CHRIS Study

ECG – HRV

Version 1.0

Link to platform

26th October 2022

Authors: FB, HB

1. Introduction

The ECG exam were performed for all participants at the CHRIS study center in Schlanders Hospital by trained study nurses.

Participants book a morning appointment at the CHRIS study center, ranging from 7.45 to 8.45 a.m. Each study participant is assigned a workflow at the reception. If there are ten study participants (maximum capacity), there are ten different workflows, marked with the letters from "A" to "K". The current workflow is as follows: A-B-C-D-E-F-G-H-I-K. All the workflows can be found in (J:\5-5 New Research Initiatives\5-50 Vinschgau-Study\5-50-2 Logistics & Organisation Schlanders Hospital\Workflows\Workflows_10participants_May_2013.xlsx).

When making the appointment, participants are asked to stay fasting overnight and to drink only water before coming to the study center. They are also requested to avoid wearing jewelry, piercing and alike. The ECG exams always follow the anthropometry assessment (height, weight, body fat). Participants can have breakfast only after their ECG exam is done and blood pressure has been measured and cannot have smoked or exercised in the previous hour.

The ECG exam consisted of two recordings: one of the first 10 seconds, and a longer one lasting 20 minutes, started immediately after the first one.

The preparation of the subject and the electrodes application are crucial for the quality of the recorded ECG. The measurement should be taken in a quiet place, with the participant relaxed, in a lying position. The participant is asked not to move or talk during the measurement, to lie with the back straight, to remove clothes from their torso.

The operators were trained by cardiologists to obtain a clean ECG, so when the participant was not in a relaxed state, or some technical problem occurred, they could decide to rerun the exam. Continuous feedback with the cardiologists who were responsible for interpreting the ECG were also in place, who could ask the operators to repeat a particular exam, to improve the ECG measurement.

HRV Analysis

HRV describes the oscillations over time of both the interval between consecutive heart beats and the instantaneous heart rates (HR).

As the HR is regulated by the Autonomous Nervous System (ANS) acting on the sinoatrial node, HRV is an established (non-invasive) quantitative marker to assess the functioning of cardiac autonomic regulation.

- the sympathetic branch of the ANS increases HR and its response is slow (few secs) → it decreases HRV
- the parasympathetic branch of the ANS decreases HR and its response is faster (0.2-0.6secs) → it increases HRV

In addition to this central control, arterial baroreceptor reflex and respiration are known to induce quick changes in the HR.

There are also some feedback mechanisms modulating the HR that respond to the perturbations sensed by baroreceptors and chemoreceptors.

Under resting condition, parasympathetic tone prevails. However, parasympathetic and sympathetic activity constantly interact, resulting in continuous beat-to-beat interval variations.

HRV time series is calculated after QRS detection: after each QRS complex is detected in a continuous ECG record, the so-called normal-to-normal (NN) intervals – that means, all intervals between adjacent QRS complexes resulting from sinus node depolarizations – or the instantaneous HR series is determined.

The time intervals between two consecutive heartbeats construct the RR series, with the n-th element being computed as follows:

$$RR_n = \alpha \cdot (t_n - t_{n-1})$$

where α is a conversion parameter: if the RR intervals are expressed in ms and the occurrence times are expressed in seconds, then $\alpha = 1000$.

If the HRV series is constructed as the sequence of instantaneous HR, then the n-th element is computed as:

$$HR_n = \frac{\beta}{t_n - t_{n-1}}$$

where β is also used as a conversion parameter: if the HR is expressed in bpm (beats per minute) and occurrence times are expressed in seconds, then $\beta = 60$.

The resulting RR series thus consists of a set of pairs (t_n, RR_n). It should be noted that this time series is not equidistantly sampled. This must be taken into consideration before performing frequency-domain analysis, which requires a uniformly sampled time series. There are several approaches to overcome this issue:

- Interpolation: the non-uniformly sampled RR series is transformed into an equidistant sampled one
- Tachogram: equidistant sampling is assumed by constructing tachogram, with RR intervals plotted as a function of the beat number (when using this approach, the spectrum is not a function of the frequency, rather of cycles per beat)
- Spectrum of the counts: it uses a series of impulses (delta functions) positioned at beat occurrence times. This approach relies on the Integral Pulse Frequency Modulator (IPFM) model, which simulates the modulation of the sinoatrial node.

Before performing the analysis of such constructed RR time series, a filtering operation must be performed to eliminate outliers or spurious points in the signal. Outliers originate from the detection of an artifact as a heartbeat (RR interval too short), or from the loss of a heartbeat in the detection procedure (RR interval too large). The RR time series may also contain some physiological artifacts such as ectopic beats (when the heartbeat is not triggered by the sinoatrial node, causing an extra beat).

If detection of the heartbeat has been revised and corrected manually by a physician, this step can be skipped.

Several tools perform HRV analysis, which is useful to extract physiological information, and they can be classified into three main categories: time-domain methods, frequency-domain methods and non-linear methods.

Kubios HRV Premium

Kubios HRV is a gold-standard heart rate variability (HRV) analysis software designed for research and professional use. The software is suitable for clinical and public health researchers, professionals working on human well-being, or sports enthusiasts; for anybody who want to perform detailed analyses on heart rate variability, e.g., to examine autonomic nervous system function.

The first versions of the Kubios HRV were developed as part of academic research work carried out at the Department of Applied Physics, University of Eastern Finland, Kuopio, Finland.

Kubios HRV is a scientifically validated software and the most commonly used HRV analysis software for scientific research. Kubios HRV is currently available as two alternative products:

- Kubios HRV Standard: Freeware HRV analysis software for non-commercial personal use. Supports HR data from most common HR monitor manufacturers and computes most commonly used time- and frequency domain HRV parameters. Software is operated through an easy-to-use GUI and analysis results can be saved as PDF report or text file.
- Kubios HRV Premium: Full featured HRV analysis software designed for scientific research and professional use. Supports wide range of ECG, PPG and HR data and computes all commonly used time-domain, frequency-domain and nonlinear HRV parameters. In addition, Kubios HRV Premium includes improved preprocessing (beat correction and noise handling), ECG derived respiration, time-varying analysis, and extended exporting options. Analysis results can be saved as illustrative PDF reports, CSV text file, MATLAB MAT file and also in a “SPSS friendly” batch file.

