MrOS Heart Rate Variability Data Release

Original Author: Farhad Kaffashi, Kenneth Loparo, Lily Lui

The Heart Rate Variability (HRV) dataset was obtained from ancillary study #108: Electrophysiologic and Respiratory Predictors of Incident Arrhythmia in Sleep Apnea. The data is from the PSG recording at the sleep visit, and was put together at Case Western.

There are 2911 MrOS participants with usable PSG data, but not all variables included in the dataset are complete for all participants. Overall, this data contains 2388 men with HRV data.

Among the 523 men who didn't have HRV data, 121 were pacemaker participants or the quality of the ECG was bad, 136 had prevalent PSG-identified arrhythmias, and there are 266 participants whose files cannot be found continuous 5 minutes epoch without any artifact or arrhythmia.

Software and Method

Heart Rate Variability (HRV) includes both conventional and novel measures that are based on time domain and frequency domain analysis. Conventional time domain HRV measures have been defined for NN (normal-to-normal) intervals, which correspond to intervals that are detected between adjacent QRS complexes of a continuous artifact free ECG [5]. NN specifically refers to the intervals between normal heartbeats resulting from sinus node depolarizations that are included in the analysis. There are two main sources of ECG artifact, measurement and physiological, depending on the source,. For example, movement and bad electrode connections are measurement artifacts and arrhythmias (PVC or PAC) are physiological artifacts. Therefore continuous artifact free ECG segments are to be used in HRV analysis and the criteria for artifact removal used in this work will be discussed in details later.

Conventional HRV uses statistical measures such as mean, standard deviation (SD), and root mean square (RMS) to quantify the variability. Statistical measures can depend on the length (window) of the data being analyzed so it is common to use 5-minute artifact free ECG epochs for short-term variability analysis and 24-hour recordings for longer-term studies. Comparisons of HRV statistical measures should be done only for data of similar length. Typical statistical HRV quantities include: MNN (mean of the NN intervals), SDNN (the mean of the 5-minute standard deviations of NN intervals calculated over 24 hours), SDANN (standard deviation of the average NN intervals calculated over 5-minute periods), RMSSD (square root of the mean squared differences of successive NN intervals). NN50 (number of interval differences of successive NN intervals greater than 50 ms), and pNN50 (proportion derived by dividing NN50 by the total number of NN intervals), see [5] for more details.

Other time series measures, such as those derived from the Poincare′ plot [7, 8], provide useful information obtained from R-R interval data by presenting and quantifying the variations in the ECG time series over different periods of time, both short- and long-term variability measures can be derived from the analysis, which in complementary to to the standard time series variability measures discussed in [5]. Consecutive beat-to-beat data is displayed on a scatter plot as shown in Figure 1 and the distribution of points is analyzed by defining SD1C and SD2C, that are measures of the spread of the distribution points along orthogonal direction that define the minor and major axes of a hypothetical ellipse that fits the data [9]. SD1C represents for the Standard Deviation of the short-term variability of the data and SD2C represents the Standard Deviation of long-term variability.

Standard spectral analysis for equally spaced (uniformly sampled) time series data is based on the Fast Fourier Transform (FFT). However, the R-R interval data is not uniformly sampled because the time between two consecutive heartbeats is not constant. Therefore a technique for calculating the spectral representation of unevenly sampled data has been used [3] for spectral analysis of the R-R time series. Using this approach the normalized low frequency power (from 0.04Hz to 0.15Hz) and the normalized high frequency power (from 0.15Hz to 0.4HZ) have been calculated [1].

