Associations Between Heart Rate Variability Measured With a Wrist-Worn Sensor and Older Adults’ Physical Function: Observational Study

Sarah Anne Graham¹,², PhD; Dilip V Jeste¹,²,³, MD; Ellen E Lee¹,², MD; Tsung-Chin Wu⁴, BS; Xin Tu¹,², PhD; Ho-Cheol Kim⁵, PhD; Colin A Depp¹,²,⁶, PhD

¹Sam and Rose Stein Institute for Research on Aging, University of California San Diego, La Jolla, CA, United States
²Department of Psychiatry, University of California San Diego, La Jolla, CA, United States
³Department of Neurosciences, University of California San Diego, La Jolla, CA, United States
⁴Department of Mathematics, University of California San Diego, La Jolla, CA, United States
⁵Scalable Knowledge Intelligence, IBM Research – Almaden, San Jose, CA, United States
⁶VA San Diego Healthcare System, San Diego, CA, United States

Corresponding Author:
Sarah Anne Graham, PhD
Sam and Rose Stein Institute for Research on Aging
University of California San Diego
9500 Gilman Drive, #0664
La Jolla, CA, 92093-0664
United States
Phone: 1 858 534 5433
Email: sagraham@ucsd.edu

Abstract

Background: Heart rate variability (HRV), or variation in beat-to-beat intervals of the heart, is a quantitative measure of autonomic regulation of the cardiovascular system. Low HRV derived from electrocardiogram (ECG) recordings is reported to be related to physical frailty in older adults. Recent advances in wearable technology offer opportunities to more easily integrate monitoring of HRV into regular clinical geriatric health assessments. However, signals obtained from ECG versus wearable photoplethysmography (PPG) devices are different, and a critical first step preceding their widespread use is to determine whether HRV metrics derived from PPG devices also relate to older adults’ physical function.

Objective: This study aimed to investigate associations between HRV measured with a wrist-worn PPG device, the Empatica E4 sensor, and validated clinical measures of both objective and self-reported physical function in a cohort of older adults living independently within a continuing care senior housing community. Our primary hypothesis was that lower HRV would be associated with lower physical function. In addition, we expected that HRV would explain a significant proportion of variance in measures of physical health status.

Methods: We evaluated 77 participants from an ongoing study of older adults aged between 65 and 95 years. The assessments encompassed a thorough examination of domains typically included in a geriatric health evaluation. We collected HRV data with the Empatica E4 device and examined bivariate correlations between HRV quantified with the triangular index (HRV TI) and 3 widely used and validated measures of physical functioning—the Short Physical Performance Battery (SPPB), Timed Up and Go (TUG), and Medical Outcomes Study Short Form 36 (SF-36) physical composite scores. We further investigated the additional predictive power of HRV TI on physical health status, as characterized by SF-36 physical composite scores and Cumulative Illness Rating Scale for Geriatrics (CIRS-G) scores, using generalized estimating equation regression analyses with backward elimination.

Results: We observed significant associations of HRV TI with SPPB (n=52; Spearman ρ=0.41; P=.003), TUG (n=51; ρ=−0.40; P=.004), SF-36 physical composite scores (n=49; ρ=0.37; P=.009), and CIRS-G scores (n=52, ρ=−0.43; P=.001). In addition, the HRV TI explained a significant proportion of variance in SF-36 physical composite scores (R²=0.28 vs 0.11 without HRV) and CIRS-G scores (R²=0.33 vs 0.17 without HRV).
Conclusions: The HRV TI measured with a relatively novel wrist-worn PPG device was related to both objective (SPPB and TUG) and self-reported (SF-36 physical composite) measures of physical function. In addition, the HRV TI explained additional variance in self-reported physical function and cumulative illness severity beyond traditionally measured aspects of physical health. Future steps include longitudinal tracking of changes in both HRV and physical function, which will add important insights regarding the predictive value of HRV as a biomarker of physical health in older adults.

