculated mean bias (with Bland-Altman), MAE and MA%E for accelerometry calibration between Actiheart and PATCH. We also computed Pearson correlations with 95% confidence intervals as well as Lin's CCC.

To test if HR and acceleration estimates were equivalent between the PATCH and Actiheart, we used the TOSTER (version 0.4.6)(43) to conduct two, one-sided tests of equivalence (TOST) (44). We used 90% CIs around the mean differences for both acceleration and HR. If the confidence interval does not overlap or exceed the equivalence bounds, then the monitors are considered equivalent ($p < 0.05$). (36). For HR equivalence bounds were set as 5%.(45) For acceleration, based on previous research and best practice, (23, 46) bounds were set at 10% of the mean value of the ActiHeart. HR data were analyzed at the epoch level (1-minute) separately by activity. Acceleration data were analyzed at the epoch level (5-second) across the whole trial to avoid issues with restricted variability within each activity.

Sample Size Calculation. We anticipated an intra-class correlation (ICC) of 0.8 between the PATCH and criterion/comparison devices. An a priori power calculation showed that a sample size of 52 provides a 95%CI from 0.701 (minimal acceptable reliability) to 0.899 based on the sample size estimation formula for precise ICC estimation.(47, 48)

RESULTS

Shaker Table Study (Phase 0 validation)

In terms of reliability, the ICC was .980 between all PATCH devices (95% CI .929 to .993). For validity, the Pearson correlation between all patch readings and the criterion shaker speed was .915 (95% CI .875 to .942). The Lin's CCC between PATCH and the criterion was .868 (95% CI .822 to .903) (see Supplemental Figure B, Supplemental Digital Content, <http://links.lww.com/MSS/C954>). MAE was 21.55 mg (95% CI 2.25 to 79.43). The sample mean bias was 9.1 mg (95%CI -26.95 to 79.53) and Bland-Altman plot showed a trend such that the error was higher at faster speeds ($B = 0.207$, $p < .05$; see Figure 2)

Semi-Structured Study with Children (Phase 1 validation)

Participant demographics. A participant flow diagram is presented in Supplemental Figure C (Supplemental Digital Content, <http://links.lww.com/MSS/C954>). Three trials were excluded due to equipment malfunction where no data could be collected. One trial was excluded because it was incomplete due to the child no longer wishing to participate. Three trials were excluded for heart rate analyses due to equipment malfunction with the PATCH device that resulted in unusable PPG signal quality. The decision to exclude a trial was made based on the signal quality indices along with expert interpretation of spectrogram signal, in line with best practice

for determining PPG signal quality.(49) The final analytic sample included 56 children (see Supplemental Figure C, Supplemental Digital Content, <http://links.lww.com/MSS/C954>).

PPG Signal Quality & Data Coverage. Of the 56 protocols included in the analytic sample, PPG signal quality metrics ranged from 0.34 to 0.97 (median = 0.81, Inner quartile range [IQR] 0.75 to 0.87) for self-consistency and 1.86 to 9.81 for STD line width (median 5.27, IQR 4.47 to 6.31). Individual level signal quality metrics are presented in Supplemental Table C Supplemental Digital Content, <http://links.lww.com/MSS/C954>). Heart Rate data coverage for Actiheart was 89%, whereas it was 99% for PATCH. Visual inspection revealed that in 38% of protocols there was some degree of signal drop out from the Actiheart device (21 of 56). Data coverage for acceleration was greater than 99% for both PATCH (99.2%) and Actiheart (99.7%).

Heart rate accuracy. Average (median and mean) bias, MAE and M%E overall and separated by each protocol task are presented in Table 2. Across all protocol tasks, the sample median MAE was 0.46 beats per minute (interquartile range 2.09 to 3.77) compared to ActiHeart ECG. The Bland-Altman plot and 95% limits of agreement at the minute level are presented in Figure 3. There was a significant linear trend ($B = -0.07$, $p < .01$) such that PATCH underestimated heart rate at higher heart rate values.

The overall median MAE value across the sample was 2.86 (Interquartile range 1.22 to 5.72). The sample median MA%E was 2.64% (Interquartile range 1.16% to 5.33%) (see Table 2). Pearson correlation between PATCH and Actiheart was .887 (95% CI .883 to .890). Lin's CCC was .883 (95%CI .880 to .887) indicating strong absolute agreement between PATCH and

Actiheart estimates of heart rate (41) (see Supplemental Figure D, Supplemental Digital Content, <http://links.lww.com/MSS/C954>). The largest disagreement was observed during running activities, and the smallest was observed during standing (see Table 2). Individual person level heart rate accuracy is available in Supplemental Table C (Supplemental Digital Content, <http://links.lww.com/MSS/C954>).

Heart rate equivalence. We conducted equivalence tests between HR estimates from the PATCH and Actiheart. Figure 4 shows the mean difference (along with 90% CI), and equivalence bounds for the Actiheart. For each of the activities, The TOST procedure indicated that the observed difference and 90% CI were completely within the equivalence bounds the Actiheart HR at each activity level ($p < .01$).

Acceleration accuracy. Average (median and mean) bias and MAE separated by each protocol task are presented in Table 3. Compared to Actiheart, the mean bias was 1.61 mg (interquartile range 0.73 to 7.11). Across all protocol tasks, MAE was 12.24 mg (interquartile range 3.48 to 9.73). The Bland-Altman plot and 95% limits of agreement are presented in Figure 5. Pearson correlation between PATCH and Actiheart was .953 (95% CI .952 to .954). Lin's CCC was .952 (95%CI .951 to .953) indicating strong absolute agreement between PATCH and Actiheart estimates of acceleration (41) (see Supplemental Figure E, Supplemental Digital Content, <http://links.lww.com/MSS/C954>).

Acceleration equivalence. We conducted equivalence tests between ENMO estimates from PATCH and Actiheart. Figure 4 shows the mean difference and 90% CI, and equivalence bounds

for PATCH and Actiheart. Across all activities, The TOST procedure indicated that the 90% CI of the difference in ENMO was completely within the equivalence bounds the Actiheart ($p < .05$).

DISCUSSION

The current study found that the PATCH, a novel device with open-source hardware and processing, was capable of measuring heart rate and acceleration in a semi-structured protocol of children aged 3-8 years old.

Heart Rate

Mean bias for heart rate was 2.26 BPM, less than the widely used validity criterion of < 3 BPM.⁽⁵⁰⁾ Additionally, our standard error of the mean bias was 2.03 BPM (SE = $14.79/\sqrt{56}$) which is less than previously established ≤ 5 BPM used in previous studies.^(50, 51) We did find that errors were more substantial at higher heart rates, which is likely due to motion artifacts when children are running.⁽⁵²⁾ Regardless, PATCH estimates of HR were equivalent to Actiheart for each level of activity.

One approach to address motion artifacts is to handle them during post-processing, where algorithms are used to distinguish between good and bad signal. This can involve filtering or interpolation techniques to remove or minimize the effects of motion artifacts on the PPG signal. However, this approach may have limitations in terms of accuracy and reliability, as it relies on algorithms that may not always accurately detect and correct motion artifacts. Another strategy to mitigate motion artifacts in PPG measurements is to consider the device placement and adhesive strategy.⁽⁵³⁾ Properly securing the PPG sensor or device to the subject's skin with the

correct amount of pressure (54) can reduce the interference of motion artifacts. This may involve using appropriate adhesive materials or mounting techniques to minimize movement of the PPG sensor during data acquisition.

