The Promise of Ultra-Short-Term (UST) Heart Rate Variability Measurements

Fredric Shaffer, PhD, BCB, Steven Shearman, and Zachary M. Meehan
Truman State University, Center for Applied Psychophysiology, Kirksville, MO

Keywords: biofeedback, heart rate variability, concurrent validity

Researchers have investigated whether ultra-short-term (UST) heart rate variability values can replace traditional 5-minute values in clinical and optimal performance settings. Concurrent validity is the extent to which the results of a measurement correspond to a previously validated assessment of the same construct. Several studies either failed to specify their concurrent validity criteria or used an inappropriate statistical test. The authors proposed a rigorous standard and demonstrated that artifacted resting ultra-short-term heart rate variability values can achieve strong concurrent validity for diverse time-domain, frequency-domain, and nonlinear measurements in healthy undergraduates. Based on these findings, resting baselines as brief as 1 minute should be sufficient to measure heart rate, the standard deviation of the interbeat interval for normal beats (SDNN), and the square root of the mean squared difference of adjacent NN intervals (RMSSD) in clinical, optimal performance, and personal health assessment with individuals who resemble Truman State University undergraduates.

Heart rate variability (HRV) is the change in the time intervals between successive heartbeats. These millisecond-scale time periods are called interbeat intervals (IBIs; Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology [Task Force], 1996). Several interdependent regulatory mechanisms, which function on different time scales, produce HRV. Circadian rhythms and fluctuations in core body temperature and metabolism operate on the slowest time scale and influence 24-hour HRV recordings. The autonomic (sympathetic and parasympathetic), cardiovascular, and respiratory systems function on the fastest time scale and contribute to short-term (e.g., 5-minute) HRV measurements (Moss, 2004).

The interaction between the cardiovascular and respiratory systems via the baroreceptor reflex mechanism helps generate short-term HRV and contribute to the very slow to high frequencies of the HRV frequency spectrum. Baroreceptors detect changes in blood pressure from the aortic arch and internal carotid arteries. When you breathe, heart rate increases because inhalation briefly disables the vagal (parasympathetic) brake. When blood pressure rises about 5 seconds later, baroreceptors respond to this change by firing more rapidly. When you exhale, heart rate decreases because the vagal brake is re-engaged. Blood pressure falls after about 5 seconds (Gevirtz, Lehrer, & Schwartz, 2016).

HRV is an emerging marker for disease and adaptability. The ability of the brainstem cardiovascular center to vary the time intervals between heartbeats appears important to health. “An optimal level of variability within an organism’s key regulatory systems is critical to the inherent flexibility and adaptability or resilience that epitomizes healthy function and well-being. While too much instability is detrimental to efficient physiological functioning and energy utilization, too little variation indicates depletion or pathology” (Shaffer, McCraty, & Zerr, 2014). The respiratory-driven speeding and slowing of the heart, called respiratory sinus arrhythmia, is associated with improved regulation of blood pressure, heart rate, and gas exchange (Bernardi, Gabutti, Porta, & Spicuzza, 2001; DeBoer, Karemaker, & Strackee, 1987; Lehrer, 2013).

Clinicians and researchers describe HRV using three sets of metrics: time-domain, frequency-domain, and nonlinear measurements. Time-domain indices quantify the degree of variability across successive interbeat intervals. Frequency-domain indices quantify the absolute, normalized, or relative power (percentage of total power) within four HRV frequency bands. Finally, nonlinear indices quantify the unpredictability of a series of interbeat interval values. All three measurement categories are important because of their ability to predict disease and mortality (Bigger et al., 1992, 1995; Kleiger, Miller, Bigger, & Moss, 1987; Stein, Domitrovich, Huikuri, & Kleiger, 2005). The authors have defined all of the reviewed measurements in the Glossary.