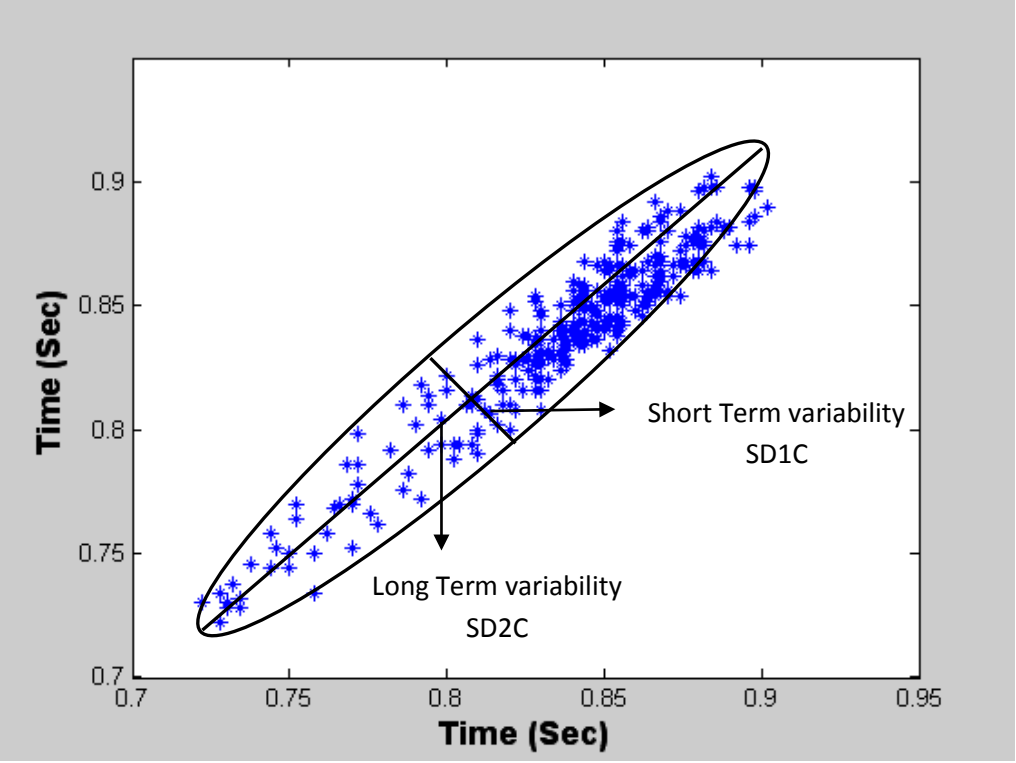

Figure 1. Poincare′ plot x-axes represent current R-R intervals and y-axes previous R-R intervals.

For nonlinear analysis, Detrended Fluctuation Analysis (DFA) has been considered.

DFA is a technique introduced by Peng at al. [10] to quantify the long-range correlation behavior in nonstationary physiological time series data. To apply DFA to ECG time series data, long-range correlations between interbeat intervals separated by several beats are detected by investigating the scaling behavior of heartbeat fluctuations on different time scales disregarding trends and nonstationary characterisitcs in the data. Penzel at al. [11] have defined small and large time scales to determine if there are short-term and long-term correlations in the heartbeat time series signal. Parameters Alpha1 and Alpha2 are estimated from the time series data, where Alpha1 corresponds to time scales between 10 and 40 heartbeats characterizing short-term correlation behavior, and Alpha2 corresponds to time scales between 70 and 300 heartbeats to observe long-term correlation behavior. The instantaneous R-R interval, DFA analysis results, calculated Aplha1 and calculated Alpha2 have been shown in Figure 2. In our analysis, we have used the same approach to estimate Alpha1 and Alpha2.