For a more detailed description of the features, please refer to Kubios Users’ Guide.

Analysis of 20m ECGs with Kubios HRV Premium

Kubios HRV Premium software version in use for analysis: 3.5.0

The analysis of the ECG signals is pretty straightforward: given the raw data (CHRIS ECGs are originally stored as a binary format) they first need to be converted into a comma-separated-value format, in order to let Kubios read them.

One channel out of the 12 recorded must be selected, always making sure that the correct sampling frequency is specified, and the AID of the patient has to be inserted in the Patient Name field (just for later sanity checks).

At this point the signal is read and analyzed.

Once the analysis is done and the results presented, these results can be saved either as a MATLAB, pdf or csv file (or any combination of those).

All Kubios settings are left as their default values.

In first analysis, just one channel was selected. The choice of the channel was on V5 (see following paragraph for the rationale behind the choice of V5). Then, all the other channels were also analyzed.

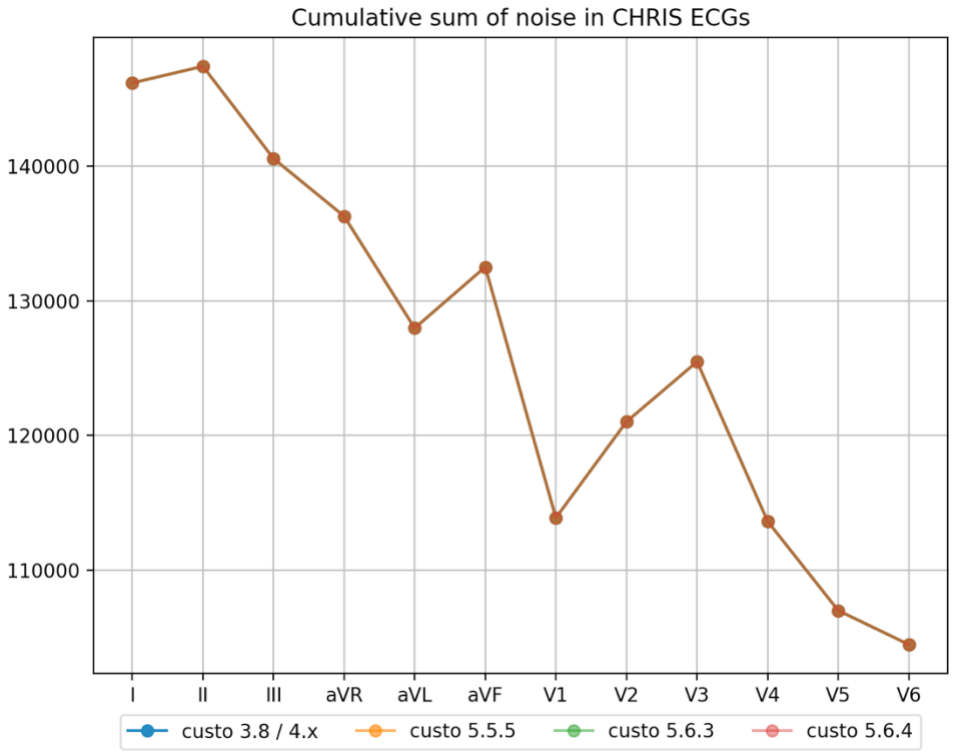
Unfortunately, Kubios is not designed for batch analysis, and this issue was solved by implementing a script for emulating mouse clicks.

Settings

All the settings in Kubios were left with their default values. See [2] for further details.

Choice of the lead for HRV analysis (V5)

To choose the best possible lead for HRV analysis, the noise level estimation as reported by Customed was taken into account. The noise value for each lead was summed over the 13 thousand exams that were available (from CHRIS participants), and the result was that V5, together with V6, were the least noisy leads, with much lower noise level compared to the other leads.



Notes:

The automation script for the HRV analysis on V5 is a bit different from what was used for all the other leads.

In V5, mouse clicks were emulated. In all the other channels, keyboard shortcuts were instead used (which is a little more stable and allowed to speed up the analysis – in terms of delay time between one click / press and the subsequent one).

Apart from this, the major difference from the two script regards the selection of the entire length of the RR series used to compute all the HRV parameters:

- For V5, the entire length is selected for each exam
- In all the other channels, this selection was made unnecessary by setting the preference in Kubios accordingly.

2. History version changes

The current module can be considered as a single version of the measurements.