Acceleration - Shaker Table

The intra-device reliability ICC values for the PATCH in the shaker table study were comparable to similar studies examining the intra-device reliability for research grade accelerometers.(55) Also consistent with similar studies, our results showed a trend for higher error at higher speeds.(34) However, it is worth noting the magnitude of these differences. For context, the mean bias between PATCH and the criterion shaker speed was 9.1mg, which is less than the intra-device bias reported for research grade devices including the widely used Actigraph GT3X (20mg) and Genea (110mg).(34) Overall, the results of this study suggest that the PATCH accelerometer has comparable intra-device reliability to other research grade accelerometers, with only minimal differences observed in bias at higher speeds.

Acceleration - Semi-Structured Child Protocol

There was a strong correlation between Actiheart and PATCH acceleration, as well as strong absolute agreement. Average mean bias was 1.61mg and MAE was 12.24mg. Furthermore, the PATCH ENMO values were equivalent to Actiheart across the aggregated trial. However, we did find higher errors at higher acceleration (running MAE was 39.53mg). For context, the Hildebrand waist-based intensity threshold(38, 39) between sedentary and light physical activity is 40mg. However, the threshold between light and moderate is 140, and moderate to vigorous is 456. Thus, we should still retain the ability to distinguish between

minimal physical activity and high intensity physical activity. However, it is worth noting that currently there are no validated intensity thresholds for chest mounted acceleration data among children or adults.

It is not uncommon to find differences between raw signal data from different types of accelerometers. John et al. (34) found that mean differences between two research grade accelerometers (GTX3 and GENEa) on a highly controlled shaker table ranged from -0.001 to 0.242, an average of 76mg - which is substantially higher than the average error in the current study between Actiheart and PATCH (1.6mg). Similarly, the error observed in human studies comparing different accelerometers has been reported to be as much as 54.6mg.(30) Even studies of different generations of the same accelerometer have demonstrated substantial differences that can translate to as much as 70kcal/day differences in energy expenditure estimation. (56) Ultimately, these discrepancies may indicate that cut point intensity thresholds are not truly device agnostic as proposed by Hildebrand et al,(38, 39) even with access to raw accelerometer data. Thus, alternative processing methods such as physical activity recognition, may prove more useful than threshold criteria when predicting energy expenditure.(57)

Predicting Physical Activity Energy Expenditure among Children

The next steps of this research line are to develop and validate energy expenditure equations for children, using the patch inputs of heart rate and acceleration. This is based on evidence that a combination of heart rate and acceleration can provide better estimates of energy expenditure compared to either metric alone.(1-10)

There is a clear need for devices designed to specifically measure children's energy expenditure. Currently, there are few devices capable of capturing both acceleration and heart rate. Devices that do exist are often proprietary and designed for adults. Furthermore, the few research grade devices that do allow some degree of access to raw data such as the Actiheart have additional challenges including device failure and dropout. Indeed, over 88% of Actiheart data were missing for heart rate, and device dropout occurred in 21 of 56 protocols. While we acknowledge that we did have 3 cases which produced unusable PPG data (3 of 60; 5%), the PATCH technology is still under development and some degree of device failure is expected. By contrast, the Actiheart is an established research grade monitor, and one of the only options currently available that allows access to raw data. Given the suboptimal data quality in a well-controlled laboratory setting, it seems likely that free-living data quality would also suffer, possibly to a greater degree.⁽⁵⁸⁾ Indeed, other studies have concluded that Actiheart is simply too fragile and logistically complicated to be used in free living settings with children.⁽¹³⁾ Taken together, it seems unlikely that public health researchers will adopt the Actiheart technology at a large scale – thus limiting its public health potential. By contrast, the PATCH technology is designed to be both open source in terms of hardware as well as software processing. This open-source approach for a device specifically designed for children has the potential to make a meaningful impact on the field of energy expenditure research and inform public health guidelines.

Strengths and Limitations

As noted, the current study did have three device protocols that produced unusable data. Additionally, further research and optimization can be done to improve PATCH signal quality.

An unfortunate limitation of this study was the device failure of the Actiheart. Although we used a bandpass filter, there were multiple trials in which the Actiheart produced unlikely heart rate estimates, which likely skewed the agreement metrics for the PATCH device. While the Actiheart was selected as a criterion in order to allow children to ambulate freely, a more cumbersome but accurate measure, such as a Holter monitor, may have produced better agreement.

This study boasts several key strengths. Firstly, the study design for device validation was based on the Keadle et al (2019)(14) framework, which is considered a best practice approach for device validation. By adhering to this rigorous framework, the study ensured that its results were robust and reliable. Additionally, the use of open-source hardware and software adds to the study's transparency and reproducibility, allowing for others to validate and build upon our findings. Furthermore, by examining acceleration and heart rate data at the within-person level using timeseries data approaches, the study was able to avoid the common practice of aggregating data at the protocol level, which has been shown to lead to biased results. This approach aligns with recent recommendations in the field (59) and allows for a more nuanced understanding of the data. Together, these strengths lend credibility to the study's conclusions and enhance its impact on the field.

CONCLUSIONS

The integration of physical activity and heart rate signals could lead to more accurate estimates of energy expenditure and 24-hour movement behaviors, while also laying the foundation for sleep staging through the derivation of heart rate variability. Overall, this proof-

of-concept calibration study demonstrates the validity of a novel open-source wearable device that can collect accurate data on heart rate and physical activity in children aged 3-8 years old.

ACCEPTED

Acknowledgments

Funding supported by the National Institute of General Medical Sciences of the National Institutes of Health P20GM130420 and R21DK131387. All authors declared no conflicts of interest. The results of the study are presented clearly, honestly, and without fabrication, falsification, or inappropriate data manipulation. The results of the present study do not constitute endorsement by the American College of Sports Medicine.

REFERENCES

1. Brage S, Brage N, Franks P, Ekelund U, Wareham N. Reliability and validity of the combined heart rate and movement sensor Actiheart. *Eur J Clin Nutr*. 2005;59(4):561-70.
2. Brage S, Brage N, Franks PW, et al. Branched equation modeling of simultaneous accelerometry and heart rate monitoring improves estimate of directly measured physical activity energy expenditure. *J Appl Physiol (1985)*. 2004;96(1):343-51.
3. Corder K, Brage S, Wareham NJ, Ekelund U. Comparison of PAEE from combined and separate heart rate and movement models in children. *Med Sci Sports Exerc*. 2005;37(10):1761-7.
4. Eston RG, Rowlands AV, Ingledew DK. Validity of heart rate, pedometry, and accelerometry for predicting the energy cost of children's activities. *J Appl Physiol (1985)*. 1998;84(1):362-71.
5. Sánchez-López M, Pardo-Guijarro MJ, del Campo DG-D, et al. Physical activity intervention (Movi-Kids) on improving academic achievement and adiposity in preschoolers with or without attention deficit hyperactivity disorder: study protocol for a randomized controlled trial. *Trials*. 2015;16:456.
6. Strath SJ, Bassett DR, Swartz AM, Thompson DL. Simultaneous heart rate-motion sensor technique to estimate energy expenditure. *Med Sci Sports Exerc*. 2001;33(12):2118-23.
7. Villars C, Bergouignan A, Dugas J, et al. Validity of combining heart rate and uniaxial acceleration to measure free-living physical activity energy expenditure in young men. *J Appl Physiol (1985)*. 2012;113(11):1763-71.