(JMIR Mhealth Uhealth 2019;7(10):e13757) doi: 10.2196/13757

KEYWORDS
wearable technology; aging; electrocardiogram; geriatric assessment

Introduction

Background

Heart rate variability (HRV), or variation in beat-to-beat intervals of the heart, is a quantitative measure of autonomic regulation of the cardiovascular system that reflects the ability of the system to react to stressors [1-3]. Low HRV indicates improper coordination between the sympathetic and parasympathetic nervous systems and is a well-established predictor of future cardiovascular disease [4-8]. HRV is also linked to other aspects of health that are directly impacted by autonomic function such as self-regulatory capacity and psychological and physiological stress [5,9,10]. Most HRV metrics are shown to decline normally with age, primarily during younger decades, and some of these metrics increase again after the seventh decade [4,11-15]. In studies of older adults, low HRV measured by electrocardiogram (ECG) recordings has shown preliminary relationships with physical frailty [16-19]. Thus, measurement of HRV may add important information to an assessment of older adults’ physical functioning and help to identify individuals at higher risk for physical decline.

Heart Rate Variability Measurement

HRV is traditionally measured in clinical or laboratory settings using standard ECG equipment or in the field with a 24-hour Holter monitor [20]. Although accurate, these measures are time consuming, require expert setup, and are obtrusive, making them less appropriate for health assessments and unlikely to be routine. Recent advances in technology enable measurement of HRV in more ecologically valid settings with unobtrusive wearable devices [21,22]. For example, wrist-worn devices capture HRV via photoplethysmography (PPG) sensors that detect blood volume changes in the microvasculature with each heartbeat [22-24]. These blood volume changes allow the determination of a PPG-derived peak-to-peak (P-P) interval, which is a valid proxy measure of the R-R interval derived from ECG recordings [25-27]. P-P or R-R intervals are also known as interbeat intervals and reflect the durations between successive heartbeats. HRV can be calculated from interbeat intervals with many different metrics (eg, time domain: standard deviation of normal-to-normal index; geometric: triangular index [TI]; and frequency domain: low frequency to high frequency power ratio) [3,20,28]. Selecting the appropriate HRV metric requires careful consideration of the length of recording and quality of the data. Wrist-worn PPG devices for HRV monitoring are increasing in popularity because of their ease of use; however, there is very little literature regarding the clinical use of these devices [29]. A critical step toward understanding the utility of these devices is to determine if and how HRV metrics derived from wrist-worn PPG devices relate to the physical function of older adults. To our knowledge, this is the first study using a wrist-worn PPG device to derive HRV in an older adult population.

Study Purpose

If HRV metrics derived from a wrist-worn wearable are related to clinical measures of physical function and further explain variability in physical health status, a wearable tool could be a useful addition to regular clinical geriatric health assessments for older adults. The primary aim of this study was to investigate associations between HRV measured with a wrist-worn PPG device, the Empatica E4 sensor (Empatica Inc) [30,31], and widely used and validated clinical measures of physical function, including Short Physical Performance Battery (SPPB) scores, Timed Up and Go (TUG) scores [32-34], and self-reported physical function (SF-36 physical composite scores) [35], in a cohort of older adults living independently in a continuing care senior housing community (CCSHC). We selected the SPPB, TUG, and SF-36 as objective and subjective measures of physical capacity and function, respectively, because they are a few of the most widely used measures of physical function in clinical geriatric research and health assessments [36-39]. Our primary hypothesis was that lower HRV would be associated with lower physical function. We further investigated whether HRV could explain additional variability in physical health status, as measured by the SF-36 physical composite scores and the presence and severity of physical comorbidities via the Cumulative Illness Severity Scale for Geriatrics (CIRS-G) [40], beyond traditionally documented aspects of health such as age, gender, race, blood pressure, medication and alcohol use, smoking status, and anthropometric measurements.

Methods

Participants

The University of California San Diego (UCSD) Human Research Protections Program (HRPP) approved the study protocol. Research staff recruited participants from a CCShC in San Diego County via short presentations using an HRPP-approved script and flyers. We recruited 77 participants living in the independent living sector of the CCShC, and all of them provided written informed consent before study participation. The inclusion and exclusion criteria for enrollment were (1) English-speaking individuals older than 65 years, (2) ability to complete study assessments, and (3) no known diagnosis of dementia or other disabling illness.
Procedures
We evaluated participants during the baseline assessment in an ongoing study of older adults between 65 and 95 years of age [41]. This assessment encompassed a thorough examination of domains typically included in a geriatric health evaluation: sociodemographic and medical health information, physical function measurements, cognitive measurements, and additional assessments of everyday function. The mean duration of the assessments was 2:34:31 (HH:MM:SS; range 1:09:30-4:38:10). The range of activities performed during this assessment reflected the types of tasks that an individual may encounter in their daily lives, and thus, the measure of HRV obtained from this time period is more similar to HRV metrics obtained over longer durations (eg, 24 hours), reflecting the cardiovascular system’s response to a range of environmental stimuli and workloads, as opposed to short-term measurements (eg, <5 min) that reflect immediate responses to a particular stimulus [3].