The current standards for HRV assessment were specified by the Task Force of the European Society of Cardiology and
Despite the strong association between HRV time-domain, frequency-domain, and nonlinear measures and disease, and the availability of short-term norms for adults (Nunan, Sandercock, & Brodie, 2010; Umetani, Singer, McCraty, & Atkinson, 1998) and elite athletes (Berkoff, Cairns, Sanchez, & Moorman, 2007), HRV assessment has not been integrated into routine medical care outside of obstetrics or personal health assessment. The reasons that HRV measurement has not become mainstream are undoubtedly complex. One of these reasons is that a 5-minute HRV assessment is prohibitively long when compared with routine office or home measurements of blood glucose, blood pressure, core body temperature, heart rate, oxygen saturation, and weight.

To reduce the time cost of HRV measurement in medical, optimal performance, and personal health assessment, several researchers have investigated the concurrent validity of values obtained from ultra-short-term (UST) monitoring periods that are briefer than 5 minutes. Researchers establish concurrent validity when they demonstrate that the results of a measurement correspond to a previously validated measure of the same construct that is administered at the same time (Miller & Lovler, 1996). HRV concurrent validity research has statistically compared UST values with 5-minute recordings.

McNames and Aboy (2006) compared 10-second to 10-minute electrocardiogram (ECG) recordings with 5-minute recordings using archival data that consisted of ambulatory recordings on PhysioNet from 30 men (28.5–76 years) and 24 women (58–73 years). They evaluated the reliability and accuracy of four time-domain, six frequency-domain, and one nonlinear measurement. They concluded that their HRV indices were affected by ECG segment length and that values derived from different segment lengths were not comparable. Segments briefer than 1 minute estimated 5-minute mean heart rate more accurately than the other HRV metrics they calculated, although they did not specify their concurrent validity criterion.

Nussinovitch et al. (2011) obtained resting ECG data from 70 healthy volunteers, 37 men and 33 women (25–58 years) and compared 10-second and 1-minute values with 5-minute values using intraclass correlation coefficients. They found that 10-second and 1-minute values each achieved good correlations with 5-minute values for mean RR duration (the time between successive heartbeats) and RMSSD. They found no statistical correlation between their UST measurements and HRV triangular index, NN50, pNN50, SDNN, and absolute very low frequency (VLF), low frequency (LF), or high frequency (HF) power. They also failed to specify their concurrent validity criterion.

Finally, Baek, Cho, Cho, & Woo (2015) recorded resting photoplethysmograph (PPG) data from 467 healthy volunteers, 249 men and 218 women (8–69 years), and compared measurements from 10–270 seconds (10, 20, 30, 60, 90, 120, 150, 180, 210, 240, and 270 seconds) with 5-minute values. They calculated the Pearson product-moment correlation coefficient to assess the linear relationship between the UST and 5-minute values, and the Kruskal-Wallis test to determine whether the UST values were statistically different from the 5-minute values. They reported the following minimum durations required to estimate 5-minute values: 10 seconds for HR; 20 seconds for HF power; 30 seconds for RMSSD; 60 seconds for pNN50; 90 seconds for LF and normalized VLF, LF, and LF/HF; 240 seconds for SDNN and time frequency (TF); and 270 seconds for VLF. They also found that the minimum required durations varied with decade of age.

Research Question
The three reviewed studies suffered two limitations. First, two of the three studies failed to specify their concurrent validity criterion and the third used the wrong criterion. While the Kruskal-Wallis test is a nonparametric test for independent groups, Baek et al. (2015) used it to compare metrics calculated from different segment lengths obtained from the same individuals, violating its underlying assumptions. They should have used the Friedman test instead.

Second, two of the three reviewed studies did not address nonlinear HRV measurements, while the third (McNames & Aboy, 2006) only examined approximate entropy (ApEn).