8. Zakeri IF, Adolph AL, Puyau MR, Vohra FA, Butte NF. Cross-sectional time series and multivariate adaptive regression splines models using accelerometry and heart rate predict energy expenditure of preschoolers. *J Nutr.* 2013;143(1):114-22.
9. de Zambotti M, Baker FC, Willoughby AR, et al. Measures of sleep and cardiac functioning during sleep using a multi-sensory commercially-available wristband in adolescents. *Physiol Behav.* 2016;158:143-9.
10. de Zambotti M, Goldstone A, Claudatos S, Colrain IM, Baker FC. A validation study of Fitbit Charge 2 compared with polysomnography in adults. *Chronobiol Int.* 2018;35(4):465-76.
11. Malina RM, Bouchard C, Bar-Or O. *Growth, Maturation, and Physical Activity.* Champaign, IL:: Human Kinetics; 2004.
12. Chen KY, Janz KF, Zhu W, Brychta RJ. Re-defining the roles of sensors in objective physical activity monitoring. *Med Sci Sports Exerc.* 2012;44(1 Suppl 1):S13-23.
13. Wilson H, Dickinson F, Griffiths P, Bogin B, Varela-silva MI. Logistics of using the Actiheart physical activity monitors in urban Mexico among 7-to 9-year-old children. *Am J Hum Biol.* 2011;23(3):426-8.
14. Keadle SK, Lyden KA, Strath SJ, Staudenmayer JW, Freedson PS. A framework to evaluate devices that assess physical behavior. *Exerc Sport Sci Rev.* 2019;47(4):206-14.
15. Massa GG, Gys I, Op't Eyndt A, et al. Evaluation of the FreeStyle® Libre flash glucose monitoring system in children and adolescents with type 1 diabetes. *Horm Res Paediatr.* 2018;89(3):189-99.
16. Schneller MB, Bentsen P, Nielsen G, et al. Measuring children's physical activity: compliance using skin-taped accelerometers. *Med Sci Sports Exerc.* 2017;49(6):1261-9.

17. Duncan S, Stewart T, Mackay L, et al. Wear-time compliance with a dual-accelerometer system for capturing 24-h behavioural profiles in children and adults. *Int J Environ Res Public Health*. 2018;15(7):1296.
18. Bent B, Goldstein BA, Kibbe WA, Dunn JP. Investigating sources of inaccuracy in wearable optical heart rate sensors. *NPJ Digit Med*. 2020;3:18.
19. Godfrey A, Bourke A, O'laighin G, Van De Ven P, Nelson J. Activity classification using a single chest mounted tri-axial accelerometer. *Med Eng Phys*. 2011;33(9):1127-35.
20. Open Hardware: World Famous Electronics llc; [Available from: <https://pulsesensor.com/pages/open-hardware>.
21. Williamson J, Liu Q, Lu F, et al., editors. Data sensing and analysis: challenges for wearables. The 20th Asia and South Pacific Design Automation Conference; 2015: IEEE.
22. One Technology. Small, Low Power, 3-Axis ± 16 g Accelerometer ADXL326 Norwood, MA: Analog Devices, Inc; 2009 [Available from: <https://www.analog.com/media/en/technical-documentation/data-sheets/ADXL326.pdf>.
23. Rowlands AV, Mirkes EM, Yates T, et al. Accelerometer-assessed physical activity in epidemiology: are monitors equivalent? *Med Sci Sports Exerc*. 2018;50(2):257-65.
24. McLean MK, Weaver RG, Lane A, et al. A sliding scale signal quality metric of photoplethysmography applicable to measuring heart rate across clinical contexts with chest mounting as a case study. *Sensors (Basel)*. 2023;23(7):3429.
25. Singh O, Sunkaria RK. A unified approach for heart rate estimation from electrocardiogram and arterial blood pressure pulses. *Adv Sci Eng Med*. 2020;12(5):588-92.

26. Dao D, Salehizadeh SM, Noh Y, et al. A robust motion artifact detection algorithm for accurate detection of heart rates from photoplethysmographic signals using time–frequency spectral features. *IEEE J Biomed Health Inform.* 2016;21(5):1242-53.
27. Van Hees VT, Gorzelniak L, Dean León EC, et al. Separating movement and gravity components in an acceleration signal and implications for the assessment of human daily physical activity. *PLoS One.* 2013;8(4):e61691.
28. Texas Instruments. LP2985 150-mA Low-Noise Low-Dropout Regulator with Shutdown SLVS522O. Technical Report. Dallas, TX: Texas Instruments. ; 2023.
29. Ode KL, Shi S, Katori M, et al. A jerk-based algorithm ACCEL for the accurate classification of sleep–wake states from arm acceleration. *iScience.* 2022;25(2):103727.
30. Clevenger KA, Pfeiffer KA, Montoye AH. Cross-generational comparability of raw and count-based metrics from ActiGraph GT9X and wGT3X-BT accelerometers during free-living in youth. *Meas Phys Educ Exerc Sci.* 2020;24(3):194-204.
31. Halliday D, Resnick R, Walker J. *Fundamentals of physics.* John Wiley & Sons; 2013.
32. MINI ORBITAL-GENIE™ Variable Speed Benchtop Orbital Shaker Operating Instructions: Scientific Industries, Inc.; 2022 [Available from: https://www.scientificindustries.com/media/file/file/l/t/ltp0087-mini_orbital_genie-rev_b.pdf.
33. Davoudi A, Wanigatunga AA, Kheirkhahan M, et al. Accuracy of Samsung Gear S Smartwatch for activity recognition: validation study. *JMIR Mhealth Uhealth.* 2019;7(2):e11270.

34. John D, Sasaki J, Staudenmayer J, Mavilia M, Freedson PS. Comparison of raw acceleration from the GENEa and ActiGraph™ GT3X+ activity monitors. *Sensors (Basel)*. 2013;13(11):14754-63.
35. John D, Miller R, Kozey-Keadle S, Caldwell G, Freedson P. Biomechanical examination of the 'plateau phenomenon' in ActiGraph vertical activity counts. *Physiol Meas*. 2012;33(2):219-30.
36. Dixon PM, Saint-Maurice PF, Kim Y, Hibbing P, Bai Y, Welk GJ. A primer on the use of equivalence testing for evaluating measurement agreement. *Med Sci Sports Exerc*. 2018;50(4):837-45.
37. Burke D, Sundlof G, Wallin BG. Postural effects on muscle nerve sympathetic activity in man. *J Physiol*. 1977;272(2):399-414.
38. Hildebrand M, VAN Hees VT, Hansen BH, Ekelund U. Age group comparability of raw accelerometer output from wrist-and hip-worn monitors. *Med Sci Sports Exerc*. 2014;46(9):1816-24.
39. Hildebrand M, Hansen BH, van Hees VT, Ekelund U. Evaluation of raw acceleration sedentary thresholds in children and adults. *Scand J Med Sci Sports*. 2017;27(12):1814-23.
40. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropr Med*. 2016;15(2):155-63.
41. Altman DG. *Practical Statistics for Medical Research*. CRC press; 1990.
42. Saint-Maurice PF, Kim Y, Welk GJ, Gaesser GA. Kids are not little adults: what MET threshold captures sedentary behavior in children? *Eur J Appl Physiol*. 2016;116(1):29-38.