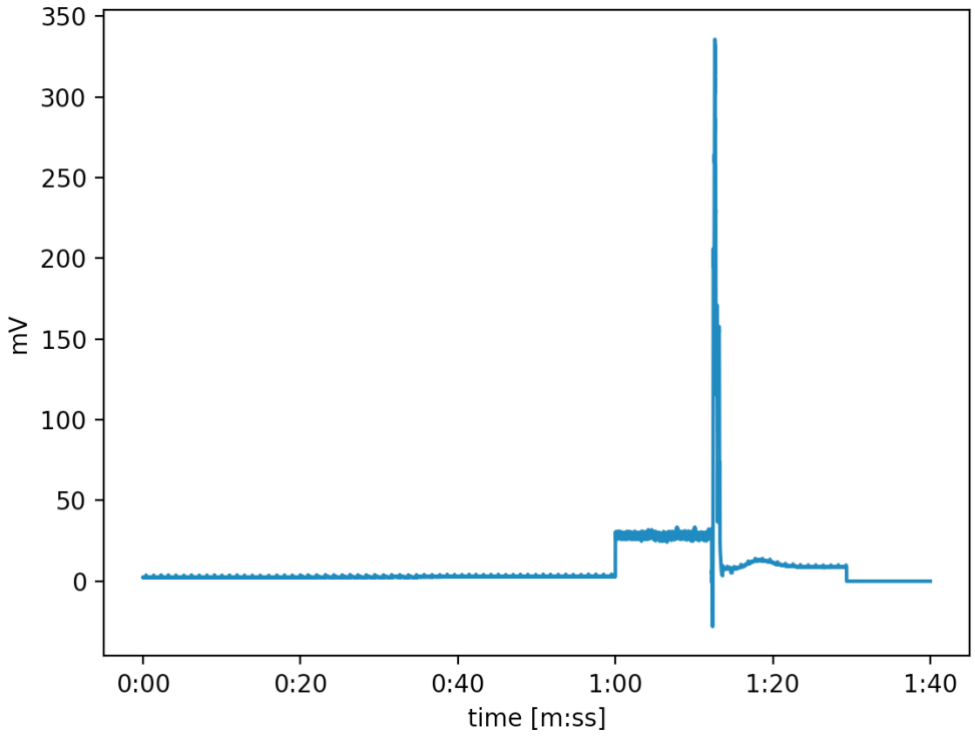
3. Data Cleaning

No actual data cleaning performed. All the ECG exams were fed into Kubios software, which automatically dealt with noisy data. When data quality was too low, or the signal was too short (minimum sample length is 60 seconds) the analysis was empty (the MATLAB file was generated with all the fields left empty). Some other cases, Kubios was stuck in opening couple of exams.

Those cases were fixed manually for V5 only.

- in one case, the filtered signal instead of the raw data was fed into Kubios software for the analysis.

- in one case, no other action was performed. Kubios remained stuck in the QRS complex detection process, prior any type of analysis could be performed. Looking at the raw data, a huge artefact is present in the early minutes of the recording:



This could explain why the analysis could not be performed.

Other cases were checked manually since the following problem emerged: as mentioned before, the total length of the ECG recording was selected for the analysis. To check that this was done correctly, the data length and the end time point of the analyzed segment were compared, expecting them to be equal, unless the final portion of the signal was flagged as noisy. There were cases however, where, beside analyzing the same signal multiple times, a certain gap persisted between the two measures, which could not be explained by the fact that the final portion of recording was noisy. Inspecting those cases, it was found that the raw data were all zero from a certain point on, confirming that Kubios was actually analyzing all the useful signal present in there. This problem affects roughly 70 cases.

4. Data structure

The variables here listed constitute the variables computed by Kubios software and exported from the MATLAB file.

For the definition of the pathologic limits, Kubios refers to the current scientific literature.

4.1 Metadata regarding the ECG recording, RR series, Noise segments

ECG metadata

These measurements indicate the original length of the ECG raw data (x0ec202), the effective data length that was used to compute the HRV measures, expressed in absolute value (x0ec203) and as percentage of the total length of the ECG (x0ec204); the start and the end point in time of the analyzed portion of ECG is also indicated (x0ec205a, x0ec205b).

RR series metadata

Those measures indicate the length of the RR series (x0ec206), the start and the end point in time of the series (x0ec207a, x0ec207b), the total number of beats (x0ec208) and the number of corrected beats, both in absolute value (x0ec209) and expressed as percentage (x0ec210).

Noise metadata

Kubios also features a noise detection algorithm that is able to localize noisy regions in the ECG trace. x0ec211 expresses the number of noisy regions that Kubios detected, while x0ec212 reports the start and the end time point of each localized region.

4.2 PNS & SNS Summary

Parasympathetic cardiac activity is known to 1) decrease heart rate, 2) increase HRV via enhanced respiratory sinus arrhythmia (RSA) component, and 3) decrease the ratio between lower frequency and higher frequency oscillations in HRV time.

Based on the above, the parasympathetic nervous system index (PNS index) is computed in Kubios HRV software based on the following three parameters:

- **Mean RR interval**. Longer mean RR interval means lower heart rate and higher parasympathetic cardiac activation.
- **Root mean square of successive RR interval differences (RMSSD),** which is a commonly used time-domain HRV parameter that captures the quick beat-to-beat changes in RR interval, and

therefore, strongly linked to RSA component magnitude. High values of RMSSD indicate strong RSA component and high parasympathetic cardiac activation.

- **Poincaré plot index SD1 in normalized units**. A commonly used approach for estimating the sympathovagal balance of the ANS is to compute the low frequency (LF) to high frequency (HF) power ration from HRV spectrum. However, in case of spontaneous breathing, the RSA component is partially or even completely overlapping with the LF component. In such cases the LF/HF ratio gives invalid interpretation of ANS status. Since Poincare plot index SD1 is known to be linked to RMSSD, the normalized SD1 value is used in Kubios HRV as the third input parameter for the PNS index computation.

Each parameter value is first compared to their normal population values as presented in Nunan et al. 2010. The normal value for the SD1 is derived based on its dependency on the time-domain variable RMSSD as described in Brennan et al. 2001. The parameter values are then scaled with the standard deviations of normal population and finally a proprietary weighting is applied to obtain robust and reliable PNS index value.

Sympathetic cardiac activity is known to 1) increase heart rate, 2) decrease HRV, reducing especially quick RSA related changes in RR interval, and 3) increase the ratio between lower frequency and higher frequency oscillations in HRV data.

Based on the above, the sympathetic nervous system index (SNS index) is computed in Kubios HRV software based on the following three parameters:

- **Mean HR interval.** Higher heart rate is linked to higher sympathetic cardiac activation.
- **Baevsky's stress index (SI)**, which is a geometric measure of HRV reflecting cardiovascular system stress. High values of SI indicate reduced variability and high sympathetic cardiac activation.
- **Poincaré plot index SD2 in normalized units**. As mentioned above (in PNS index description) LF/HF power ratio is commonly used assessing sympathovagal balance of the ANS, which however is sensitive to breathing rate. Thus normalized Poincare plot index SD2, which is known to be linked to SDNN and to correlate with LF/HF ratio, is used in Kubios HRV as the third input parameter for the SNS index computation.