For this investigation, we characterized participants with sociodemographic and clinical information including age, sex, race, education, body mass index (BMI), waist-to-hip ratio, blood pressure, medication use (antihypertensives and antidepressants), presence and severity of physical comorbidities reported with the CIRS-G [40], and capacity for everyday functioning with the UCSD Performance-Based Skills Assessment Brief [42].

Physical Function Measurement
We assessed physical functioning with both objective, capacity-based measures and subjective, self-reported measures. Capacity-based measures included the SPPB and TUG test, both of which have been shown to be valid, objective measures of lower extremity function and mobility [32-34]. The SPPB has 3 subcategories: the ability to stand for 10 seconds with feet in 3 different positions (side by side, semitandem, and tandem), 2 timed trials of a 3- or 4-m walk (faster of the two), and the time it takes to rise from a chair 5 times. Scores range from 0 to 12, with higher scores indicating better lower extremity function. The TUG test is scored by the amount of time it takes to rise from a chair, walk 3 m at a comfortable pace, turn, return to the chair, and sit down. A shorter time to complete the test indicates better mobility. For self-reported physical functioning, we used the physical composite score from the Medical Outcomes Study Short Form 36 (SF-36), which is a globally used questionnaire for assessing 8 dimensions of health-related quality of life [35]. The physical composite score is an aggregate of the 8 scale scores reflecting self-reported physical health [43].

Heart Rate Variability Measurement
We collected raw PPG signals and interbeat intervals for the calculation of HRV with a wrist-worn device called the Empatica E4 [30,31]. Research staff placed the Empatica E4 wristband on a participant’s nondominant wrist at the onset of the assessment. During the first few minutes, the staff member checked the quality of the PPG signal in real time through the E4 application to ensure proper wrist placement as recommended by Empatica. Wrist-worn PPG sensors tend to be more accurate at rest than during exercise because of contamination from movement artifacts and often require accelerometry technology to measure consistent or repetitive movements to minimize influence of these artifacts [21,44-47]. The PPG sensor in the Empatica E4 is designed to be robust against movement artifacts in that it can attenuate noise even when movements are not repetitive in nature, using an artifact removal technique based on a combination of multiple infrared light wavelengths [30].

The Empatica E4 provided continuous heart rate and interbeat intervals that were not interpolated. It removed interbeat intervals that corresponded to regions where the PPG signal was not clear. For postprocessing, we downloaded the interbeat intervals provided in a comma-separated value file format from the E4 connect application. We viewed each interbeat interval time series in MATLAB R2016a (MathWorks, Inc) and calculated the percentage of gaps in the data (ie, nonconsecutive interbeat intervals) for each participant. We removed 19 participants from data analyses because of poor-quality interbeat interval recordings with greater than 20% gaps between interbeat intervals that resulted in less than 20 min of recorded high-quality interbeat interval data. We then processed the remaining data with Kubios HRV standard software version 3.1.0 using a low threshold artifact correction to adjust for any remaining ectopic beats and the smoothing priors detrending method (default λ=500) [48,49].

Selecting an HRV metric requires careful consideration of the strengths and limitations of wrist-worn PPG devices. The Empatica E4 uses algorithms to remove errant interbeat intervals and provides the remaining clean data that contain regions where interbeat intervals are not necessarily consecutive [30,31,50]. To overcome this limitation and provide a robust and valid HRV metric, we selected a metric that is less sensitive to gaps between interbeat intervals called the HRV TI [44,51,52]. The HRV TI is a geometric index calculated as the integral of the density distribution of interbeat intervals divided by the maximum of the density distribution [28] with larger numbers indicating more favorable HRV. The HRV TI requires a longer recording period (approximately 20 min), is robust against missing interbeat intervals, and has good intraindividual reproducibility [28]. For these reasons, we selected the HRV TI and analyzed the longest recording period available for each participant (ie, full-length assessment) for greater measurement stability.