43. Lakens D. Equivalence tests: a practical primer for t tests, correlations, and meta-analyses. *Soc Psychol Personal Sci*. 2017;8(4):355-62.
44. Schuirmann DJ. A comparison of the two one-sided tests procedure and the power approach for assessing the equivalence of average bioavailability. *J Pharmacokinet Biopharm*. 1987;15(6):657-80.
45. 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;7(2):3.
46. Welk GJ, Bai Y, Lee J-M, Godino J, Saint-Maurice PF, Carr L. Standardizing analytic methods and reporting in activity monitor validation studies. *Med Sci Sports Exerc*. 2019;51(8):1767-80.
47. Watson P, Petrie A. Method agreement analysis: a review of correct methodology. *Theriogenology*. 2010;73(9):1167-79.
48. Bonett DG. Sample size requirements for estimating intraclass correlations with desired precision. *Stat Med*. 2002;21(9):1331-5.
49. Pradhan N, Rajan S, Adler A, Redpath C, editors. Classification of the quality of wristband-based photoplethysmography signals. 2017 IEEE international symposium on medical measurements and applications (MeMeA); 2017 7–10 May; Rochester, MN, USA. Piscataway, NJ, USA: IEEE.
50. 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(3):540-7.
51. Terbizan DJ, Dolezal BA, Albano C. Validity of seven commercially available heart rate monitors. *Meas Phys Educ Exerc Sci*. 2002;6(4):243-7.

52. 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.* 2015;39(5):264-71.
53. Reiser M, Breidenassel A, Amft O, editors. Simulation framework for reflective PPG signal analysis depending on sensor placement and wavelength. 2022 IEEE-EMBS International Conference on Wearable and Implantable Body Sensor Networks (BSN); 2022: IEEE.
54. Wang Q, Sheng D, Zhou Z, Liu Z, editors. Numerical and experimental study of the influence of device pressure on PPG signal acquisition. *Optical Interactions with Tissue and Cells XXXI*; 2020: SPIE.
55. Santos-Lozano A, Marín PJ, Torres-Luque G, Ruiz JR, Lucía A, Garatachea N. Technical variability of the GT3X accelerometer. *Med Eng Phys.* 2012;34(6):787-90.
56. Rothney MP, Apker GA, Song Y, Chen KY. Comparing the performance of three generations of ActiGraph accelerometers. *J Appl Physiol (1985).* 2008;105(4):1091-7.
57. Pfeiffer KA, Clevenger KA, Kaplan A, Van Camp CA, Strath SJ, Montoye AH. Accessibility and use of novel methods for predicting physical activity and energy expenditure using accelerometry: a scoping review. *Physiol Meas.* 2022;43(9):09TR01.
58. Bruno E, Böttcher S, Viana PF, et al. Wearable devices for seizure detection: practical experiences and recommendations from the Wearables for Epilepsy And Research (WEAR) International Study Group. *Epilepsia.* 2021;62(10):2307-21.
59. Alfonso C, Garcia-Gonzalez MA, Parrado E, Gil-Rojas J, Ramos-Castro J, Capdevila L. Agreement between two photoplethysmography-based wearable devices for monitoring

heart rate during different physical activity situations: a new analysis methodology. *Sci Rep.* 2022;12(1):15448.

ACCEPTED

FIGURE LEGENDS

Fig 1. Example of high-quality spectrogram with protocol order identified.

Note: Protocol Order:(1) Supine Rest;(2) Rest (seated);(3) Stand;(4) Rest (seated);(5) Walk;(6) Rest (seated); (7) Jog/Run.

All activities last 5 minutes, aside from (1) which is 10 minutes.

Fig 2. The Bland-Altman plot of bias at the minute level with 95% limits of agreement (orange) for PATCH ENMO (mg) compared to the shaker table (criterion). Note: Grey solid line shows mean bias. Black dotted line represents the trend in bias ($y = 0.21x + -5.94$)

Fig 3. The Bland-Altman plot of bias at the minute level with 95% limits of agreement (orange) for PATCH heart rate compared to Actiheart. Note: Grey solid line shows mean bias. Black dotted line shows trend in bias ($y = -0.07x + 9.86$)

Fig. 4. Equivalence of heart rate (HR) and ENMO (mg) between PATCH and Actiheart.

Note: Grey bars show the equivalence zone for the Actiheart. The orange bars indicate the 90% confidence interval around the estimated difference between PATCH and Actiheart monitors. Measures are equivalent if the orange bars fall completely within the grey bars.

Fig 5. The Bland-Altman plot of bias with 95% limits of agreement (orange) for PATCH ENMO compared to Actiheart. Note: Black dotted line shows trend in bias trend in bias ($y = -0.34x + 4.77$)

