Evaluating utility and compliance in patient-based eHealth study using continuous-time heart rate and activity trackers

William Speier¹, Eldin Dzubur², Mary Zide¹, Chrisandra Shufelt³, Sandy Joun³, Jennifer E. Van Eyk³,⁴, C. Noel Bairey Merz³, Mayra Lopez², Brennan Spiegel², Corey Arnold¹,⁵,⁶*

Corresponding author:
Corey Arnold
924 Westwood Blvd Ste 420
Los Angeles, CA 90024
(310) 794-3538 (p)
(310) 794-3546 (f)
CWArnold@mednet.ucla.edu

Affiliations:
¹Medical Imaging and Informatics group, University of California, Los Angeles, USA
²Center for Outcomes Research and Education, Cedars-Sinai Medical Center, Los Angeles, USA
³Barbra Streisand Women’s Heart Center, Cedars-Sinai Smidt Heart Institute, Los Angeles, USA
⁴Department of Biomedical Sciences, Cedars-Sinai Medical Institute, Los Angeles, USA
⁵Department of Radiological Sciences, University of California, Los Angeles, USA
⁶Department of Pathology & Laboratory Medicine, University of California, Los Angeles, USA

Keywords: mHealth, health apps, self-monitoring, mobile health, informatics

Word count: 1996
ABSTRACT

Telemedicine has been used to remotely diagnose and treat patients, yet previously applied telemonitoring approaches have been fraught with adherence issues. The primary goal of this study was to evaluate the adherence rates using a consumer-grade continuous-time heart rate and activity tracker in a mid-risk cardiovascular patient population. As a secondary analysis, we show the ability to utilize the information provided by this device to identify information about a patient’s state by correlating tracker information with patient reported outcome survey scores. We showed that using continuous-time activity trackers with heart rate monitors can be effective in a telemonitoring application as patients had a high level of adherence (90.0% median usage) and low attrition (0.09% decrease per day) over a 90 day period. Furthermore, data collected correlated significantly with clinically relevant patient surveys ($r^2=0.15$ for PROMIS global health scores, $p<0.00001$), and therefore might provide an effective signal for identifying patients in need of intervention.

INTRODUCTION

There has been significant prior work in research and clinical settings in the use of telemedicine to remotely diagnose and treat patients. However, previously applied telemonitoring approaches have been fraught with adherence issues and often exhibit non-conclusive results.[1–11] Additionally, studies have indicated that device fatigue limits adherence,[12–14] a phenomenon known as the law of attrition.[15] Little work has been done to demonstrate how readily available commercial devices may limit intervention burden by automating data collection, such as passive accelerometry. However, previous studies have shown that activity trackers are capable of accurately documenting health indicators such as physical activity, are a popular low-cost option with older patients, and often have higher adherence rates than other devices.[16–19]

Activity trackers in combination with smartphones are perceived to be easy-to-use, accessible means for providing feedback and support directly to patients.[20] This feedback loop has been shown to positively impact health interventions with the goal of lifestyle changes.[21–23] Although demonstrating potential as intervention methods, these approaches have largely been used in studies with small samples or in healthy subjects, which may not accurately represent true adherence in a telemedicine application.[1,4,24,25] Furthermore, these studies have largely been based on data aggregated over the course of days or weeks, which limits the precision with which patients can be monitored and restricts the ability to detect problems or intervene when beneficial to patients.[1,26,27]
The primary goal of this study was to evaluate the adherence rates using a consumer-grade continuous-time heart rate (HR) and activity tracker over 90 days in a group of patients with ischemic heart disease (IHD), a sample representative of a chronic disease patient population. Patients with stable IHD can develop precursor indicators that are not clinically detected, yet may progress to major adverse cardiac events (MACE) if not recognized, a potentially catastrophic and expensive outcome. There is a pressing need to validate and scale cost-effective techniques to monitor this large population between healthcare visits, and adherence levels relate directly to the feasibility of clinical use. We compare the ability to detect activity and adherence using this device to methods used with standard accelerometers. As a secondary analysis, we show the ability to utilize the information provided by this device to identify information about a patient’s state by correlating tracker information with survey scores.

METHODS

Data Collection

A group of 200 patients with IHD were recruited from Cedars-Sinai Medical Center as part of a larger study seeking to predict MACEs using biometrics, biomarkers, wearable sensors, and patient-reported surveys. Nine subjects withdrew from the study and five were lost to follow-up. Wearable data was available for the remaining 186 subjects (93%) for inclusion in this analysis. The study was approved by the Cedars-Sinai institutional Review Board (IRB). Subjects were given Fitbit Charge 2 (Fitbit Inc., San Francisco, CA, USA) HR trackers at enrollment and followed for 90 days.