Statistical Analyses
Given the exploratory nature of this study, we did not perform a formal a priori power analysis. We performed correlational statistical analyses in SPSS version 25 and generalized estimating equation (GEE) regression with backward elimination in R version 3.4.1. We used Spearman correlation coefficients to assess bivariate correlations between the HRV TI measured across the full assessment period and SPPB, TUG, and SF-36 physical composite scores. We removed an additional 5 participants from correlational analyses for the suspected presence of arrhythmia based on an HRV TI greater than 20.42 [53]. We were also missing SPPB scores for 1 participant, TUG scores for 2 participants, and SF-36 physical composite scores for 4 participants. We included the remaining participants in analyses (n=52, SPPB; n=51, TUG; and n=49, SF-36 physical scores for 2 participants, and SF-36 physical composite scores for 2 participants, and SF-36 physical composite scores for 4 participants. We included the remaining participants in analyses (n=52, SPPB; n=51, TUG; and n=49, SF-36 physical scores for 2 participants, and SF-36 physical composite scores for 4 participants. We included the remaining participants in analyses (n=52, SPPB; n=51, TUG; and n=49, SF-36 physical
composite). We conducted Mann-Whitney U tests to determine if there were differences between the participants removed from analyses and those retained. For our primary analyses, we first examined correlations between the HRV TI and variables that have been previously shown to be related to HRV including age, BMI, blood pressure, and mean heart rate [4,54]. We also looked for group differences in the HRV TI between men and women and between participants on versus off antihypertensive and antidepressant medications [16]. We used these results to determine whether to adjust for any of these variables using partial correlations (ie, control for one or more of these variables using nonparametric partial correlation syntax in SPSS if any covariates could potentially modify our primary bivariate relationships of interest). For the regression analyses, we included n=42 participants with complete data on all potential covariates (age, gender, race, blood pressure, BMI, waist-to-hip ratio, SPPB and TUG scores, antihypertensive and antidepressant use, alcohol use, smoking status, SF-36 physical composite or CIRS-G scores depending on the model, and HRV). We checked variance inflation factors (VIFs) for each covariate and considered a covariate with VIF greater than 3 as having high multicollinearity [55]. We rebuilt linear models and recalculated VIFs for each covariate after excluding covariates with a VIF greater than 3 and repeated this exclusion process until all covariates had a VIF less than 3. To build the regression model, we started with all variables of VIF less than or equal to 3 in the GEE model [56]. We then iteratively removed the variable with the largest P value, rebuilt a new GEE model based on the remaining variables, and recalculated P values for each variable. We repeated this process until all remaining variables had a P value at a threshold of less than or equal to .2. The backward elimination procedure ensured minimum bias in the final model [57]. We derived $R^2$ values using the linear model. We established statistical threshold for each family of statistical tests: correlational significance to a Bonferroni adjusted $P \leq .017$ (0.05/3 primary outcomes of interest) and $P \leq .025$ for the regressions (0.05/2 regression analyses).

## Results

We did not find significant differences for any variable in Table 1 between the individuals removed from correlational analyses and those that we retained ($P = .17$ to $P = .99$). Heart rate remained within a normal range during the baseline assessment: mean of 76 beats per minute (range 63-93). Mean heart rate over the course of the assessment was inversely correlated with the mean interbeat interval length $\rho = -0.78$ and $P \leq .001$ (Figure 1) as expected, because a higher heart rate corresponds to a shorter duration between beats and vice versa.

There were no significant differences in the HRV TI based on sex ($U = 242.5; P = .40$), antihypertensive medication use ($U = 271.5; P = .37$), or antidepressant medication use ($U = 147.0; P = .59$). We also did not observe significant correlations between any potential covariate and both HRV TI and physical function measures (ie, variables that could potentially modify these bivariate relationships; Table 2). Therefore, we did not adjust for any variables using partial correlations for our primary hypothesis regarding relationships between HRV TI and physical function measures. HRV TI was significantly related to SPPB, TUG, and SF-36 physical composite scores (Table 2 and Figure 2). We also observed an inverse relationship between HRV TI and CIRS-G scores (Figure 2).

Results of the regression revealed that the HRV TI explained a significant proportion of variance in physical health status as characterized by both SF-36 physical composite and CIRS-G scores. $R^2$ values increased from 0.11 to 0.28, with the HRV TI included in estimating SF-36 physical composite scores, and from 0.17 to 0.33, with the HRV TI included in estimating CIRS-G scores. Additional significant correlates for these measures of physical health status following the backward elimination procedure included gender, BMI, medication use, and smoking status (Table 3). VIF values were all less than 2.0, suggesting minimal collinearity.
Table 1. Sociodemographic information and other variables collected from participants.