Each parameter value is first compared to their normal population values as presented in Nunan et al. 2010. The normal value for the SD2 is derived based on its dependency on the time-domain variable SDNN as described in Brennan et al. 2001. The normal values for the Baevsky's stress index are taken from Baevsky 2009. The parameter values are then scaled with the standard deviations of normal population and finally a proprietary weighting is applied to obtain the SNS index value.

4.3 Time Domain Analysis

The time domain measurements can be derived into two classes: statistical methods and geometrical methods.

Among statistical methods, these parameters are derived either from the RR interval time series, composed by N successive beat intervals, such as:

$$RR = (RR_1, RR_2, \dots, RR_N,)$$

or from the series of successive RR interval differences, with the n-th element being computed as:

$$\Delta RR_n = RR_{n+1} - RR_n$$

Mean RR interval (x0ec215)

$$\overline{RR} = \frac{1}{N}\sum_{n=1}^{N} RR_n$$

Standard Deviation of normal-to-normal intervals (SDNN) (x0ec216)

The best-known time analysis statistic is the standard deviation of the RR interval, which is computed as:

$$SDNN = \sqrt{\frac{1}{N-1}\sum_{n=1}^{N}(RR_n - \overline{RR})^2}$$

Since the variance is mathematically equal to the total power of the spectral analysis, SDNN reflects the power of all the cyclic components responsible for variability in the period of recording (both short- and long-term variations within the RR series). When it is calculated over a 24h period, the measure encompasses both the short-term high frequency variations as well as the lowest frequency components seen in a 24h. as the period of monitoring decreases, SDNN estimates shorter and shorter cycle lengths.

It should also be noted that the total variance of the HRV increases with the length of the analyzed recording. Thus, on arbitrarily selected ECGs, SDNN may not be an appropriate statistical quantity because it depends on the recording's length.

Square Root of the Mean Squared Differences of successive NN intervals (RMSSD) (x0ec217)

The root mean square of successive differences is given by:

$$RMSSD = \sqrt{\frac{1}{N-1}\sum_{n=1}^{N}(\Delta RR_n)^2}$$

Number of successive NN intervals greater than 50 ms (pNN50) (x0ec218)

It is defined as the number of successive RR (or NN) intervals that differ more than 50 ms.

$$NN50 = \#\,\Delta RR > 50ms$$

Percentage of interval differences of successive NN intervals greater than 50 ms (pNN50) (x0ec219)

Proportion of the number of interval differences of successive RR intervals greater than 50ms (NN50) with respect to the total number of RR intervals.

$$pNN50 = \frac{\#\,\Delta RR > 50ms}{N} \times 100\%$$

All these measurements of short-term variation estimate HF variations in HR and thus are highly correlated.

Standard Deviation of the average NN interval in all 5 min. of the whole record (SDANN) (x0ec220)

To avoid the issue of SDNN being length-dependent, statistical variables calculated from segments of the total monitoring period may be used. Among these, the SDANN is defined as the standard deviation of the average NN (RR) intervals calculated over short periods (usually 5 minutes), which is an estimate of the changes in HR due to cycles longer than 5 min.

SDNN Index (SDNNI) (x0ec221)

It is of the 5-min standard deviation of the NN interval calculated over 24h, which measures the variability due to cycles shorter than 5 min.

Mean HR (x0ec222)

$$\overline{HR} = \frac{60}{\overline{RR}}$$

Standard Deviation of HR (x0ec223)

$$SDT\_HR = \sqrt{\frac{1}{N-1}\sum_{n=1}^{N}(HR_n - \overline{HR})^2}$$

Min and Max HR (x0ec224- x0ec225)

$$HR_{min} = min(HR_1, HR_2, \dots, HR_N,)$$
$$HR_{max} = max(HR_1, HR_2, \dots, HR_N,)$$

(In Kubios they are computed using N beat moving average (default value: N=5))

Baevsky's stress index (SI) (x0ec226)

Stress Index, Triangular index and TINN are classified as geometric measures, and they are calculated from the RR interval histogram.

The Baevsky's stress index (SI) is computed according to the formula:

$$\mathrm{SI} = \frac{\mathrm{AMo} \times 100\%}{2\mathrm{Mo} \times \mathrm{MxDMn}}$$

where AMo is the so-called mode amplitude presented in percent, Mo is the mode (the most frequent RR interval) and MxDMn is the variation scope reflecting degree of RR interval variability. The mode Mo

is simply taken as the median of the RR intervals. The AMo is obtained as the height of the normalised RR interval histogram (bin width 50 msec) and MxDMn as the difference between longest and shortest RR interval values. In order to make SI less sensitive to slow changes in mean heart rate (which would increase the MxDMn and lower AMo), the very low frequency trend is removed from the RR interval time series by using the smoothness priors method. In addition, the square root of SI is taken to transform the tailed distribution of SI values towards normal distribution.

HRVI – HRV Triangular Index (x0ec227)

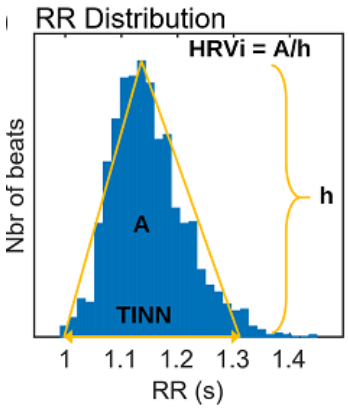
The HRV triangular index is defined as the integral of the density distribution (i.e., the number of RR intervals) divided by the maximum of the density distribution. The density distribution can be estimated using a histogram; thus, the size of the bins must be specified. Using this measurement of RR intervals on a discrete scale, this measure is approximated by the value:

$$HRVI = {Tot.number\ of\ RR\ intervals}/{Number\ of\ RR\ intervals\ in\ the\ modal\ bin}$$

As the measure is dependent on the bin length, a bin width of 1/128 s is recommended in order to obtain comparable results

TINN – Triangular Interpolation of NN (RR) interval histogram (x0ec228)

It is defined as the baseline width of the distribution measured as the base of a triangle, approximating the RR interval distribution (a triangular interpolation of the histogram may be used, and the minimum square difference is used to find such a triangle). It is usually expressed in ms.