SUPPLEMENTAL DIGITAL CONTENT

SDC 1: Supplementary Materials_v4.pdf

ACCEPTED

Figure 1

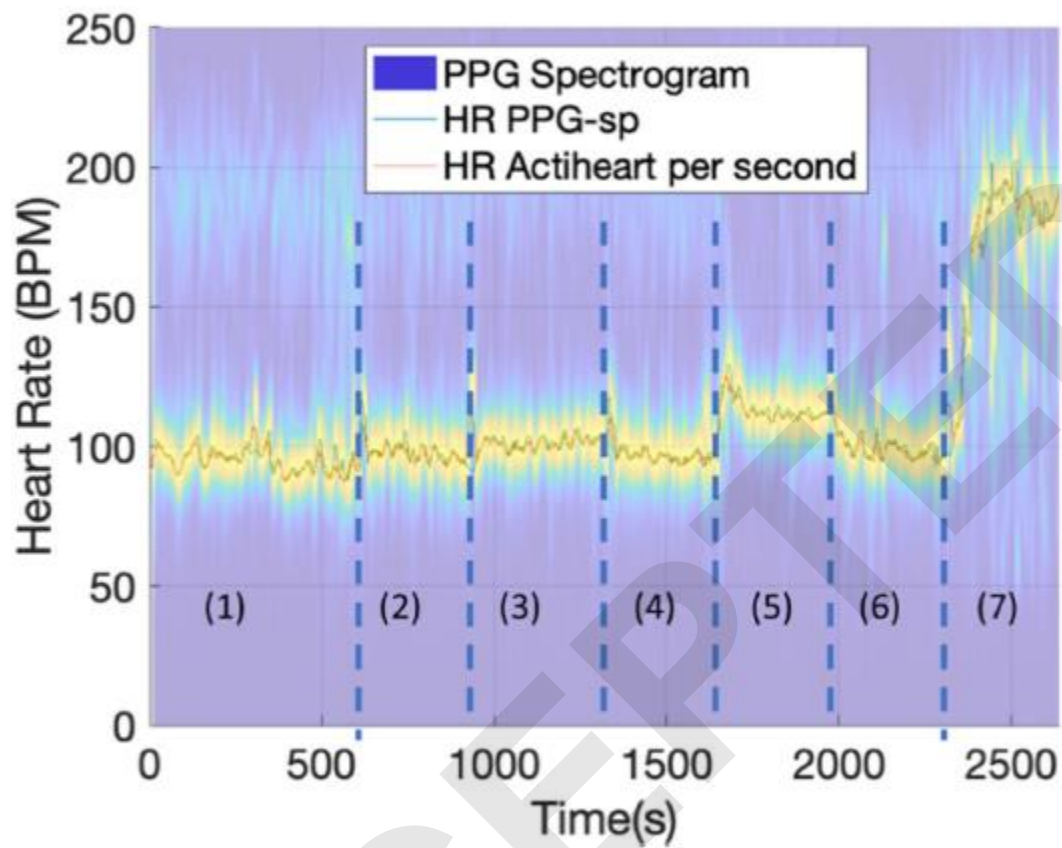


Figure 2

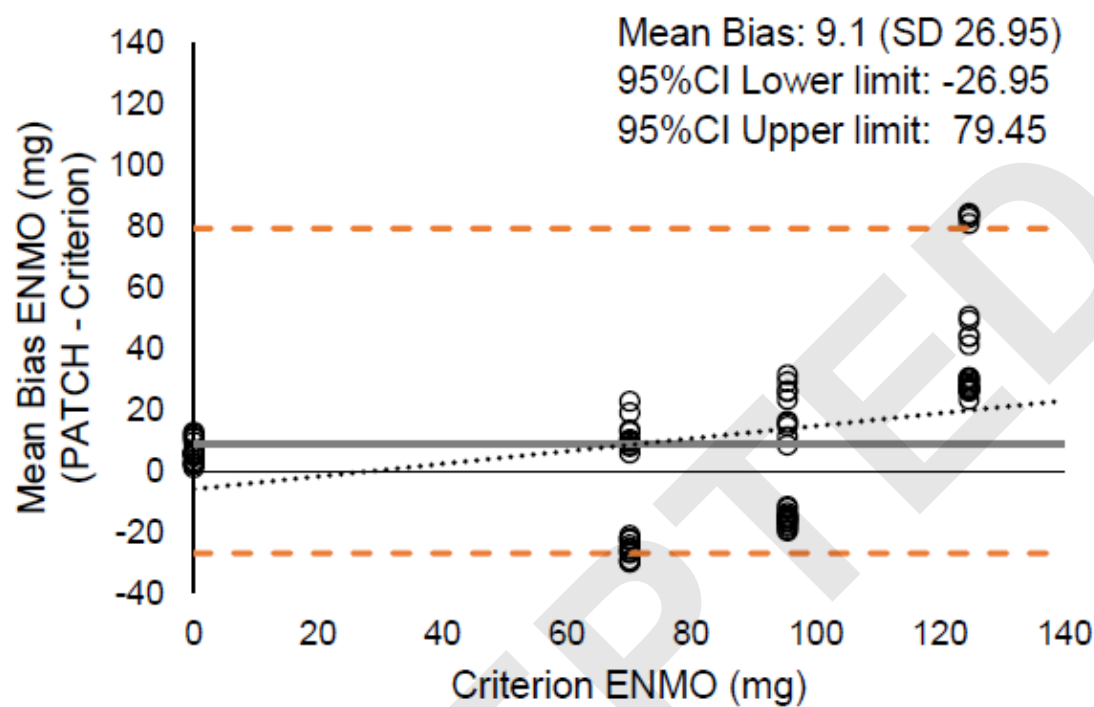


Figure 3

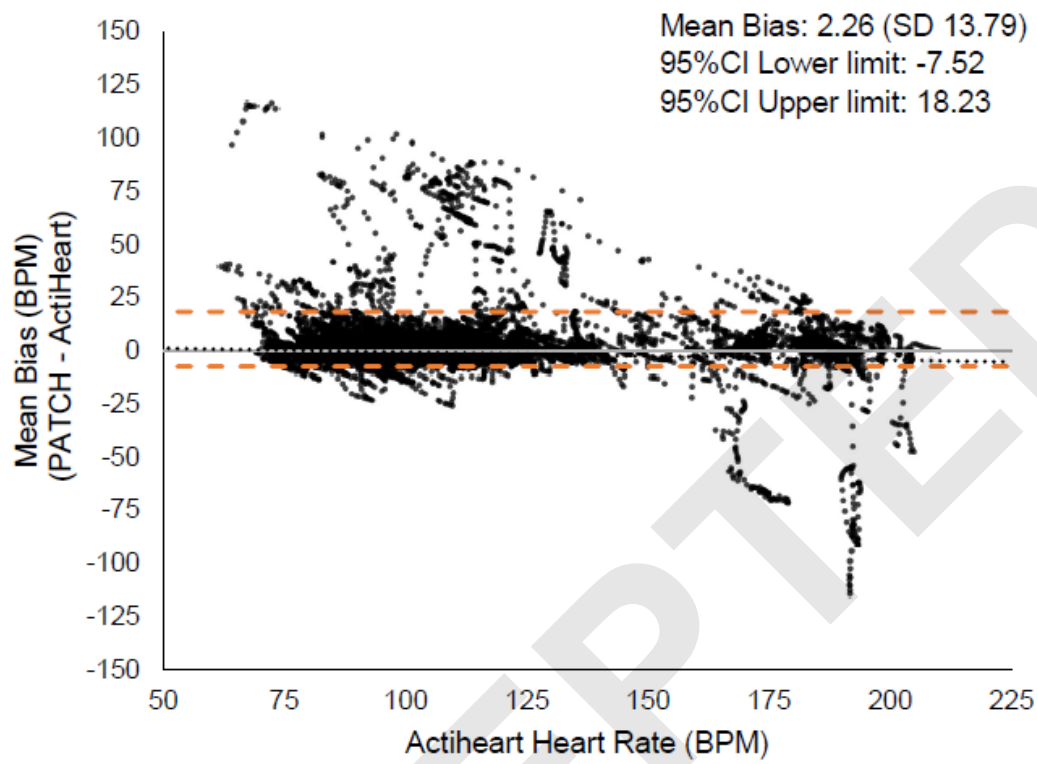


Figure 4

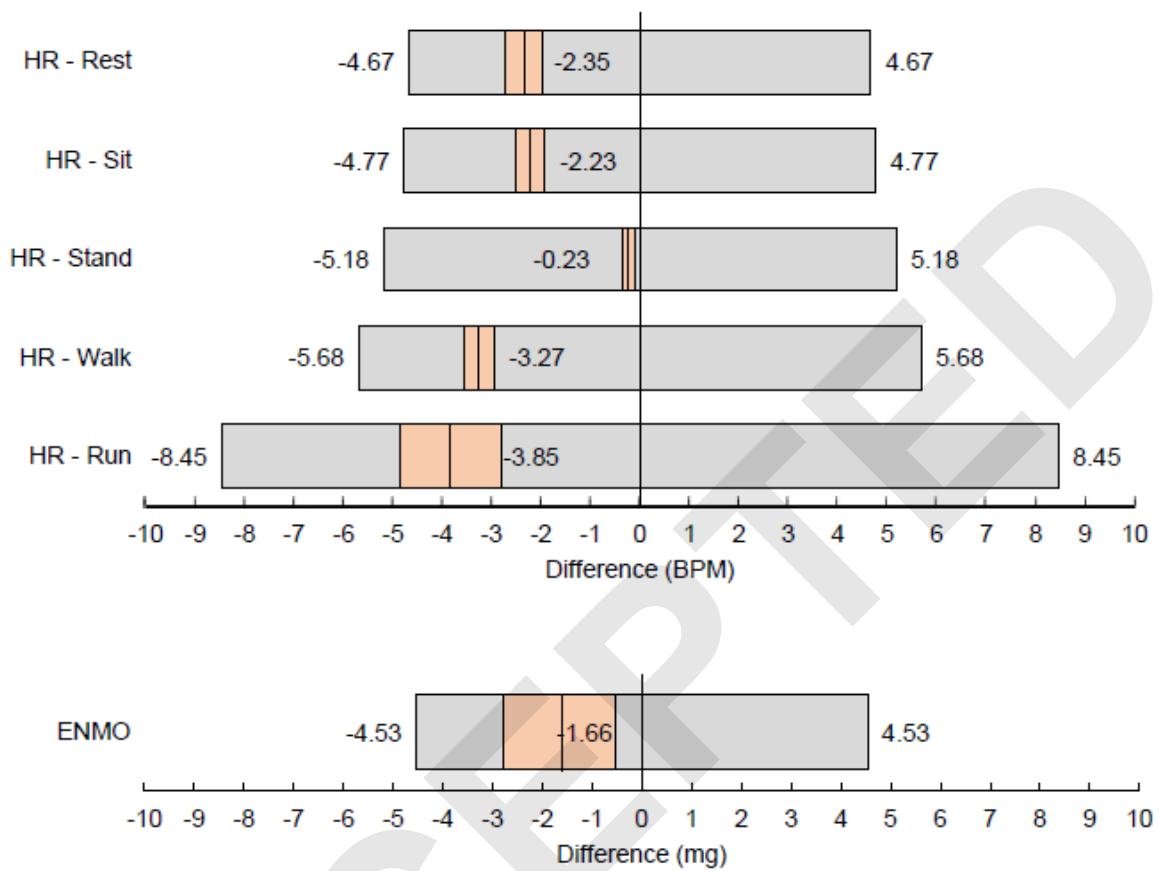


Figure 5

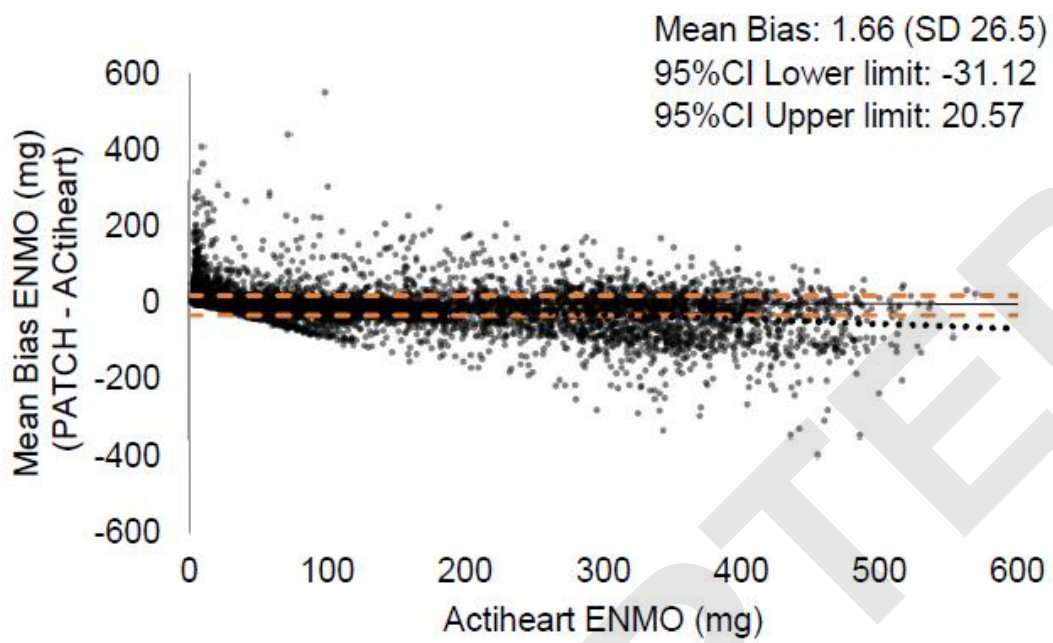


Table 1. Sample Demographics

	N/Mean	%/(SD)
Female	25	40%
Age (years)	6.3	(1.7)
Fitzpatrick Scale		
1	6	10%
2	20	32%
3	17	27%
4	13	21%
5	4	6%
6	3	5%
Weight (kg)	25.7	(9.35)
Height (cm)	119.8	(12.63)
Race		
Asian	2	3%
Black or African		
American	10	16%
White	47	75%
Other; More than 1 race	4	6%

Table 2. Heart Rate Error – Median, mean and interquartile lower limit (LL) and upper limit (UL) ranges for average bias, mean absolute error (MAE) & mean absolute percent error (MA%E)

Activity		Median	Mean	SD	Interquartile Range	
					(LL	UL)
Overall	Mean Bias	0.46	2.26	14.79	-2.09	3.77
	MAE	2.86	6.67	13.39	1.22	5.72
	MA%E	2.64	5.99	12.15	1.16	5.33
Supine Rest	Mean Bias	-0.03	2.34	12.79	-2.33	2.74
	MAE	2.50	5.75	11.67	1.04	5.44
	MA%E	2.59	6.18	12.18	1.15	5.76