Patient State Surveys

Subjects were each administered the Patient-Reported Outcomes Measurement Information System (PROMIS), Seattle Angina (SAQ), and Kansas City Cardiomyopathy (KCCQ) questionnaires at the conclusion of the 90-day study.[28–30] PROMIS is a group of validated measures of global and domain-specific physical, mental, and social health.[28] PROMIS includes domain-specific measures for depression, emotional distress, fatigue, physical function, sleep disturbance, and social isolation, as well as a global short form version (PROMIS 10), which queries physical, mental, and social health. The SAQ and KCCQ are short form surveys with Likert-type items specific to cardiology patients, focusing on quality of life/functionality, and symptoms.[29,30]
Adherence

Adherence to ambulatory measurement of physical activity was measured using four methods: 1) HR hours (HR-hour); 2) HR minutes (HR-minute); 3) standard Actigraph non-wear without sleep imputation (NHANES-wake); and 4) standard Actigraph non-wear with fixed sleep (NHANES-sleep). The Fitbit Charge 2 device provides a continuous stream of HR data using photoplethysmography (PPG), which allows us to disentangle moments of sedentary time and non-wear. Further, PPG can be used, in tandem with accelerometer data, to accurately estimate sleep time. For HR-minute, we examined data at 60-second epochs, where bouts with missing HR data were considered non-wear. HR-hour was calculated by aggregating data to the hour and defining non-wear as hours with no available HR data.

In order to provide a standard comparison to prior work in accelerometer studies, we used two protocols based on criteria used in the Actigraph module of the National Health and Nutrition Examination Survey (NHANES). Non-wear was estimated using continuous bouts of zero activity counts lasting longer than 60 minutes, allowing for up to two minutes of activity.[31] To calculate adherence for NHANES-wake, we made no assumptions about sleep time and examine all available data in a single stream. For NHANES-sleep, we assumed 16 waking hours per day and exactly eight hours of sleep.[32] Because the NHANES study combines non-use and sleep, we subtracted the eight hours of sleep directly from the non-use period.

Summary Statistics

Summary statistics for telemonitoring devices were reported as averages per day across valid days. Valid days were defined as those having at least 10 hours of wear time per day, irrespective of algorithm.[31] Consistency within patients’ recording was evaluated by calculating intraclass correlation (ICC) using a two-way random effects model. Correlation metrics between the average number of steps per day for each patient and the questionnaire scores were calculated. The statistical significance of the correlations was tested using a t-test with Bonferroni correction for multiple comparisons.

RESULTS

Adherence

Time spent using the device can be separated into hours spent sedentary, engaged in moderate-vigorous physical activity (MVPA), light physical activity (LPA), or asleep. Table 1 presents medians and interquartile ranges of active living components, including non-wear time. The HR based criteria generally had lower
non-wear times and higher sedentary times than the NHANES criteria. Figure 1 illustrates the proportion of usage across all subjects based on each of the criteria. The median usage percentages were 90.0%, 83.7%, 43.8%, and 77.1%, when calculated by the HR-hour HR-minute, NHANES-wake, and NHANES-sleep criteria, respectively.

Table 1: Median and IQR for the hours per day that subjects were sedentary, engaged in light (LPA) or moderate-vigorous (MVPA) physical activity, asleep, or not wearing the device. These values were calculated with non-wear time determined four ways: by heart rate at the hour (HR-hour) or minute (HR-minute) level, and by activity level without considering sleep (NHANES-wake) and with an assumed sleep period (NHANES-sleep).

<table>
<thead>
<tr>
<th></th>
<th>Sedentary</th>
<th>LPA</th>
<th>MVPA</th>
<th>Sleep</th>
<th>Non-wear</th>
</tr>
</thead>
<tbody>
<tr>
<td>HR-hour</td>
<td>11.34</td>
<td>2.50</td>
<td>0.24</td>
<td>5.66</td>
<td>2.43</td>
</tr>
<tr>
<td></td>
<td>(9.92 – 12.67)</td>
<td>(1.50 – 3.59)</td>
<td>(0.07 - 0.51)</td>
<td>(3.60 – 7.22)</td>
<td>(9.02 – 7.31)</td>
</tr>
<tr>
<td>HR-minute</td>
<td>10.02</td>
<td>2.50</td>
<td>0.24</td>
<td>5.66</td>
<td>3.92</td>
</tr>
<tr>
<td></td>
<td>(8.47 – 11.40)</td>
<td>(1.50 – 3.59)</td>
<td>(0.07 - 0.51)</td>
<td>(3.60 – 7.22)</td>
<td>(2.30 – 8.43)</td>
</tr>
<tr>
<td>NHANES-wake</td>
<td>7.88</td>
<td>2.50</td>
<td>0.24</td>
<td>NA</td>
<td>13.48</td>
</tr>
<tr>
<td></td>
<td>(5.19 – 9.37)</td>
<td>(1.50 – 3.59)</td>
<td>(0.07 - 0.51)</td>
<td></td>
<td>(10.70 – 16.94)</td>
</tr>
<tr>
<td>NHANES-sleep</td>
<td>7.88</td>
<td>2.50</td>
<td>0.24</td>
<td>8</td>
<td>5.48</td>
</tr>
<tr>
<td></td>
<td>(5.19 – 9.37)</td>
<td>(1.50 – 3.59)</td>
<td>(0.07 - 0.51)</td>
<td></td>
<td>(2.70 – 8.94)</td>
</tr>
</tbody>
</table>