Both HRVI and TINN express overall HRV usually measured over 24h and are more influenced by the lower than by the higher frequencies.

Deceleration Capacity (DC) (x0ec230)

Overall measures of HRV such as SDNN do not distinguish between vagal and sympathetic effects. An approximate distinction of these two effects might be made possible by separate assessment of the deceleration related and acceleration related HRV.

DC of HRV is a measure of the parasympathetic modulation on the heart, as it captures the lengthening of RR interval within 2-4 successive beats.

To compute this index, heartbeat intervals longer than the preceding one are identified as anchors. To suppress errors due to artifacts, RR interval prolongations of more than 5% are excluded.

Segments of interval data around the anchors are selected. All segments have the same size (chosen according to the lowest frequency to be visualized). They are aligned at the anchors (phase rectification).

The signal X(i) is obtained by ensemble averaging the signals within the aligned segments - i.e., phase rectified signal averaging (PRSA) is applied.

In this new signal, X(0) is the average of the RR intervals at all anchors; X(1) and X(-1) are the averages of the RR intervals immediately following and preceding the anchors.

DC is then quantified as follows:

$$DC = \frac{[X(0) + X(1) - X(-1) - X(-2)]}{2}$$

The obtained curve characterizes the average capacity of the heart to decelerate the cardiac rhythm from one beat to the next.

NB. In this formula, DC is divided by 2, instead of 4 as in the original paper, in order to reflect the absolute change in RR interval (in msec).

Acceleration Capacity (AC) – variable x0ec229 in the Query Builer

Opposite to the DC index, AC captures the shortening of RR interval within few successive beats.

To compute AC, heartbeat intervals shorter than the preceding one are identified as anchors. To suppress errors due to artifacts, RR interval shortening of more than 5% are excluded. Segments of interval data around the anchors are selected and aligned at the anchors (phase rectification). The PRSA signal X(i) is obtained by ensemble averaging the signals within the aligned segments.

AC is then quantified as follows:

$$AC = \frac{[X(0) + X(1) - X(-1) - X(-2)]}{2}$$

NB. In this formula, DC is divided by 2, instead of 4 as in the original paper, in order to reflect the absolute change in RR interval (in msec).

Acceleration and Deceleration Capacity modified indexes (x0ec231, x0ec232)

To better capture the instantaneous deceleration and acceleration, a modified DC and AC parameters are also computed according to (Nasario-Junior et al. 2014) as follows:

$$AC_{mod} = [RR(0) - RR(-1)]$$

$$DC_{mod} = [RR(0) - RR(-1)]$$

4.4 Frequency Domain Analysis

Due to the different speed of response of the sympathetic and parasympathetic branches on the HR, the frequency analysis can be useful to split these two contributions and analyze them singularly. Three frequency bands can be identified, each one reflecting one or many components of the ANS:

- VLFs (below 0.04Hz) provide information on the renin-angiotensin system
- LFs (0.04-0.15Hz) provide information on both sympathetic and parasympathetic system
- HFs (0.15-0.4Hz) provide information on the parasympathetic system

The most-used frequency domain analysis technique is the Power Spectral Density (PSD), which provides the basic information of how power (i.e., variance of RR intervals) distributes as a function of frequency. Independent of the method employed, only an estimate of the true PS of the signals can be obtained.

Methods for the calculation of the PSD may be classified as non-parametric and parametric. In most instances, they provide comparable results.

Advantages of non-parametric:

- Simplicity of the algorithm employed (FFT)
- High processing speed

Advantages of parametric:

- Smoother spectral components which can be distinguished independently of preselected frequency bands
- Easy post-processing of the spectrum with automatic calculation of LF and HF power and easy identification of the central frequency of each component
- Accurate estimate of the PSD even on small number of samples on which the signal is supposed to maintain stationarity

Disadvantages of parametric:

- Need to verify the suitability of the chosen model and its complexity (i.e., order of the model)

In Kubios HRV software, the HRV spectrum is calculated with FFT based Welch's periodogram method and with the AR method. Spectrum factorization in AR method is optional. In the Welch's periodogram

method the HRV sample is divided into overlapping segments. The spectrum is then obtained by averaging the spectra of these segments, which decreases the variance of the FFT spectrum.

The generalized frequency bands in case of short-term HRV recordings are the very low frequency (VLF, 0–0.04 Hz), low frequency (LF, 0.04–0.15 Hz), and high frequency (HF, 0.15–0.4 Hz). The frequency-domain measures extracted from a spectrum estimate for each frequency band include absolute and relative powers of VLF, LF and HF bands; LF and HF band powers in normalized units; the LF/HF power ratio; and peak frequencies for each band. In the case of FFT spectrum, absolute power values for each frequency band are obtained by simply integrating the spectrum over the band limits. In the case of AR spectrum, on the other hand, if factorization is enabled distinct spectral components emerge for each frequency band with a proper selection of the model order and the absolute power values are obtained directly as the powers of these components. If factorization is disabled, the AR spectrum powers are calculated as for the FFT spectrum. The band powers in relative and normalized units are obtained from the absolute values.

Total Power (TP) (x0ec239, x0ec243, x0ec256, x0ec260)

TP expresses the variance of the total number of NN intervals. It is calculated by integrating the PSD within the overall bandwidth (between 0 and 0.4Hz).

It is measured both in ms^2 and with natural logarithm transformation and reflects overall autonomic activity.

Very Low Frequency (VLF) (x0ec233, x0ec236, x0ec240, x0ec250, x0ec253, x0ec257)

The physiological explanation of the VLF components is much less defined than LF and HF. The non-harmonic component which does not have coherent properties, and which is affected by algorithms of baseline or trend removal is commonly accepted as a major constituent of VLF.