The present study addressed these concerns by choosing a conservative concurrent validity criterion and nine nonlinear measurements, in addition to seven time-domain and six frequency-domain indices. This study addressed two questions. First, can UST resting measurements provide comparable estimates of 5-minute HRV measurements? Second, for each metric, at what epoch length will UST measurements account for at least 81% of the variance ($r = .90$) in 5-minute measurements?

Method
The investigators extracted 10-, 20-, 30-, 60-, 90-, 120-, 180-, and 240-second segments from 5-minute resting ECG
recordings of 38 healthy undergraduates, 20 men and 18
women (18–23), who participated in a study (Zerr et al.,
2014) that received Institutional Review Board approval.
These data had already been artifacted (i.e., false IBI values
due to problems like movement were corrected).

The researchers calculated UST time-domain, frequency-
domain, and nonlinear measurements for each segment
length from the start of the 5-minute recordings for each of
22 HRV parameters using CardioPro and Kubios 2.2
software. Seven time-domain indices included heart rate,
SDNN, RMSSD, NN50, pNN50, HRV triangular index,
and the width of the RR interval histogram baseline
(TINN). Six frequency-domain measures included absolute
VLF power, absolute and normalized LF and HF power, and
LF/HF ratio. Finally, nine nonlinear measures included
SD1, SD2, ApEn, sample entropy (SampEn), correlation
dimension (D2), detrended fluctuation analysis (DFA)
indices $\alpha_1$ (short-term fluctuations) and $\alpha_2$ (long-term
fluctuations), and recurrence plot analysis (RPA) indices
determinism (DET) and Shannon entropy (ShanEn).

Since IBI values are continuous variables (they can
assume an infinite number of possible values) and satisfy
the requirements of a ratio scale of measurement (e.g.,
equal intervals between adjacent values and an absolute
zero), the investigators measured concurrent validity
between the UST and 5-minute measurements using a
Pearson product-moment correlation coefficient. Because
the calculation of UST and 5-minute measurements from
the same data sample should be expected to inflate
correlation values, they selected a conservative criterion ($r$
= .90) for acceptable concurrent validity. This threshold
ensured that an acceptable UST measurement accounted for
at least 81% of the variability ($r^2 = 0.81$) in the
corresponding 5-minute value and a vanishingly low risk
of committing a Type 1 error (falsely concluding that the
two measurements were correlated).

**Results**

Resting UST measurements achieved acceptable concurrent
validity for all but one of the 5-minute HRV metrics
examined in this study. A 10-second segment successfully
estimated heart rate. A 60-second segment estimated
SDNN, RMSSD, NN50, and pNN50. A 90-second segment
estimated TINN, LF power, SD1, and SD2. A 120-second
segment estimated HRV triangular index and DFA $\alpha_1$. A
180-second segment estimated, LFnu, HF power, HFnu, LF/
HF power, SampEn, DFA $\alpha_2$, and DET. A 240-second
segment estimated ShanEn. No UST measurement success-
fully estimated correlation dimension (Table 1).

**Summary**

There is an inescapable trade-off between HRV sample
length and concurrent validity. While UST samples
impose a lower time cost than 5-minute samples, convenience is meaningless if the measurements achieve
poor concurrent validity. The criterion for concurrent
validity should be sufficiently rigorous that UST and 5-
minute values are interchangeable. The present study
employed a cut-off of $r = .90$, so that UST values would
account for at least 81% of the variability in 5-minute
values. Using this stringent standard, artifacted resting
UST measurements obtained from healthy undergraduates
achieved strong concurrent validity when compared with
5-minute mean HRV values.

Based on these findings, resting baselines as brief as 1
minute should be sufficient to measure heart rate, SDNN,
and RMSSD for individuals who resemble Truman State
University undergraduates as long as the data are carefully
artifacted. The authors encourage further research employ-
ing this study’s concurrent validity criterion to determine
minimum sample lengths for major demographic groups.
This is a prerequisite for the development of UST norms
and the adoption of routine HRV assessment in clinical,
opimal performance, and personal assessment applications
where time constraints might preclude traditional 5-minute
measurements.