Seated Rest	Mean Bias	0.72	2.19	9.46	-1.62	3.55
	MAE	2.62	4.76	8.46	1.16	4.97
	MA%E	2.76	5.06	8.46	1.21	5.22
Stand	Mean Bias	-0.09	0.24	3.94	-2.02	2.26
	MAE	2.13	2.89	2.69	0.98	3.84
	MA%E	2.09	2.88	2.84	0.93	3.78
Walk	Mean Bias	1.13	3.27	9.90	-1.76	5.28
	MAE	3.33	5.77	8.68	1.41	6.49
	MA%E	2.97	5.15	7.73	1.27	5.74
Jog/Run	Mean Bias	1.75	3.80	30.05	-3.18	8.53
	MAE	5.59	17.05	25.03	2.38	16.27
	MA%E	3.26	12.39	22.39	1.36	10.05

Table 3. ENMO (mg) - Median, mean and interquartile lower limit (LL) and upper limit (UL) ranges for mean bias & mean absolute error (MAE)

Activity		Median	Mean	SD	Interquartile Range	
					(LL -	UL)
Overall	Mean Bias	4.35	1.61	28.71	0.73	7.11
	MAE	5.87	12.24	23.51	3.48	9.73
Supine Rest	Mean Bias	5.51	6.18	10.75	3.32	7.76
	MAE	5.54	6.34	10.66	3.38	7.77
Seated Rest	Mean Bias	4.80	5.33	11.00	2.66	6.87
	MAE	4.95	6.16	10.55	2.90	7.05
Stand	Mean Bias	4.83	5.89	14.13	3.06	10.95
	MAE	4.91	6.47	13.87	4.95	6.57
Walk	Mean Bias	-7.51	-9.43	24.78	-16.02	-1.47
	MAE	9.00	14.87	21.95	0.94	17.79
Jog/Run	Mean Bias	-9.82	-18.25	56.78	-41.40	6.62
	MAE	23.15	39.53	44.65	8.34	54.65