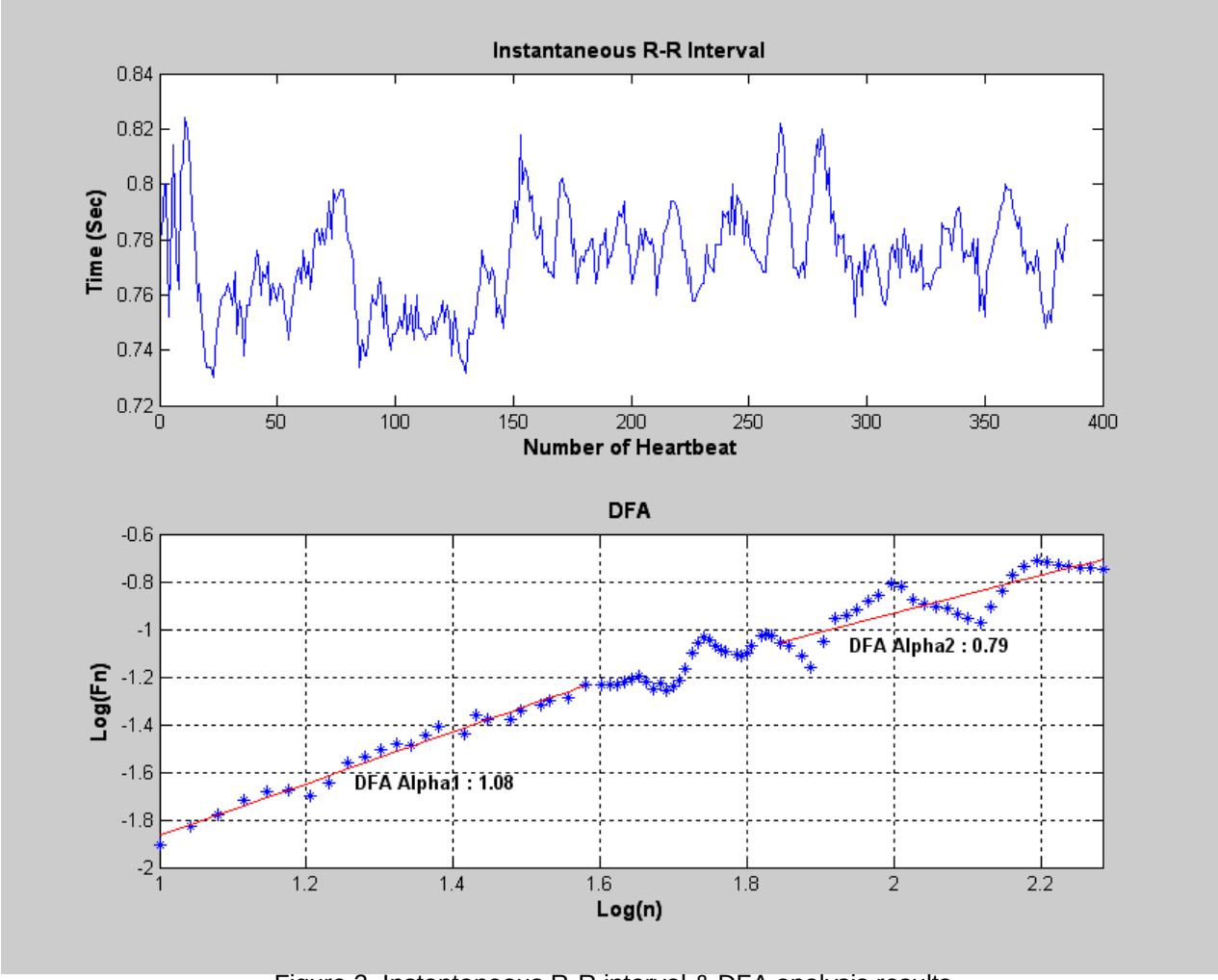

Figure 2. Instantaneous R-R interval & DFA analysis results

In order to compute the various HRV measures and visualize the ECG data and analysis results we have developed two tools referred to as “HRV Compute” and “HRV

View". The HRV Compute tool shown in Figure 4 allows users to select the studies to be analyzed, select the HRV measures to be computed, adjust the processing parameters, automatically process the studies, and save the analysis results of each study to a file. The following HRV measures are implemented in HRV Compute tool:

MNN : Mean of normal to normal heart beats
SDNN : Standard deviation of normal to normal heart beats
CV : Coefficient of variation of normal to normal heart beats
SD1C : Short term variability from Poincare plot
SD2C : Long term variability from Poincare plot
SD Ratio : Short to long term variability of Poincare plot
LFP : Normalized low frequency power
LHP : Normalized high frequency power
LHR : Low to high frequency power ratio
DFA-Alpha1 : Detrended Fluctuation Analysis for short term correlations in short time scales corresponding to log of 10 < t < 40 beats
DFA-Alpha2 : Deterend Fluctuation Analysis for long term correlations in long time scales corresponding to log of 70 < t < 300 beats