**Glossary**

**Approximate entropy (ApEn):** nonlinear index of HRV that
measures the regularity and complexity of a time series.
**Correlation dimension (D2):** the number of variables
required to model nonlinear system dynamics.
**Detrended fluctuation analysis (DFA):** nonlinear index of
HRV that extracts the correlations between successive RR
intervals over different time scales and yields estimates of
short-term ($\alpha_1$) and long-term ($\alpha_2$) fluctuations.
**Determinism (DET):** the determinism of a time series.
**Frequency-domain measurements:** indices that quantify the
distribution of absolute, normalized, or relative power into
four HRV frequency bands.
**High frequency (HF) band:** HRV frequency range from
0.15–0.40 Hz that represents the inhibition and activation
of the vagus nerve by breathing at normal rates (respiratory
sinus arrhythmia).
**HFnu:** normalized power of the HF band.
**HRV triangular index:** the integral of the density of the RR
interval histogram divided by its height.
**Low frequency (LF) band:** HRV frequency range of 0.04–
0.15 Hz that may represent the influence of PNS, SNS, and
baroreflex activity when breathing at an individual’s resonance frequency.

LFnu: normalized power of the LF band.

NN50: the number of adjacent NN intervals that differ from each other by more than 50 milliseconds.
pNN50: the percentage of adjacent NN intervals that differ from each other by more than 50 milliseconds.

Poincaré plot analysis (PPA): visual display that plots every RR interval against the prior interval, creating a scatter plot, to identify patterns buried within a time series.

Recurrence plot analysis (RPA): nonlinear analysis which measures the frequency and duration of the recurrences of a dynamical system.

RMSSD: the square root of the mean squared difference of adjacent NN intervals.

RR duration: the time between successive heartbeats.

Sample entropy (SampEn): nonlinear index of HRV that was designed to provide a less biased measure of signal regularity and complexity than ApEn.

SD1: the standard deviation of the distance of each point from the $y = x$ axis that measures short-term HRV.

SD2: the standard deviation of each point from the $y = x + \bar{R}R$ average RR interval that measures short- and long-term HRV.

SDNN: the standard deviation of the interbeat interval for normal beats measured in milliseconds.

Time-domain measurements: indices that quantify the total amount of heart rate variability.

Time frequency (TF): total signal power; the sum of VLF, LF, and HF power.

TINN: the width of the RR interval histogram baseline.