VLF band is also related to the renin-angiotensin system and is affected by algorithms of baseline removal.

Low Frequency (LF) (x0ec234, x0ec237, x0ec241, x0ec251, x0ec254, x0ec258)

The LF interpretation is a subject of controversy. Some consider the LF phenomena being of both sympathetic and parasympathetic origin, others suggest that the sympathetic system predominates.

Generally, it is a strong indicator of sympathetic activity, whereas parasympathetic influence is represented by LF when respiration rate is lower than 7 breaths per minute or during taking a deep breath. Thus, when subject is in the state of relaxation with a slow and even breathing, the LF values can be very high indicating increased parasympathetic activity rather than increase of sympathetic regulation.

This band also includes the component referred to as the 10-second rhythm or the Mayer wave, caused by oscillations in baroreceptor and chemoreceptor reflex control systems.

High Frequency (HF) (x0ec235, x0ec238, x0ec242, x0ec252, x0ec255, x0ec259)

HF bands provides information mostly on the parasympathetic branch. Moreover, among all the HF mechanisms involved in the HR modulation, we also find the so-called Respiratory Sinus Arrhythmia (RSA), indicator of the heartbeat synchronization with the respiratory rhythm.

Relative power [%] of VLF, LF and HF bands (x0ec244, x0ec245, x0ec246, x0ec261, x0ec262, x0ec263)

VLF [%] = VLF [ms^2] / total power [ms^2] x 100%
LF [%] = LF [ms^2] / total power [ms^2] x 100%
HF [%] = HF [ms^2] / total power [ms^2] x 100%

Powers of LF and HF bands in normalised units [n.u] (x0ec247, x0ec248, x0ec264, x0ec265):

Normalized Low Frequency (LF Norm) - Normalized Low Frequency is the ratio between absolute value of the Low Frequency and difference between Total Power and Very Low Frequency. This measure minimizes an effect of changes in Very Low Frequency power and emphasizes changes in sympathetic regulation. Normalized LF is calculated in percentile units.

$$LF\ [n.u.] = LF\ [ms^2] / (total\ power\ [ms^2] - VLF\ [ms^2]) \times 100\%$$

Normalized High Frequency (HF Norm) - Normalized High Frequency is the ratio between absolute value of the High Frequency and difference between Total Power and Very Low Frequency. This measure minimizes an effect of changes in Very Low Frequency power and emphasizes changes in parasympathetic regulation. Normalized HF is calculated in percentile units.

$$HF\ [n.u.] = HF\ [ms^2] / (total\ power\ [ms^2] - VLF\ [ms^2]) \times 100\%$$

LF/HF Ratio (x0ec249, x0ec266)

The LF/HF ratio is the ratio between the power of LF and HF bands. It indicates the overall balance between sympathetic and parasympathetic systems (sympatho/vagal balance).

EDR: ECG derived respiration (x0ec267)

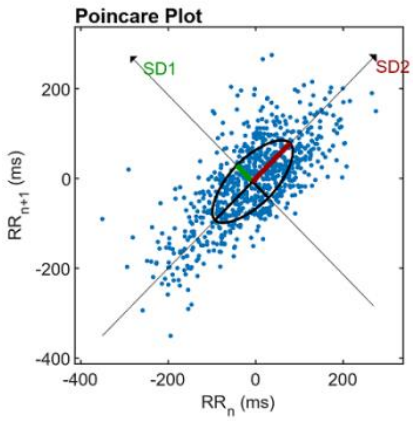
Expressed in Hz.

4.5 Nonlinear Domain Analysis

This type of analysis is useful to extract valuable information from the HRV since it responds to a complex control system. Nonlinear phenomena are certainly involved in the genesis of HRV. It has been speculated that analysis of the HRV based on nonlinear methods may elicit valuable information for the physiological interpretation of HRV and for the assessment of the risk of sudden cardiac death.

Poincaré plot ()

It is a graphical representation of the correlation between successive RR intervals, i.e. plot of RRn+1 as a function of RRn, thus creating a scatter plot. The shape of the plot is essential and a common approach to parameterize the shape is to fit an ellipse to the plot.



This ellipse is oriented according to the line of identity (RRn=RRn+1).

After fitting the ellipse, we can derive three nonlinear measures. The first one is the area of the ellipse (S), which represents total HRV, correlates with baroreflex sensitivity (BRS), LF and HF power, and RMSSD.

The other two measures are the standard deviation SD1 and SD2, described below.

Their ratio SD1/SD2 measures the unpredictability of the RR time series, and it is used to measure the autonomic balance when the monitoring period is sufficiently long and there is sympathetic activation.

SD1/SD2 is correlated with the LF/HF ratio.

Standard deviation (longitudinal or diagonally of the angle bisector) in Poincare plot (x0ec268)

The standard deviation of the points perpendicular to the line of identity, denoted by SD1, describes the short-term variability which is mainly caused by the RSA.

SD1 is related to the time-domain measure SDSD according to the equation:

$$SD1^2 = \frac{1}{2}SDSD^2$$

Where SDSD is the standard deviation of successive RR interval differences, defined by

$$SDSD = \sqrt{E\{\Delta RR_n^2\} - E\{\Delta RR_n\}^2}$$

Standard deviation (diagonally or longitudinal of the angle bisector) in Poincare plot (x0ec269)

The standard deviation of the points along to the line of identity, denoted by SD2, on the other hand, describes the long-term variability.

SD2 is related to the time-domain measures SDNN and SDSD according to the equation:

$$SD2^2 = 2 * SDNN^2 - \frac{1}{2}SDSD^2$$

Ratio SD2-SD1 (x0ec270)

Simply given by the ratio between the two above defined indexes: SD2 / SD1

EnoughData (x0ec271)

It is a Boolean flag to indicate whether the nonlinear analysis (with all the following indexes) could be performed or not.