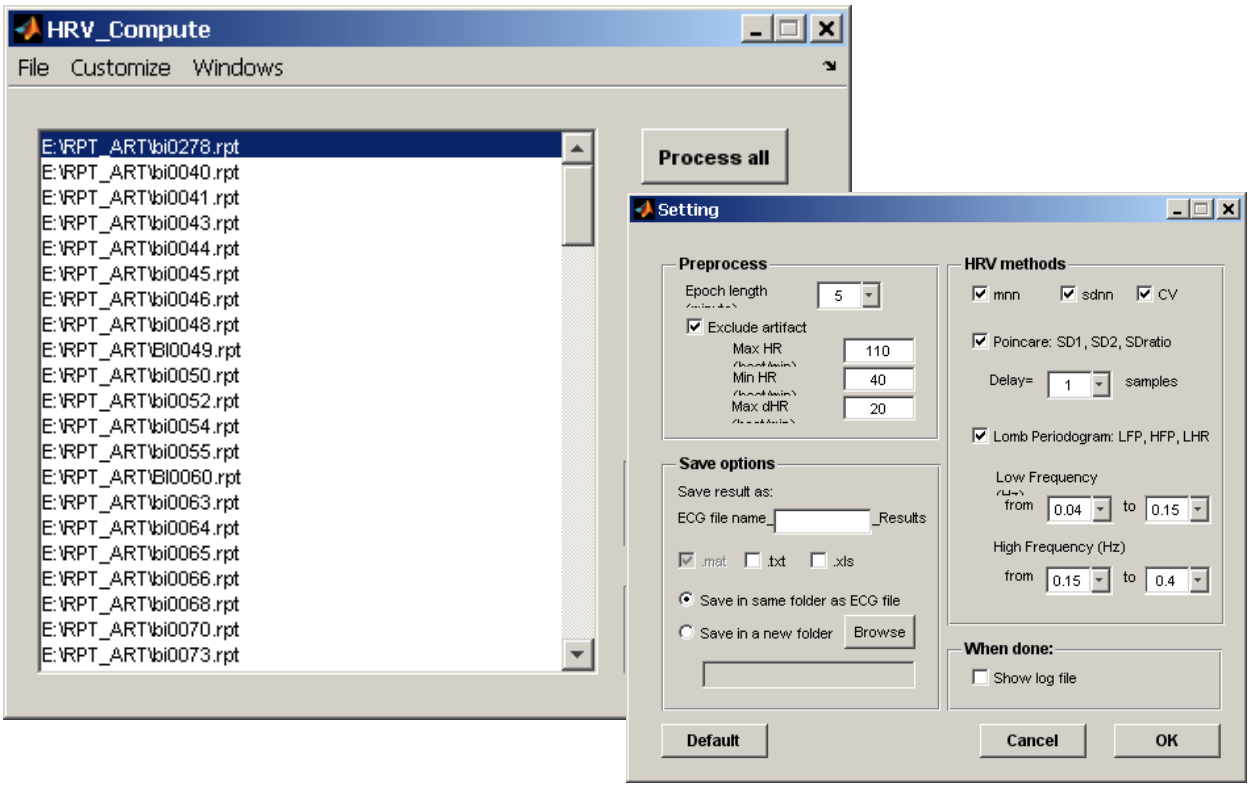

Figure 4. HRV Compute Tool

The HRV View, Figure 5, is able to load the analysis results and verify the entire collection of selected 5 minutes epochs quickly. In addition, users can review the raw data associated with each epoch.