<table>
<thead>
<tr>
<th>Variable name</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years; n=77), mean (SD); range</td>
<td>82.9 (6.7); 67-98</td>
</tr>
<tr>
<td>Gender (n=77); number of females, n (%)</td>
<td>52 (68)</td>
</tr>
<tr>
<td>Education (years; n=77), mean (SD); range</td>
<td>15.8 (2.4); 12-20</td>
</tr>
<tr>
<td>Race (n=77); number of whites, n (%)</td>
<td>68 (88)</td>
</tr>
<tr>
<td>Body mass index (kg/m²; n=76), mean (SD); range</td>
<td>27.9 (4.9); 19-43</td>
</tr>
<tr>
<td>Waist-to-hip ratio (ratio; n=75), mean (SD); range</td>
<td>0.87 (0.08); 0.71-1.06</td>
</tr>
<tr>
<td>Systolic blood pressure (mm Hg; n=74), mean (SD); range</td>
<td>134 (17); 100-167</td>
</tr>
<tr>
<td>Diastolic blood pressure (mm Hg; n=74), mean (SD); range</td>
<td>75 (9); 56-94</td>
</tr>
<tr>
<td>University of California San Diego Performance-Based Skills Assessment Brief (score; n=77), mean (SD); range</td>
<td>76.1 (12.5); 37.4-100.0</td>
</tr>
<tr>
<td>Cumulative Illness Rating Scale for Geriatrics (score; n=76), mean (SD); range</td>
<td>8.8 (3.2); 2-15</td>
</tr>
<tr>
<td>Antihypertensives (n=75); number of participants using, n (%)</td>
<td>48 (62)</td>
</tr>
<tr>
<td>Antidepressants (n=75); number of participants using, n (%)</td>
<td>13 (17)</td>
</tr>
<tr>
<td>Current smoker (n=75); number of participants smoking, n (%)</td>
<td>1 (1)</td>
</tr>
<tr>
<td>Alcohol use (n=68), n (%)</td>
<td></td>
</tr>
<tr>
<td>Lifetime abstainer (number of yes; [% abstaining])</td>
<td>6 (9)</td>
</tr>
<tr>
<td>Current infrequent drinker (number of yes; [% drinking])</td>
<td>36 (53)</td>
</tr>
<tr>
<td>Current regular drinker (number of yes; [% drinking])</td>
<td>17 (25)</td>
</tr>
<tr>
<td>Former drinker (number of yes; [% used to drink])</td>
<td>9 (13)</td>
</tr>
<tr>
<td>Short Physical Performance Battery (score out of 12; n=76), mean (SD); range</td>
<td>8.2 (2.7); 0-12</td>
</tr>
<tr>
<td>Timed up and Go time (seconds; n=74), mean (SD); range</td>
<td>11.0 (3.1); 6.1-23.0</td>
</tr>
<tr>
<td>Short Form 36 physical composite (score out of 100; n=70), mean (SD); range</td>
<td>41.8 (10.7); 18.3-60.1</td>
</tr>
<tr>
<td>Heart rate variability triangular index (n=53), mean (SD); range</td>
<td>11.4 (2.9); 3.7-17.2</td>
</tr>
</tbody>
</table>

Figure 1. Scatterplot of the relationship between mean heart rate across the entire evaluation and mean interbeat interval length: n=52; Spearman correlation=-0.78; P≤.001. bpm: beats per minute.
Table 2. Bivariate correlations among all variables. Correlations for primary hypotheses are in italics.