Approximate Entropy (x0ec272)

ApEn measures the irregularity or complexity of a time series. It was designed for short time series in which some noise is present and makes no assumptions regarding underlying systems dynamics. Applied to HRV, large ApEn values indicate low predictability of fluctuations in successive RR intervals. Small ApEn values mean that the signal is more regular and predictable.

The ApEn is computed as follows:

- First, a set of length m vector *uj* is formed

$$u_j = (\mathrm{RR}_j, \mathrm{RR}_{j+1}, \ldots, \mathrm{RR}_{j+m-1}), \quad j = 1, 2, \ldots N - m + 1$$

  where *m* is called the embedding dimension and *N* is the number of measured RR intervals. The distance between these vectors is defined as the maximum absolute difference between the corresponding elements, i.e.

$$d(u_j, u_k) = \max\left\{|\mathrm{RR}_{j+l} - \mathrm{RR}_{k+l}| \,\middle|\, l = 0, \ldots, m-1\right\}$$

- Next, for each *uj* the relative number of vectors uk for which d(uj, uk) <=r is calculated. This index can be written in the form

$$C_j^m(r) = \frac{\text{nbr of } \left\{u_k \,\middle|\, d(u_j, u_k) \le r\right\}}{N - m + 1} \quad \forall\, k.$$

  Due to the normalization, this value is always smaller or equal to 1. Note that this value is at least 1/(N-m+1) since uj is also included in the count.

- Then, take the natural logarithm of each Cj^m(r) and average over j to yield:

$$\Phi^m(r) = \frac{1}{N - m + 1} \sum_{j=1}^{N-m+1} \ln C_j^m(r).$$

- Finally, the approximate entropy is obtained as:

$$\mathrm{ApEn}(m, r, N) = \Phi^m(r) - \Phi^{m+1}(r).$$

Thus, the value of the estimate ApEn depends on three parameters, the length m of the vectors uj, the tolerance value r, and the data length N.

When N is increased the ApEn approaches its asymptotic value. The tolerance r has a strong effect on ApEn and it should be selected as a fraction of the standard deviation of the RR interval data (SDNN).

This selection enables the comparison of RR data from different subjects. A common selection for r is r = 0.2SDNN.

Sample Entropy (x0ec273)

SampEn is similar to ApEn, i.e., it measures the complexity of a time series. Large values indicate high complexity, smaller values characterize more regular signals. There are two important differences with respect to the calculation of ApEn: for ApEn, in the calculation of the number of vectors uk for which d(uj, uk) <= r, also the vector uj itself is included. This ensures that Cj^m(r) is always larger than 0 and the log can be applied, but at the same time it causes bias to ApEn. In SampEn instead, this self-comparison of uj is eliminated by calculating Cj^m(r) as

$$C_j^m(r) = \frac{\text{nbr of } \left\{u_k \,\middle|\, d(u_j, u_k) \le r\right\}}{N - m} \quad \forall\, k \neq j.$$

In this way Cj^m(r) is between 0 and 1. Next, the values of Cj^m(r) are averaged to yield

$$C^m(r) = \frac{1}{N - m + 1} \sum_{j=1}^{N-m+1} C_j^m(r)$$

And sample entropy is obtained as

$$\mathrm{SampEn}(m, r, N) = \ln\left(C^m(r)/C^{m+1}(r)\right).$$

Both ApEn and SampEn are estimates for the negative natural logarithm of the conditional probability that a data of length N, having repeated itself within a tolerance r for m points, will also repeat itself for m+1 points. SampEn was designed to reduce the bias of ApEn and has a closer agreement with the theory for data with known probabilistic content.

D2 – correlation dimension (x0ec274)

D2 estimates the minimum number of dynamic variables required to model the underlying system; the more variables required to predict the time series, the greater its complexity. The D2 index measures

the system's attractor dimension (the set of values toward which a variable in the system converges over time), which can be integer or fractal.

It can be obtained as follows:

- Similar to the calculation of ApEn and SampEn, form length m vectors uj as

$$u_j = (\mathrm{RR}_j, \mathrm{RR}_{j+1}, \ldots, \mathrm{RR}_{j+m-1}), \quad j = 1, 2, \ldots, N-m+1$$

- Calculate the number of vectors uk for which d(uj, uk)<=r, that is

$$C_j^m(r) = \frac{\text{nbr of } \left\{u_k \,\middle|\, d(u_j, u_k) \le r\right\}}{N-m+1} \quad \forall\, k$$

- Where the distance function d(uj, uk) is now defined as

$$d(u_j, u_k) = \sqrt{\sum_{l=1}^{m} (u_j(l) - u_k(l))^2}.$$

- An average of the term Cj^m(r) is taken as:

$$C^m(r) = \frac{1}{N-m+1} \sum_{j=1}^{N-m+1} C_j^m(r)$$

  That is the so-called correlation integral.

- The correlation dimension D2 is defined as the limit value

$$D_2(m) = \lim_{r \to 0} \lim_{N \to \infty} \frac{\log C^m(r)}{\log r}.$$

This limit value is approximated by the slope of the regression curve (log r, log Cm(r)). The slope is calculated from the linear part of the log-log plot.

The slope of the regression curves tends to saturate on the finite value of D2 when m is increased. A common choice for m is m=10.

Short term and long term scaling exponent of detrended fluctuation analysis (DFA alpha1 and alpha2) (x0ec275, x0ec276)

DFA extracts the correlations between successive RR intervals over different time scales.

The correlation is extracted for different time scales as follows:

- The RR interval time series is first integrated

$$y(k) = \sum_{j=1}^{k} (\mathrm{RR}_j - \overline{\mathrm{RR}}), \quad k = 1, \ldots, N$$

  where $\overline{RR}$ is the average RR interval.

