Limits of Agreement (LoA) Determination of Minimum Epochs for Estimating 5-Min Ultra-Short-Term (UST) Heart Rate Variability Measurements

Hannah Urban, Marilise Stamps, Jacqueline Pecoraro, and Fredric Shaffer, PhD, BCB-HRV (Faculty Mentor)

Truman State University
Center for Applied Psychophysiology

1

Gratitude

We are deeply grateful to the Foundation for Education and Research in Biofeedback and Related Sciences for student travel support, which has made our participation possible.

Foundation for Education and Research in Biofeedback and Related Sciences

2

Introduction

Ultra-short-term (UST) HRV measurements are based on less than 5 min of data.

There is growing interest in whether these values can achieve acceptable concurrent validity, which is the degree to which values obtained from proposed and established measurement procedures are correlated.

3

Introduction

McNames and Aboy (2006) compared the accuracy of HRV metrics calculated from 10-s to 10-min resting ECG recordings to those calculated from 5-min recordings, using 571 15-min records from three databases.

The strongest correlations were achieved with HF ms^2, SDSD, and RMSSD. The authors did not specify a minimum correlation.

4

Introduction

Salahuddin et al. (2007) obtained 5 min of resting ECG data during baseline and Stroop test conditions from 24 healthy students.

Their criterion was a nonsignificant Kruskal-Wallis comparison between UST and 150-s values.

5

Introduction

Their comparison was flawed because:
(1) the Kruskal-Wallis test is designed to compare *nominal* values from *different* subjects,

(2) nonsignificance can be due to insufficient statistical power, and

(3) nonsignificance does not ensure that UST values fall within an acceptable range

6

Introduction

Baek et al. (2015) recorded 5 min of resting PPG data from 467 healthy volunteers.

Their concurrent validity criteria were a Pearson correlation greater than 0.70 and a nonsignificant Kruskal-Wallis test.

7

Introduction

Their Pearson criterion ($r > 0.70$) was insufficiently rigorous since it allowed UST values to account for less than 50% of the variability in corresponding 5-min values.

As with Salahuddin et al. (2007), the Kruskal-Wallis test was inappropriate because the measurements were not independent of each other.

8

Introduction

Munoz et al. (2015) measured SDNN and RMSSD in 3,387 adults using the ECG method and analyzed data using Pearson's correlation coefficients, the Bland-Altman LoA method, and Cohen's *d*.

9

Introduction

The authors did not specify their maximum Bland-Altman allowable difference.

At 120 s, recordings achieved nearly perfect agreement with 240–300 s values ($r = 0.96$, bias = 0.41 for SDNN and $r = 0.99$, bias = 0.01 for RMSSD).

10

Introduction

Shaffer et al. (2016) recorded 5 min of resting ECG data from 38 healthy undergraduates and manually artifacted the IBIs.

They correlated 10-, 20-, 30-, 60-, 90-, 120-, 180-, and 240-s HRV metrics with 5-min metrics.

11

Introduction

The authors selected a conservative criterion of $r \geq 0.90$ to ensure that UST values would account for at least 81% of the variability in 5-min values.

Dr. Gevirtz recommended the Munoz et al. (2015) article and encouraged the authors to reanalyze these data using the Bland-Altman LoA method.

12

Introduction

Gravett et al. (in press) re-analyzed these data using the Bland-Altman LoA method (allowable difference within ± 5% of a 5-min value's range).

While they found that the Pearson and LoA methods agreed on the minimum sample required to estimate 5-min values for 11 metrics, they disagreed on 5.

13

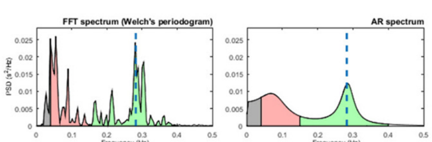
Introduction

The present study expanded our previous study with a larger sample (85 versus 35) and the addition of autoregression (AR) to Fast Fourier Transformation (FFT) frequency-domain values.

14

Introduction

Autogression "smooths" the spectral curve as shown below (graphic from Tarvainen & Niskanen, 2017).



15

Method

Participants

Seventy-five undergraduates (49 women and 26 men), 18 to 28 years of age, participated in this study.

16

Method

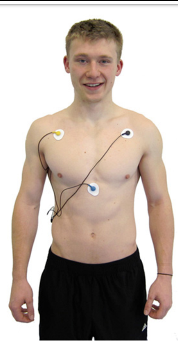
Apparatus

A Thought Technology ProComp Infiniti ™ system monitored ECG, HRV, and respiration. Active ECG electrodes were located on the upper torso.

A respirometer was positioned over the navel to measure excursion and respiration rate.

17

Method



18

Method

19

Method

Procedure

Subjects were stabilized for 5 min and then monitored for 5 min sitting upright, with eyes open, no feedback, and instructions to breathe normally.