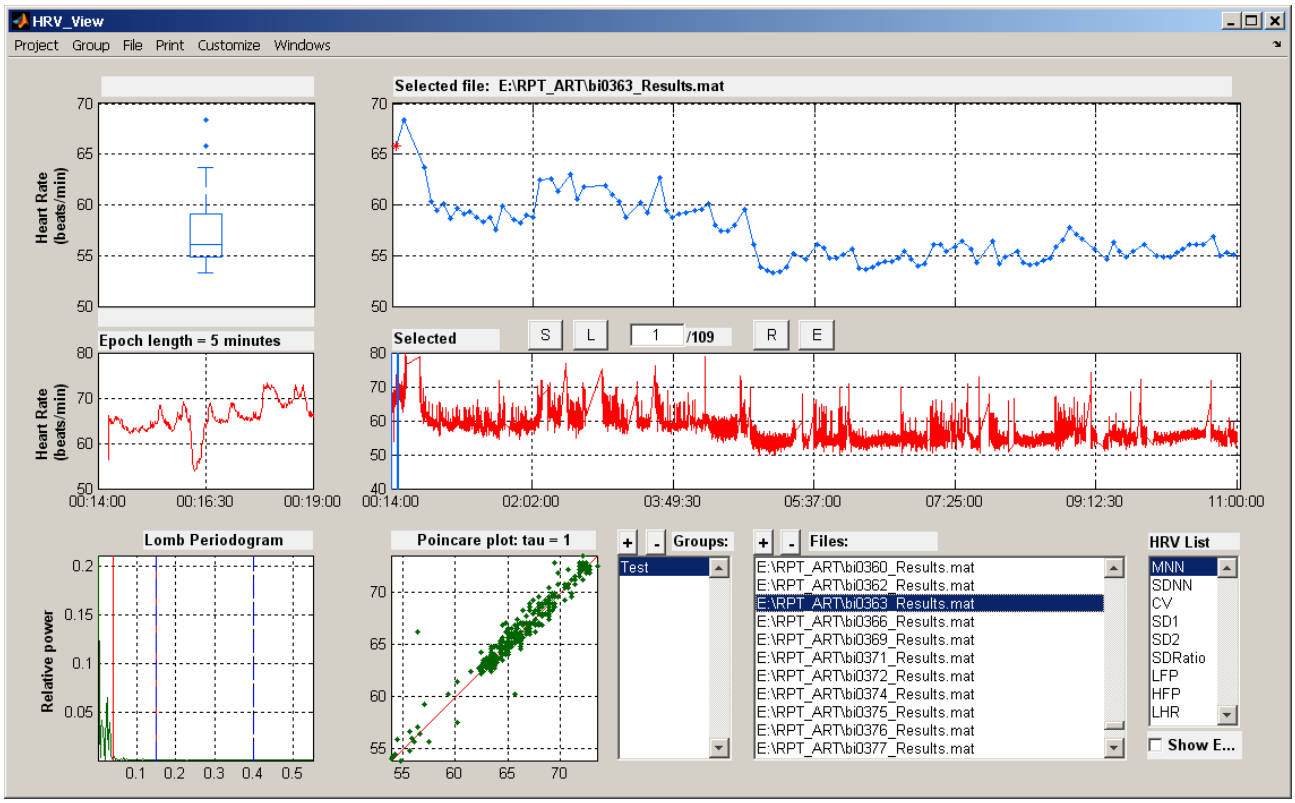

Figure 5. HRV View Tool

The following criteria has been considered to identify the artifact free segments of ECG to compute the HRV measures:

- 5 minutes of continuous non-overlapping ECG has been considered as the epoch length.
- Somte program (Compumedics) has been used during clinical review of the ECG data within the sleep period. The program automatically processes the data, identifies the R-beats, identifies arrhythmias (PVC or PAC), as well as other artifacts in the data prior to clinical review. Reviewers then go through the results from the automated Somte analysis and make sure the identified points are correct and make corrections if needed. In particular, reviewers go through the results of detected and annotated arrhythmias and visually validate each point. Then, the results of each detected heartbeat along with type information are saved in a database file, which is then used in processing of the ECG to

generate the HRV measures. In the database file, the type information for each identified R-beat using the following coding system:

- 1 – sinus beat
- 2 – VE (PVC)
- 3 – SVE (PAC)
- Otherwise – artifact

Therefore, only R-beats that are sinus beats (Type: 1) are accepted as normal beats subsequently used in the HRV analysis.