<table>
<thead>
<tr>
<th>Variables</th>
<th>( \rho ) (( P ) value)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SPPBA (^a)</td>
</tr>
<tr>
<td>HRV TI ((n=53))</td>
<td>0.41(^i)(;) ((.003))</td>
</tr>
<tr>
<td>SPPB</td>
<td>(-0.79(^k);) ((\leq .001))</td>
</tr>
<tr>
<td>TUG</td>
<td>(-0.39(^l);) ((.007))</td>
</tr>
<tr>
<td>SF-36 phys comp</td>
<td>(-0.65(^l);) ((\leq .001))</td>
</tr>
<tr>
<td>CIRS-G</td>
<td>(-0.20;) ((.16))</td>
</tr>
<tr>
<td>Age</td>
<td>(-0.14;) ((.34))</td>
</tr>
<tr>
<td>BMI</td>
<td>(-0.01;) ((.95))</td>
</tr>
<tr>
<td>Waist to hip ratio</td>
<td>(-0.13;) ((.36))</td>
</tr>
<tr>
<td>Sys BP</td>
<td>(-0.64(^l);) ((\leq .001))</td>
</tr>
<tr>
<td>Dia BP</td>
<td>(-0.03;) ((.85))</td>
</tr>
</tbody>
</table>

\(^a\)SPPB: Short Physical Performance Battery.  
\(^b\)TUG: Timed Up and Go.  
\(^c\)SF-36 phys comp: Short Form 36 physical composite score.  
\(^d\)CIRS-G: Cumulative Illness Rating Scale for Geriatrics.  
\(^e\)BMI: body mass index.  
\(^f\)Sys BP: systolic blood pressure.  
\(^g\)Dia BP: diastolic blood pressure.  
\(^h\)HR: heart rate.  
\(^i\)HRV TI: heart rate variability triangular index.  
\(^j\)\( P \leq .017.\)  
\(^k\)Not applicable.  
\(^l\)\( P \leq .001.\)
Figure 2. Scatter plots of the relationships between heart rate variability triangular index (TI) and (a) Short Physical Performance Battery scores (n=52), (b) Timed Up and Go scores (n=51), (c) Short Form 36 physical composite scores (n=49), and (d) Cumulative Illness Rating Scale for Geriatrics (CIRS-G) scores (n=52). HRV: heart rate variability; SF-36: Short Form 36.

Table 3. Results of the regression analyses (backward selection): remaining significant correlates of physical health status characterized by Short Form 36 physical composite scores and Cumulative Illness Rating Scale for Geriatrics scores. Adjusted $R^2$ of the full model for SF-36 scores=0.28 and for CIRS-G scores=0.33.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>Naïve SE</th>
<th>Naïve z</th>
<th>Robust SE</th>
<th>Robust z</th>
<th>$P$ value</th>
</tr>
</thead>
<tbody>
<tr>
<td>SF-36$^a$ physical composite score</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>22.01</td>
<td>5.81</td>
<td>3.79</td>
<td>6.05</td>
<td>3.64</td>
<td>≤.001</td>
</tr>
<tr>
<td>HRV TI$^b$</td>
<td>1.58</td>
<td>0.47</td>
<td>3.38</td>
<td>0.44</td>
<td>3.58</td>
<td>≤.001</td>
</tr>
<tr>
<td>Gender</td>
<td>5.88</td>
<td>3.01</td>
<td>1.96</td>
<td>2.79</td>
<td>2.11</td>
<td>.04</td>
</tr>
<tr>
<td>Antihypertensives</td>
<td>−5.95</td>
<td>2.71</td>
<td>−2.20</td>
<td>2.53</td>
<td>−2.35</td>
<td>.02</td>
</tr>
<tr>
<td>Smoking status</td>
<td>6.72</td>
<td>2.71</td>
<td>2.48</td>
<td>2.46</td>
<td>2.73</td>
<td>.006</td>
</tr>
<tr>
<td>CIRS-G$^c$ score</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>8.62</td>
<td>3.02</td>
<td>2.85</td>
<td>2.80</td>
<td>3.07</td>
<td>.002</td>
</tr>
<tr>
<td>HRV TI$^d$</td>
<td>−0.46</td>
<td>0.14</td>
<td>−3.41</td>
<td>0.10</td>
<td>−4.56</td>
<td>≤.001</td>
</tr>
<tr>
<td>Gender</td>
<td>−1.56</td>
<td>0.89</td>
<td>−1.75</td>
<td>0.81</td>
<td>−1.93</td>
<td>.05</td>
</tr>
<tr>
<td>BMI$^e$</td>
<td>0.15</td>
<td>0.95</td>
<td>1.70</td>
<td>0.07</td>
<td>1.96</td>
<td>.05</td>
</tr>
<tr>
<td>Antihypertensives</td>
<td>2.52</td>
<td>0.83</td>
<td>3.05</td>
<td>0.71</td>
<td>3.57</td>
<td>≤.001</td>
</tr>
<tr>
<td>Antidepressants</td>
<td>1.34</td>
<td>0.91</td>
<td>1.47</td>
<td>0.95</td>
<td>1.41</td>
<td>.16</td>
</tr>
</tbody>
</table>

$^a$SF-36: short form 36.