- Next, the integrated series is divided into segments of equal length n. within each segment, a least squares line is fitted into the data. Let yn(k) denote these regression lines. Next the integrated series y(k) is detrended by subtracting the local trend within each segment and the root-mean-square fluctuation of this integrated and detrended time series is calculated by:

$$F(n) = \sqrt{\frac{1}{N}\sum_{k=1}^{N}(y(k) - y_n(k))^2}.$$

- This computation is repeated over different segment lengths to yield the index F(n) as a function of segment length n. Typically F(n) increases with segment length. A linear relationship on a double log graph indicates presence of fractal scaling and the fluctuations can be characterized by scaling exponent α (the slope of the regression line relating log F(n) to log n. Different values of α indicate the following
    - $\alpha = 1.5$: Brown noise (integral of white noise)
    - $1 < \alpha < 1.5$: Different kinds if noise
    - $\alpha = 1$: 1/f noise
    - $0.5 < \alpha < 1$: Large values are likely to be followed by large value and vice versa
    - $\alpha = 0.5$: White noise
    - $0 < \alpha < 0.5$: Large values are likely to be followed by small value and vice versa

Typically, in DFA the correlations are divided into short-term and long-term fluctuations. This analysis results in slope α1, obtained from the (lon n, log F(n)) graph within range $4 \leq n \leq 12$, which describes short-term fluctuations, and slope α2, obtained from the range $13 \leq n \leq 64$ which describes long-term fluctuations.

The short-term correlations extracted using DFA reflect the baroreceptor reflex, while long-term correlations reflect the regulatory mechanisms that limit fluctuation of the beat cycle. DFA is designed to analyze a time series that spans several hours of data.

Recurrence Plot Analysis (RPA) (x0ec277, x0ec278, x0ec279, x0ec280, x0ec281, x0ec282)

Advanced technique for the nonlinear analysis that quantifies the number and duration of the recurrences in the phase space. A recurrence is a time instant in which the trajectory returns to a phase space region it has visited before.

In this approach, vectors

$$u_j = (\mathrm{RR}_j, \mathrm{RR}_{j+\tau}, \ldots, \mathrm{RR}_{j+(m-1)\tau}), \quad j = 1, 2, \ldots, N - (m-1)\tau$$

Where m is the embedding dimension and τ the embedding lag. The vectors uj then represent the RR interval time series as a trajectory in m dimensional space. A recurrence plot is a symmetrical $[N - (m-1)\tau] \times [N - (m-1)\tau]$ matrix of zeros and ones. The element in the j-th row and k-th column of the RP matrix, i.e. RP(j,k), is 1 if the point uj on the trajectory is close to point uk. That is

$$\mathrm{RP}(j,k) = \begin{cases} 1, & d(u_j - u_k) \leq r \\ 0, & \text{otherwise} \end{cases}$$

where d(uj,uk) is the Euclidean distance and r is a fixed threshold. The structure of the RP matrix usually shows short line segments of ones parallel to the main diagonal. The lengths of these diagonal lines describe the duration of which the two points are close to each other.

The recurrence plot is therefore the graphical representation of the recurrence matrix of the RR time series.

Default values for these indexes can be m=10, $\tau = 1$, $r = \sqrt{m} * SD$

Several statistics can be derived from the RQA analysis of the RR time series.

The most important RQA statistics are:

- Recurrence rate (REC): Percentage of recurrence points in a Recurrence Plot (i.e., ratio of ones and zeros in the RP matrix). Being the number of elements in the RP matrix for $\tau = 1$ equal to $N - m + 1$, the recurrence rate is given by

$$REC = \frac{1}{(N-m+1)^2} \sum_{j,k=1}^{N-m+1} RP(j,k)$$

- Laminarity (LAM): Percentage of recurrence points that form vertical lines
- Ratio (RATIO): Ratio between DET and RR
- Longest diagonal line (Lmax): Length of the longest diagonal line
- Averaged diagonal line length (Lmean): Mean length of the diagonal lines. The main diagonal is not taken into account. It is obtained as:

$$l_{mean} = \frac{\sum_{l=lmin}^{lmax} lN_l}{\sum_{l=lmin}^{lmax} N_l}$$

  Where N_l is the number of length l lines.
- Divergence (DIV): Inverse of Lmax

$$DIV = \frac{1}{Lmax}$$

  This has been shown to correlate with the largest positive Lyapunov exponent.
- Determinism (DET): Percentage of recurrence points that form diagonal lines

$$DET = \frac{\sum_{l=lmin}^{lmax} lN_l}{\sum_{j,k=1}^{N-m+1} RP(j,k)}$$

- Longest vertical line (Vmax): Longest vertical line
- Trapping time (Vmean): Average length of the vertical lines
- Shannon Information Entropy (ShanEn): Shannon entropy of the diagonal line lengths distribution

$$ShanEn = -\sum_{l=lmin}^{lmax} n_l \ln n_l$$

  Where n_l is the number of length l lines divided by the total number of lines, that is:

$$n_l = \frac{N_l}{\sum_{l'=lmin}^{lmax} N_{l'}}$$

- Recurrence Rate (REC): Number of recurrent points depending on the distance to the main diagonal

Multiscale Entropy (MSE) (x0ec283a…x0ec283t)

MSE is an extension of SampEn in the sense that it incorporates two procedures:

1. A coarse-graining process is applied to the RR interval time series. Multiple coarse-grained time series are constructed for the time series by averaging the data points within non-overlapping windows of increasing length τ, where τ represents the scale factor and is selected to range between $\tau = 1, 2, \dots, 20$. The length of each coarse-grained time series is N/τ , where N is the number of RR intervals in the data. For scale $\tau = 1$, the coarse-grained time series is simply the original beat-to-beat RR interval time series.
2. SampEn is calculated for each coarse-grained time series. SampEn as a function of the scale factor produces the MSE. MSE for scale factor $\tau = 1$ returns standard SampEn (computed from the original data points).

5. Advice for the analysis

Being derived from the same RR series, those indexed are highly correlated with each other.

6. References

[1] Kubios_HRV_Users_Guide.pdf

[2] kubios_setting.pdf