Summary Statistics

HR-hour was selected for analysis of summary statistics and correlations. As shown in Figure 2, there was a 0.09% decrease per day in usage specified by HR-hour (t=4.50, p=0.00001). Summary statistics, including means, standard deviations, and ICCs are presented in Table 2 and Figure 3. A patient’s average number of steps per day had significant positive correlation with their KCCQ ($r^2$=0.09, p=0.0001) and SAQ ($r^2$=0.08, p=0.0008) overall scores. The average number of steps also positively correlated with the PROMIS global physical health ($r^2$=0.15, p<0.00001) and physical function short form ($r^2$=0.18, p<0.00001) scores, and negatively correlated with the fatigue short form score ($r^2$=0.13, p=0.0001). There was no significant correlation between the average number of steps and the PROMIS emotional distress short form ($r^2$=0.01, p=0.12), social isolation short form ($r^2$=0.02, p=0.07), and sleep disturbance short form ($r^2$=0.03, p=0.03) scores. There were slight correlations with other variables with the PROMIS global physical health score, such as resting HR ($r^2$=0.04, p=0.02), but none were statistically significant.
Table 2: Mean, standard deviation, and intraclass correlation coefficient (ICC) for active and resting heart rates, number of steps per day, and average number of hours spent sedentary, engaged in light (LPA) or moderate-vigorous (MVPA) physical activity, asleep, or not wearing the device.

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>SD</th>
<th>ICC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Steps/day</td>
<td>4882</td>
<td>3070</td>
<td>0.47</td>
</tr>
<tr>
<td>Active HR</td>
<td>81.8</td>
<td>9.5</td>
<td>0.41</td>
</tr>
<tr>
<td>Resting HR</td>
<td>66.2</td>
<td>7.7</td>
<td>0.35</td>
</tr>
<tr>
<td>Sleep hours/day</td>
<td>5.2</td>
<td>2.4</td>
<td>0.38</td>
</tr>
<tr>
<td>Sedentary hours/day</td>
<td>9.9</td>
<td>2.7</td>
<td>0.21</td>
</tr>
<tr>
<td>LPA hours/day</td>
<td>2.6</td>
<td>1.4</td>
<td>0.46</td>
</tr>
<tr>
<td>MVPA hours/day</td>
<td>0.4</td>
<td>0.4</td>
<td>0.39</td>
</tr>
<tr>
<td>Non-wear hours/day</td>
<td>6.0</td>
<td>5.1</td>
<td>0.37</td>
</tr>
</tbody>
</table>

DISCUSSION

In this study, we observed adherence rates up to 90.0% (Figure 1), substantially higher than those previously reported in telemonitoring studies using similar study periods. Shaw et al. reported patients with chronic illness using a variety of telemonitoring devices were only adherent 16% of days in their four-week study.[13] In the Shaw study, patients reported feeling overwhelmed with having multiple devices and eventually only used those that interested them, which included a Fitbit. Our study used a single device with high utility, which simplifies the task of remote monitoring for both researchers and patients. The ability of the device to simultaneously record multiple variables such as HR and accelerometer data also allowed us to more accurately determine the patient’s state, whether they are active, sedentary, asleep, or not currently using the device. Over the course of the study, average adherence dropped from an initial 87.7% to 72.0% on the final day (90). The reduction in adherence is consistent with the law of attrition in eHealth studies,[15] as patients will generally have high interest at the point of recruitment, but may gradually lose interest in, start to forget about, or become burdened by the study. Because the Fitbit provides data access in real time, gaps of adherence can be detected quickly and reminders could be sent as a result, possibly improving adherence.
The average number of steps recorded per day by subjects in this study correlated significantly with several attributes of the patients’ self-reported states at the conclusion of the study, including their overall health, physical function, fatigue, and KCCQ and SAQ scores. These findings are consistent with previous studies that have found relationships between physical activity and health,[33,34] as well as physical function.[35,36] Similarly, other studies have shown telemonitoring participation correlated with improvements in clinical outcomes in specific groups, however not all were statistically significant.[37–40] Given these results, telemonitoring data could possibly be used as a surrogate for these questionnaires, allowing clinicians to automatically evaluate a patient’s state and possibly intervene without needing to wait until the patient’s next appointment or patient self-report. However, some survey scores such as emotional distress did not correlate strongly with raw telemonitoring data. Also, many of the telemonitoring variables did not correlate significantly with survey scores in the univariate analyses performed here. These survey scores may still be predictable from telemonitoring data using multivariate analyses and machine learning methods, which can exploit complicated interactions between variables and outcomes.