- Generally artifacts in the ECG data will result in missing detected R-beats or detecting additional points between two real R-points. In these cases, we have the duration of the R-R interval that is either shorter or longer than the regular R-R interval due to artifact. In order to identify these points, we have defined simple thresholds on the instantaneous heart rate that is calculated from consecutive heartbeats within each epoch. Therefore, any heartbeat lower than 30 beats per min or greater than 180 beats per min is accepted as artifact and cannot be included in a HRV analysis epoch. In addition, the rate of change in instantaneous heart rate has been considered as another criteria for detecting “bad” data, and any changes in heart rate that exceed 80 beats per minute between two consecutive R-R intervals has been accepted as artifact and excluded from the HRV analysis.

[1] Bauer, Axel, Marek Malik, Georg Schmidt, Petra Barthel, Hendrik Bonnemeier, Iwona Cygankiewicz, Przemyslaw Guzik et al. "Heart rate turbulence: standards of measurement, physiological interpretation, and clinical useInternational Society for Holter and Noninvasive Electrophysiology Consensus." *Journal of the American College of Cardiology* 52, no. 17 (2008): 1353-1365.

[2] Watanabe, Mari A., and Georg Schmidt. "Heart rate turbulence: a 5-year review." *Heart rhythm* 1, no. 6 (2004): 732-738.

[3] Press, William H., and George B. Rybicki. "Fast algorithm for spectral analysis of unevenly sampled data." *The Astrophysical Journal* 338 (1989): 277-280.

[4] Park, Soojin, Farhad Kaffashi, Kenneth A. Loparo, and Frank J. Jacono. "The use of heart rate variability for the early detection of treatable complications after aneurysmal subarachnoid hemorrhage." *Journal of clinical monitoring and computing* 27, no. 4 (2013): 385-393.

[5] Heart rate variability. "Standards of measurement, physiological interpretation, and clinical use. Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology." *Circulation* 93 (1996): 1043-1065.

[6] Fishman, Mikkel, Frank J. Jacono, Soojin Park, Reza Jamasebi, Anurak Thungtong, Kenneth A. Loparo, and Thomas E. Dick. "A method for analyzing temporal patterns of variability of a time series from Poincare plots." *Journal of Applied Physiology* 113, no. 2 (2012): 297-306.

[7] Brennan, Michael, Marimuthu Palaniswami, and Peter Kamen. "Do existing measures of Poincare plot geometry reflect nonlinear features of heart rate variability?." *Biomedical Engineering, IEEE Transactions on* 48, no. 11 (2001): 1342-1347.

[8] Brennan, Michael, Marimuthu Palaniswami, and Peter Kamen. "Poincaré plot interpretation using a physiological model of HRV based on a network of oscillators." *American Journal of Physiology-Heart and Circulatory Physiology* 283, no. 5 (2002): H1873-H1886.

[9] Tulppo, MIKKO P., T. H. Makikallio, T. E. S. Takala, T. H. H. V. Seppanen, and H. V. Huikuri. "Quantitative beat-to-beat analysis of heart rate dynamics during exercise." *American Journal of Physiology-Heart and Circulatory Physiology* 40, no. 1 (1996): H244.

[10] Peng, C-K., Shlomo Havlin, H. Eugene Stanley, and Ary L. Goldberger. "Quantification of scaling exponents and crossover phenomena in nonstationary heartbeat time series." *Chaos: An Interdisciplinary Journal of Nonlinear Science* 5, no. 1 (1995): 82-87.

[11] Penzel, Thomas, Jan W. Kantelhardt, Ludger Grote, Jörg-Hermann Peter, and Armin Bunde. "Comparison of detrended fluctuation analysis and spectral analysis for heart rate variability in sleep and sleep apnea." *Biomedical Engineering, IEEE Transactions on* 50, no. 10 (2003): 1143-1151.