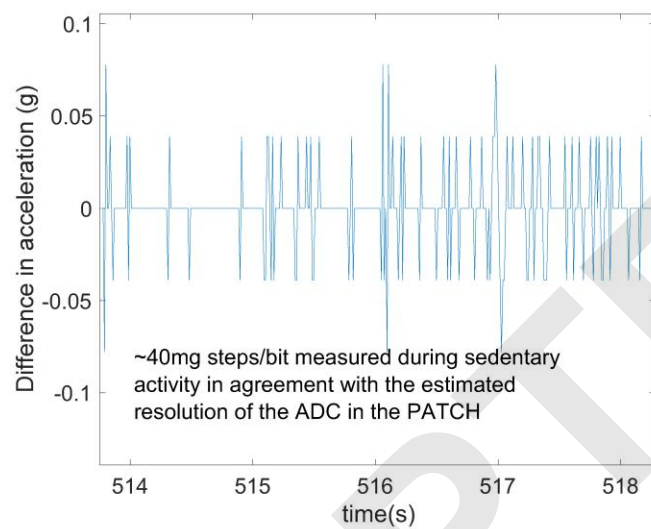
Supplemental Materials, Tables & Figures

Signal Quality.

Self-Consistency (also known as heart rate frequency difference[1]) is defined as the difference between the fundamental frequency and heart rate computed from the time domain peak calculation. This feature measures the agreement between the fundamental frequencies detected from the frequency spectrum and from the time-domain signal. It is assumed that the frequencies would be in agreement in a clean PPG segment. In a noise-corrupted segment, however, there could be large differences in the values. We compute the self-consistency metric as follows: the self-consistency between time domain and frequency domain is defined as the fraction of points that agreed to within 10 BPM i.e. $1.94 \times$ the 5BPM limits of agreement threshold chosen. Previous studies reveal a self-consistency below 30 BPM indicates potential poor-quality signal.[2]

Standard Deviation of line width (STD-width) is the second signal quality metric used. Supplementary Figure A show examples of spectrograms with bad agreement with the telemetry reference values, respectively. In the “good” signal (Figure 1), the heart rate signal is sharp and well defined in frequency, and the width of the line does not change much, leading to a small standard deviation. In the “poor” signal (Supplementary Figure A), the emergence of the $1/f$ (or, wideband) streaks from loss of contact with the skin as described above leads to very wide streaks when contact with skin is lost. When the abrupt change stabilizes, the line width (perhaps due to noise) becomes unpredictable until skin contact is re-established. This cyclical process leads to wide variations in line width due to poor mounting giving a large standard deviation. Previous studies reveal a STD-width above 10 BPM indicates potential poor-quality signal. [2]

Supplementary Figure A. Typical variation of measured vector magnitude of acceleration during sedentary activity.



Supplementary Table A. Components in the PATCH device (version 2.3.7)

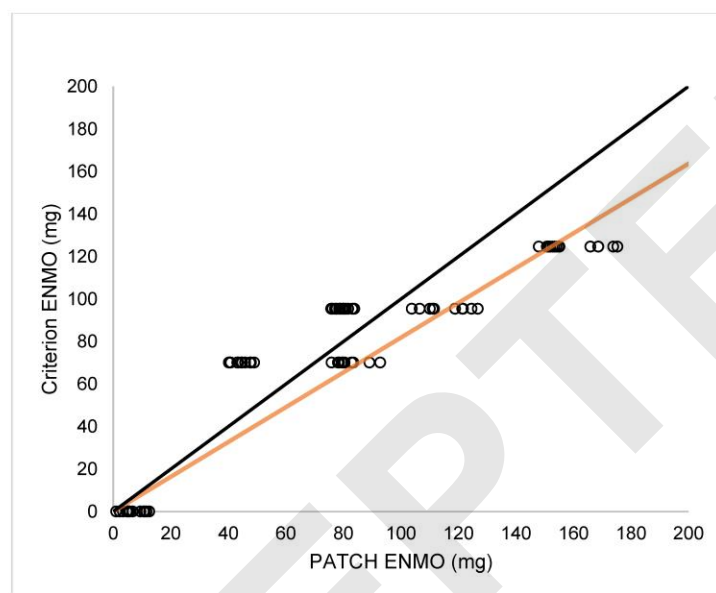
Component	Manufacturer	Model	Purchase Link
Board	Arduino	Micro	https://store-usa.arduino.cc/products/arduino-micro?selectedStore=us
PPG Sensor	PulseSensor	PulseSensor	https://pulsesensor.com/products/pulse-sensor-amped
SD	Adafruit	MicroSD card breakout board	https://www.adafruit.com/product/254
Accelerometer	Adafruit	ADXL326	https://www.adafruit.com/product/1018

Supplementary Table B. Protocol Tasks

Order	Activity	Target Physical Activity Type	Target METs	Minutes Spent in each activity
1	Supine Resting	Sedentary	1-1.4	10
2	Rest Transition (seated)	Sedentary	1-1.4	5
3	Stand	Stationary	1.5-2.0	5
4	Rest Transition (seated)	Sedentary	1-1.4	5
5	Walk	Light	2.0-3.6	5
6	Rest Transition (seated)	Sedentary	1-1.4	5
7	Jog/Run	Moderate to vigorous	5.5-10.6	5

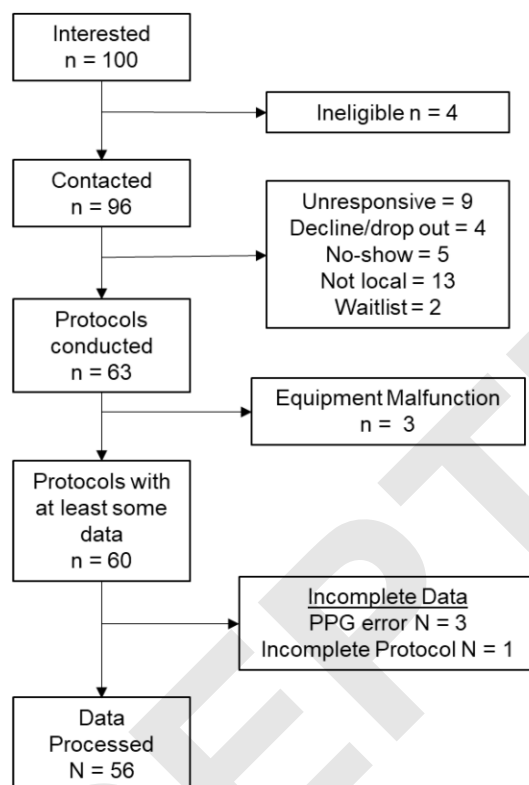
METS - Metabolic equivalent of task

Supplementary Figure B. Lin's Concordance Coefficient (CCC) for accelerometer data from Shaker table and PATCH ENMO values



Note: Orange line represents line of best fit for sample data
The black line represents perfect fit ($r = 1.0$)

Supplementary Figure C. Flow diagram of participants.