<table>
<thead>
<tr>
<th>HRV Index</th>
<th>10 s</th>
<th>20 s</th>
<th>30 s</th>
<th>60 s</th>
<th>90 s</th>
<th>120 s</th>
<th>180 s</th>
<th>240 s</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heart rate</td>
<td>.960**</td>
<td>.973**</td>
<td>.977**</td>
<td>.986**</td>
<td>.990**</td>
<td>.991**</td>
<td>.996**</td>
<td>.997**</td>
</tr>
<tr>
<td>SDNN</td>
<td>.685**</td>
<td>.855**</td>
<td>.897**</td>
<td>.950**</td>
<td>.965**</td>
<td>.963**</td>
<td>.986**</td>
<td>.994**</td>
</tr>
<tr>
<td>RMSSD</td>
<td>.597**</td>
<td>.796**</td>
<td>.893**</td>
<td>.945**</td>
<td>.948**</td>
<td>.969**</td>
<td>.988**</td>
<td>.995**</td>
</tr>
<tr>
<td>NN50</td>
<td>.537**</td>
<td>.653**</td>
<td>.795**</td>
<td>.924**</td>
<td>.948**</td>
<td>.969**</td>
<td>.985**</td>
<td>.992**</td>
</tr>
<tr>
<td>pNN50</td>
<td>.653**</td>
<td>.693**</td>
<td>.820**</td>
<td>.930**</td>
<td>.947**</td>
<td>.968**</td>
<td>.985**</td>
<td>.994**</td>
</tr>
<tr>
<td>HRV triangular index</td>
<td>.389</td>
<td>.675**</td>
<td>.647**</td>
<td>.765**</td>
<td>.863**</td>
<td>.904**</td>
<td>.929**</td>
<td>.925**</td>
</tr>
<tr>
<td>TINN</td>
<td>.713**</td>
<td>.795**</td>
<td>.808**</td>
<td>.888**</td>
<td>.906**</td>
<td>.820**</td>
<td>.953**</td>
<td>.972**</td>
</tr>
<tr>
<td>VLF power</td>
<td></td>
<td>.320</td>
<td>.016</td>
<td>.149</td>
<td>.315</td>
<td>.539**</td>
<td>.526**</td>
<td>.901**</td>
</tr>
<tr>
<td>LF power</td>
<td>.675**</td>
<td>.607**</td>
<td>.795**</td>
<td>.825**</td>
<td>.919**</td>
<td>.934**</td>
<td>.985**</td>
<td>.999**</td>
</tr>
<tr>
<td>LFnu</td>
<td>.262</td>
<td>.462**</td>
<td>.550**</td>
<td>.563**</td>
<td>.693**</td>
<td>.774**</td>
<td>.936**</td>
<td>.994**</td>
</tr>
<tr>
<td>HF power</td>
<td>.492**</td>
<td>.303</td>
<td>.590**</td>
<td>.746**</td>
<td>.813**</td>
<td>.843**</td>
<td>.974**</td>
<td>.990**</td>
</tr>
<tr>
<td>HFnu</td>
<td>.261</td>
<td>.460**</td>
<td>.549**</td>
<td>.563**</td>
<td>.692**</td>
<td>.772**</td>
<td>.936**</td>
<td>.994**</td>
</tr>
<tr>
<td>LF/HF power</td>
<td>.026</td>
<td>.108</td>
<td>.349**</td>
<td>.646**</td>
<td>.695**</td>
<td>.800**</td>
<td>.962**</td>
<td>.998**</td>
</tr>
<tr>
<td>SD1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.948**</td>
<td>.969**</td>
<td>.988**</td>
<td>.995**</td>
</tr>
<tr>
<td>SD2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.965**</td>
<td>.961**</td>
<td>.984**</td>
<td>.994**</td>
</tr>
<tr>
<td>ApEn</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.441**</td>
<td>.658**</td>
<td>.885**</td>
<td>.964**</td>
</tr>
<tr>
<td>SampEn</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.800**</td>
<td>.768**</td>
<td>.930**</td>
<td>.978**</td>
</tr>
<tr>
<td>Correlation dimension</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.621**</td>
<td>.774**</td>
<td>.822**</td>
<td>.888**</td>
</tr>
<tr>
<td>DFA $\alpha_1$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.896**</td>
<td>.935**</td>
<td>.962**</td>
<td>.986**</td>
</tr>
<tr>
<td>DFA $\alpha_2$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.620**</td>
<td>.778**</td>
<td>.925**</td>
<td>.970**</td>
</tr>
<tr>
<td>DET</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.770**</td>
<td>.805**</td>
<td>.927**</td>
<td>.960**</td>
</tr>
<tr>
<td>ShanEn</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.614**</td>
<td>.651**</td>
<td>.905**</td>
<td>.960**</td>
</tr>
</tbody>
</table>

*p < .05 level (2-tailed). **p < .01 level (2-tailed).
Very low frequency (VLF): HRV frequency range of 0.003–0.04 Hz that may represent temperature regulation, plasma renin fluctuations, endothelial, and physical activity influences, and possible PNS and SNS contributions.

References


Correspondence: Fredric Shaffer, PhD, BCB, Truman State University, Center for Applied Psychophysiology, 100 E. Normal, Kirksville, MO 63501, email: fredricshaffer@gmail.com.