The sleep estimate provided by the Fitbit may underestimate the amount of sleep for some subjects, as it found that 30.1% of the subjects in this study averaged under four hours of sleep per night. Patients may remove their Fitbits overnight, resulting in non-wear time during sleep. Additionally, prior literature has found that Fitbit devices tend to underestimate sleep in comparison to gold-standard trackers (e.g., ActiWatch), although less is known about Fitbit devices that use PPG.[41] Taking into account a patient’s normal schedule, statistical models might be able to infer the activity during these non-wear periods. We plan to investigate whether we might be able to leverage the continuous nature of the Fitbit data to overcome these censored periods when the subject is not using the device.

The ability to use activity tracker information to make clinical predictions is reliant upon the accuracy of the data provided by the tracker. Benedetto et al., demonstrated that individual HR recordings from the device used in this study could underestimate a subject’s actual HR by up to 30 bpm.[42] However, most clinical decisions are not made on the time resolution of a single recording, but rather an average over a period of time, which Benedetto showed had a more modest 5.9 bpm bias. Here, the correlation between consumer grade activity tracker data and clinically used patient surveys demonstrates that these devices provide information about a patient’s state, despite the imperfections in the data. While the correlated with clinically administered outcome surveys was statistically significant, that does not guarantee a significant effect in clinical practice. Patient reported outcomes are increasingly being used in health
care,[43] which indicates that this correlation could have a clinical impact. However, the best way to evaluate clinical utility would be to conduct a prospective trial where patient treatment is based in part on the telehealth data recorded from an activity tracker. A study could use this data to provide interventions for patients whose data correlates with low or decreasing outcome scores.

CONCLUSION

This study demonstrates that consumer-grade continuous-time activity trackers with HR monitors may be an effective tool for telemonitoring applications as patients have demonstrated a high level of adherence and relatively low attrition over 90 days. These devices can record useful patient statistics including activity level, resting and active HR, and sleep time. These data correlate with clinically used patient surveys, and therefore might be an effective way of identifying patients who require intervention. Future studies should investigate the utility of this real-time tracking as a basis for patient health surveillance and as a means for using feedback to overcome the attrition seen in eHealth studies.

CONTRIBUTIONS

WS, ED, and CWA contributed to the conception of this analysis and the interpretation of the data, and drafted and revised the manuscript. CS, SJ, JVE, NBM, ML, and BS contributed to the study design and acquisition of the data, and provided critical revisions to the manuscript. MZ contributed to the analysis, and drafted and revised the manuscript.

FUNDING

This work was supported in part by the California Initiative to Advance Precision Medicine (CIAPM) (BS, NBM, and JVE); the National Institutes of Heart, Lung, and Blood Institute (NIH/NHLBI R56HL135425, CWA; K23HL127262, CS); the National Center for Research Resources (NIH/NCRR UL1RR033176); the National Center for Advancing Translational Sciences (NCATS) and UCLA Clinical Translational Science Institute (CTSI) (NIH/NCATS UL1TR000124); the Advanced Clinical Biosystems Research institute (JVE); the Erika Glazer Endowed Chair in Women’s Heart Health (NBM and JVE); and the Barbra Streisand Women’s Cardiovascular Research and Education Program.

COMPETING INTERESTS

The authors declare no conflict of interest.
REFERENCES


17 Persons AAR. Building a Better Tracker: Older Consumers Weigh in on Activity and Sleep Monitoring Devices. 2015.


33 Tudor-Locke C. Steps to better cardiovascular health: how many steps does it take to achieve good health and how confident are we in this number? *Curr Cardiovasc Risk Rep* 2010;**4**:271–6.


41 Mantua J, Gravel N, Spencer R. Reliability of sleep measures from four personal health monitoring devices compared to research-based actigraphy and polysomnography. *Sensors*
2016;16:646.


FIGURE CAPTIONS

Figure 1: Histograms of the adherence percentage using the tracker for the four methods of calculation. Using activity to determine use of an activity tracker gives consistently lower adherence values because it can treat use during sleep and extended sedentary time as non-use. Using heart rate gives values that are more accurate because it can track use during sedentary times as well as sleep.

Figure 2: Average usage rate over the course of the study (Blue). Linear regression was used to find a slight downward trend of 0.09% per day over the course of the study.

Figure 3: Histograms of active and resting heart rates, number of steps per day, and average number of hours spent sedentary, engaged in MVPA or LPA, or asleep.