Supplementary Table C. Individual person-level heart rate estimates for mean bias mean absolute error (MAE), mean absolute percent error (MA%E), and signal quality indices of self-consistency and standard deviation (STD) line width

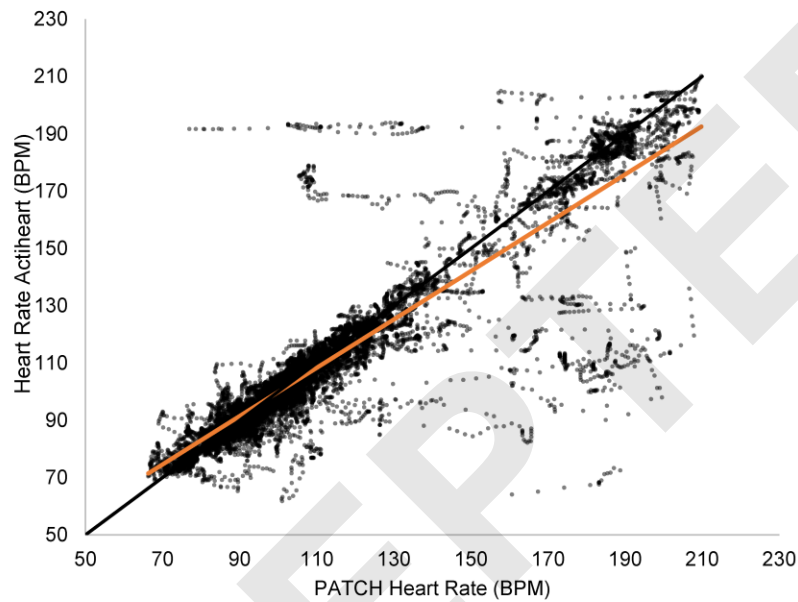
ID	Mean Bias	MAE	MA%E	Self-Consistency ^a	STD Line Width ^b
890	22.71	23.58	22.79	0.60	8.84
1106	2.70	4.20	5.05	0.79	6.30
1222	-0.37	2.05	1.64	0.77	4.66
1799	1.28	2.15	2.36	0.97	2.75
2007	0.75	2.27	1.91	0.85	5.58
2030	-1.34	2.86	2.52	0.90	3.92
2076	10.51	13.54	13.35	0.62	5.24
2077	11.51	14.25	13.44	0.50	9.81
2078	-1.04	5.55	5.83	0.71	7.31
2214	0.90	3.58	2.49	0.87	3.10
2215	-1.13	2.30	3.05	0.93	5.03
2216	1.18	5.15	4.16	0.82	5.56
2222	3.82	5.07	4.14	0.86	4.56
2469	4.49	6.01	4.33	0.84	4.95
2470	0.42	4.37	3.40	0.95	3.81
2471	-1.22	7.43	5.67	0.80	5.44
2473	-0.02	3.61	3.43	0.75	5.62
2474	-2.21	4.21	3.23	0.87	4.01
2476	0.17	1.90	1.59	0.93	3.39
2477	-0.34	10.18	10.73	0.80	5.57
2478	0.13	12.04	12.59	0.87	4.54
2479	1.24	6.84	6.81	0.87	4.61
2480	8.04	9.39	12.88	0.79	7.52
2491	9.70	12.68	10.20	0.77	4.45
2492	4.32	7.25	6.46	0.80	5.76
2517	2.76	3.64	3.37	0.77	5.67
2518	-4.85	8.73	5.89	0.81	5.24
2525	2.04	3.90	3.94	0.78	5.48
2541	1.28	2.86	2.23	0.66	5.42
2557	23.73	24.56	25.13	0.44	5.20
2568	-5.74	7.20	5.00	0.75	7.07
2569	13.12	15.76	13.63	0.50	6.50
2570	0.76	3.09	2.62	0.78	7.44
2581	3.06	5.67	4.81	0.80	5.15
2589	1.71	3.14	2.78	0.52	8.15
2596	8.02	10.80	11.23	0.81	7.13
2597	14.07	15.47	13.55	0.89	6.31
2601	0.92	1.65	1.60	0.96	1.86
2604	-0.02	4.19	3.63	0.84	5.22
2614	1.35	3.03	2.66	0.84	4.00
2616	0.84	2.40	1.99	0.77	4.42
2634	2.15	3.22	3.05	0.93	3.23
2635	5.73	9.92	9.64	0.68	4.93
2656	9.83	10.20	9.86	0.87	4.30
2657	0.69	2.51	2.32	0.79	4.57
2664	0.12	2.72	2.87	0.81	4.46
2665	-0.69	2.05	2.00	0.72	9.72

2670	-1.15	2.23	1.87	0.89	4.44
2679	-0.48	1.05	0.84	0.83	5.30
2693	-0.28	2.76	2.90	0.87	6.02
2694	0.31	7.36	6.05	0.83	5.50
2695	-12.66	14.90	9.26	0.87	4.51
2707	1.02	3.09	3.12	0.81	5.92
2708	3.41	5.39	6.00	0.63	7.56
2714	-14.01	18.08	11.63	0.68	6.53
2729	4.34	5.82	7.05	0.34	7.54

^a Self-consistency - Values closer to 1 indicate better signal quality

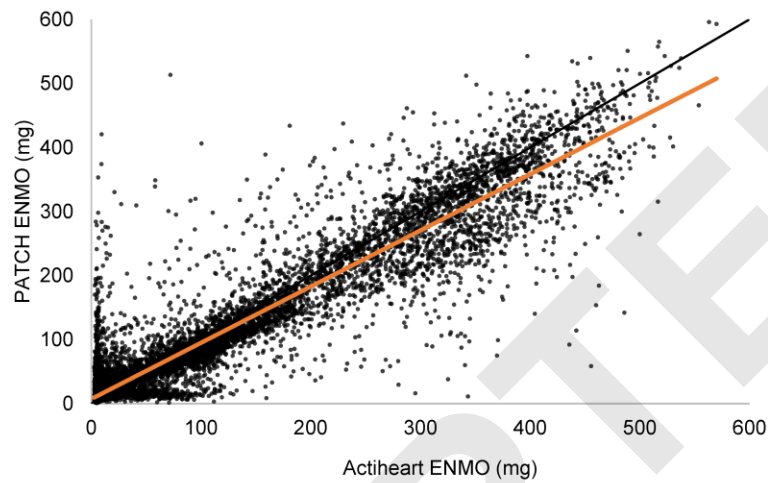
^b STD Line Width – Values closer to 0 indicate better signal quality

Supplementary Figure D. Lin's correlation concordance coefficient for heart rate



Note: Orange line shows line of best fit for sample data
The black line represents perfect fit ($r = 1.0$)
BPM – Beats per minute

Supplementary Figure E. Lin's correlation concordance coefficient for ENMO between PATCH and Actiheart accelerometers



Note: Orange line shows line of best fit for sample data
Black line represents perfect fit

References

1. Yan Y-s, Poon CC, Zhang Y-t. Reduction of motion artifact in pulse oximetry by smoothed pseudo Wigner-Ville distribution. *J Neuroeng Rehabil.* **2005**, 2,3.
2. McLean MK, Weaver RG, Lane A, Smith MT, Parker H, Stone B, McAninch J, Matolak DW, Burkart S, Chandrashekhar M. A Sliding Scale Signal Quality Metric of Photoplethysmography Applicable to Measuring Heart Rate across Clinical Contexts with Chest Mounting as a Case Study. *Sensors.* **2023**, 23 (7),3429.