$^b$HRV TI: heart rate variability triangular index

$^c$CIRS-G: Cumulative Illness Rating Scale for Geriatrics.

$^d$HRV: heart rate variability.

$^e$BMI: body mass index.
Discussion

Principal Findings
In this study, we explored relationships between HRV measured using a wrist-worn device and both capacity-based and self-reported measures of physical function obtained during a standard clinical geriatric assessment of older adults living independently in a CCShC. Our primary hypothesis was supported as we observed significant relationships between the HRV TI and both capacity-based measures of physical function—that is, the SPPB and TUG [32-34] and self-reported physical function, as measured by the SF-36 physical composite score [35,43]. In addition, we observed that the HRV TI explained a significant proportion of variance in physical health status, as characterized by SF-36 physical composite and CIRS-G scores [40].

Relationships Between Heart Rate Variability, Physical Function, and Physical Health
The relationships between the HRV TI and both the SPPB and TUG and self-reported measures of physical function are noteworthy. Self-reported and capacity-based measures of physical capacity are reported to contribute information about physical function in different ways [58]. For example, self-reported measures are more effective in distinguishing among persons at lower levels of physical function who may not be able to complete capacity-based tests [59]. In contrast, self-reported measures fail to differentiate well among persons in the mid- or high range of functioning. Capacity-based measures provide information across a broader range of physical function and discriminate more effectively than self-reported measures at higher capacities [59]. We observed significant relationships between SPPB and TUG scores and self-reported physical function in our sample. We also observed an inverse relationship between the presence of physical comorbidities and self-reported physical function, which is consistent with previous reports demonstrating that older adults who report more health problems also report lower levels of perceived function [60]. Presence of physical comorbidities (CIRS-G) was related to the HRV TI but not to either SPPB or TUG scores in this study. Therefore, HRV TI may characterize a complementary aspect of older adults’ physical health that is not captured by physical capacity measures.

Previous work has identified relationships between HRV quantified via standard ECG or Holter monitoring and aspects of older adults’ physical function [5,16,19]. In particular, frailty appears to be linked to a loss of physiologic flexibility, as quantified by nonlinear HRV dynamics [16,17]. If changes in HRV precede physical decline, then regular measurement of HRV may be a useful biomarker for older adults. Low HRV, along with male sex, older age, and smoking, has also been reported to be a significant predictor of chronic diseases including hypertension and hyperglycemia and a diagnosis of diabetes within 12 years [61]. These findings suggest that monitoring of HRV may help to identify individuals at high risk for both physical decline and development of chronic diseases.

High HRV has been described as a measure of adaptability [62], which is directly related to maintenance of physical function. The consistent relationships observed between the HRV TI and measures of physical function and physical health in this study suggest that HRV TI quantified with a wearable device is indeed related to physical function and cumulative illness burden in older adults. These relationships were not accounted for by a direct association between HRV and age, suggesting that this measure of autonomic function may be useful for aging research. This finding is not surprising given that previous relationships between HRV and age have generally been observed across a much broader age span [4,14,15,28]. Although standard ECG and Holter monitor recordings have been useful for risk stratification in a variety of pathologic conditions [62-64], our findings suggest that a wearable device may also be useful for quantifying autonomic dysfunction in older adults. Although HRV was significantly associated with physical capacity measures, mean heart rate was not, suggesting that measurement of heart rate alone does not capture the relationship between autonomic regulation and physical function.