The investigators extracted 10-, 20-, 30-, 60-, 90-, 120-, 180-, and 240-s segments from manually-artifacted 5-min resting ECG recordings.

20

Results

For each HRV metric, investigators calculated concurrent validity between each segment length (10 through 240 s) and its corresponding 5-min value using a Pearson Correlation Coefficient and the Bland-Altman LoA method.

21

Results

The authors required that an epoch satisfy a minimum Pearson correlation coefficient ($r \geq 0.90$) and a Bland-Altman allowable difference within $\pm$ 5% of a 5-min value's range.

22

Results

Epoch	HRV Metric
10 s	HR
90 s	RMSSD, SDNN, ARLFHF ms², FFTHF ms², SD1, and SD2.
120 s	pNN50 and FFTLFHF
150 s	ARHF ms², DFA α1, and SampEn
180 s	ARVLF ms², ARLF ms², ARLF nu, and ARHF nu
210 s	FFTVLF ms², FFTHF nu, and DET
240 s	D2

23

Results

No UST measurement successfully estimated NN50, FFTLF ms², or ApEn.

24

Discussion

The reviewed studies underscore the need for rigorous concurrent validity standards for UST HRV measurements.

In the present study, neither a Pearson correlation coefficient nor a Bland-Altman LoA comparison were adequate by themselves to estimate 13 of 27 metrics.

25

Discussion

A strong Pearson correlation ($r \geq 0.90$) did not ensure that an UST measurement fell within an allowable range ($\pm$ 5%) with respect to a 5-min value for 8 of 27 metrics.

Moreover, an UST measurement that fell within an allowable range did not account for at least 81% of the variability in a 5-min value for 5 of 27 metrics.

26

Discussion

The use of rigorous Pearson ($r \geq 0.90$) and Bland-Altman LoA ($\pm$ 5%) criteria successfully estimated 24 of 27 5-min metrics.

27

Discussion

Since the autoregression (AR) and Fast Fourier Transformation (FFT) methods produce different distributions, the minimum epochs needed to estimate frequency-domain values differed

28

Discussion

Epoch	HRV Metric
90 s	ARLFHF ms² and FFTHF ms²
120 s	FFTLFHF
150 s	ARHF ms²
180 s	ARVLF ms², ARLF ms², ARLF nu, and ARHF nu
210 s	FFTVLF ms² and FFTHF nu

29

Discussion

Clinicians and researchers should use the AR method to estimate 5-min HF ms² since no UST epoch successfully approximated FFTHF ms².

30

Discussion

Two applied findings from this study were:
(1) **10-s** resting baselines can estimate *heart rate.*
(2) **90-s** resting baselines can estimate *RMSSD* and *SDNN* (compared with 120 s for Munoz et al., 2015).

These findings are limited to individuals who resemble Truman State University undergraduates when data are carefully artifacted.

31

Discussion

The authors encourage further research employing this study's rigorous concurrent validity criteria to determine minimum sample lengths for major demographic groups.

32

Discussion

Clinicians should not use these measurements in place of conventional 5-min and 24-hr metrics until measurement protocols are standardized and normative values for healthy nonathlete, optimal performance, and clinical populations are established.

33

Discussion

34

References

Baek, H. J., Cho, C. H., & Woo, J. M. (2015). Reliability of ultra-short-term analysis as a surrogate of standard 5-min analysis of heart rate variability. *Telemedicine and e-Health, 21*(5), 404-414. doi:10.1089/tmj.2014.0104

Gevirtz, R. N., Lehrer, P. M., & Schwartz, M. S. (2016). Cardiorespiratory biofeedback. In M.S. Schwartz & F. Andrasik (Eds.). *Biofeedback: A practitioner's guide* (4th ed.). New York: The Guilford Press.

McNames, J., & Aboy, M. (2006). Reliability and accuracy of heart rate variability metrics versus ECG segment duration. *Med Biol Eng Comput, 44*(9), 747-756.

Munoz, M. L., van Roon, A., Riese, H., Thio, C., Oostenbroek, E., Westrik, I., de Geus, E. J., Gansevoort, R., Lefrandt, J., Nolte, I. M., & Snieder, H. (2015). Validity of (ultra) short recordings for heart rate variability measurements. *PLoS One, 10*, e0138921. doi:10.1371/journal.pone.0138921

35

References

Salahuddin, L., Cho, J., Jeong, M. G., & Kim, D. (2007). Ultra short term analysis of heart rate variability for monitoring mental stress in mobile settings. *Conf Proc IEEE Eng Med Biol Soc*, 2007, 4656-4659.

Shaffer, F., McCraty, R., & Zerr, C. L. (2014). A healthy heart is not a metronome: An integrative review of the heart's anatomy and heart rate variability. *Frontiers in Psychology*. doi:10.3389/fpsyg.2014.01040

Tarvainen, M. P., & Niskanen, J.-P. (2017). *Kubios HRV version 3.0 user's guide*. Kuopio, Finland: University of Finland.

36