Wearable Technology for Heart Rate Variability Measurement
The accessibility of wearable technology [29] and the relationships between HRV derived from a wrist-worn wearable and measures of physical function suggest that incorporating wearables for HRV monitoring into clinical geriatric assessments could potentially help in the early detection of physical decline. In particular, watch-like devices are ideal because they are noninvasive and small and do not interfere with the activities performed during clinical geriatric health assessments. These devices are also relatively low cost, do not require expert setup, and may be more easily deployed than ECG equipment in remote home-health settings. In the future, such wearables may be ideal for monitoring the health of older adults, further enabling early detection of physical decline and early intervention to improve autonomic regulation and possibly delay deterioration of physical function. However, such long-term monitoring of HRV via wearable devices will necessitate careful consideration of the adoption of technology by older adults and ethical concerns such as data privacy and informed consent [65]. It may also be important to monitor HRV in younger populations, given the observed relationships to measures of physical health status in this study and the fact that declining physical function does not happen only in adults older than 65 years. HRV is known to decline with age, particularly before the age of 60 years, but these correlations are only modest [14]. Thus, perhaps those with poorer physical health also exhibit greater tendency for decline. If HRV metrics are related to the physical health status of younger adults as well, early identification of declining HRV may enable more timely intervention to mitigate further decline.

Future Goals for Heart Rate Variability Measurement
In this study, we measured HRV during a single assessment to demonstrate relationships between HRV metrics quantified with a wearable device and clinical measures of physical function; however, we intend to track this cohort of participants longitudinally. These findings will help determine whether HRV...
from a wearable device predicts or coincides with changes in physical function. We are also interested in the feasibility of more frequent HRV measurements using wrist-worn devices for older adults and plan to examine day-to-day variability of HRV metrics. The clinical value of HRV measurement is likely to be at the level of the individual and not the group; thus, identifying changes outside of an individual's normal HRV range may lend the most insight into incipient decline.

**Study Limitations**

This study had several limitations. This was a small sample of predominantly white and middle-to-upper class older adults, so our findings may not apply to other populations. We also had a majority of female participants, which may have prevented us from detecting differences in HRV due to gender. However, gender differences in HRV have been reported to disappear after the age of 50 years [14]. Furthermore, the cross-sectional nature of this investigation precludes the ability to draw causal relationships between the variables investigated. We investigated only 1 PPG device, the Empatica E4 sensor, in this study. However, we did not use an HRV metric provided by this device (Empatica does not offer these data). We simply used the device to collect raw interbeat interval data and then used the well-known Kubios HRV software [48] to calculate HRV TI. The presence of missing interbeat intervals influenced our choice of HRV metric (we selected the HRV TI instead of others) and, in particular, prevented us from using frequency-based HRV metrics as they are affected by missing interbeat intervals. As frequency-based metrics are reported to be more descriptive of the balance between sympathetic and parasympathetic activity [66], improving the quality of interbeat interval recordings from wearable devices is of considerable importance. We also had to remove a considerable number of participants from statistical analyses because of poor-quality interbeat interval recordings. Future research directed at refining wearable technology such as the Empatica E4 to improve the quality of recordings (eg, further minimize movement artifacts) will be essential in improving the utility of these devices. Finally, we cannot easily compare the HRV TI values obtained in this study to previously established norms for several reasons, including less frequent reporting of the HRV TI metric, differences in calculation methods, and differing durations of recordings (eg, often done over a 24-hour period).

**Conclusions**

In summary, we demonstrated that the HRV TI measured using a relatively novel wearable device, the Empatica E4 sensor, was related to both objective (SPPB and TUG) and self-reported (SF-36 physical composite) measures of physical function of older adults and cumulative illness burden (CIRS-G) collected during a single assessment. In addition, the HRV TI explained a significant proportion of variance in physical health status, as characterized by the SF-36 physical composite scores and CIRS-G scores, beyond typically measured aspects of physical health. The next steps include longitudinal tracking of changes in both HRV and physical function, which will add important insights regarding the possible predictive value of HRV as a biomarker of functional outcomes in older adults.

**Acknowledgments**

The authors thank all the study participants for their contributions to this work. The funding for this study was provided, in part, by the National Institute of Mental Health T32 Geriatric Mental Health Program (grant MH019934 to Dilip V Jeste [Principal Investigator]), the Stein Institute for Research on Aging at UCSD, and by IBM Research AI through the AI Horizons Network IBM-UCSD AI for Healthy Living program.

**Conflicts of Interest**

None declared.

**References**


Abbreviations

BMI: body mass index
CCHSC: continuing care senior housing community
CIRS-G: Cumulative Illness Rating Scale for Geriatrics
ECG: electrocardiogram
GEE: generalized estimating equation
HRPP: Human Research Protections Program
HRV: heart rate variability
PPG: photoplethysmography
SF-36: Short Form 36
SPPB: Short Physical Performance Battery
TI: triangular index
TUG: Timed Up and Go
VIF: variance inflation factor