IMPACT OF THE EDITING OF PATTERNS WITH ABNORMAL *RR*-INTERVALS ON THE ASSESSMENT OF HEART RATE VARIABILITY

Dorota Wejer[a], Danuta Makowiec[a], Zbigniew R. Struzik[a,b]
Marta Żarczyńska-Buchowiecka[c]

[a]Institute of Theoretical Physics and Astrophysics, University of Gdańsk
Wita Stwosza 57, 80-952 Gdańsk, Poland
[b]Graduate School of Education, The University of Tokyo
RIKEN Brain Science Institute, Japan
[c]First Department of Cardiology, Medical University of Gdańsk
M. Skłodowskiej-Curie 3a, 80-210 Gdańsk, Poland

(Received September 8, 2014; revised version received December 4, 2014)

Indices of heart rate variability are calculated twice: firstly from signals with unperturbed normal *RR*-intervals, and subsequently from the same signals edited according to either real patterns of disturbances or random patterns of perturbations. Four methods of editing are applied: (I) deletion of abnormal *RR*-intervals, (II) replacement of abnormal *RR*-intervals by the median, (III) replacement of abnormal *RR*-intervals by a random value from the surrounding *RR*-intervals, (IV) entering the values that result from the statistics of similar patterns. The fractality indices, such as α_1 and α_2 from detrended fluctuation analysis (DFA), and a ratio of *RR*-intervals greater than 50 ms, pNN50, are found to be the most sensitive to editing, independently of the method of editing and the organization of the disturbance pattern.

DOI:10.5506/APhysPolB.45.2103
PACS numbers: 87.19.Hh, 07.05.Hd, 87.18.–h

1. Introduction

Since its discovery by Einthoven [1] in 1895, the electrocardiogram (ECG) has become a very popular medical tool used for the evaluation of the heart function. ECG is a noninvasive, easy to perform and cheap measurement. Numerous methods have been designed to establish and then improve its clinical significance [2]. Moreover, it is commonly accepted that heart rate is driven by the autonomic nervous system to maintain the actual demands of the organism for the nutrient supply [3–6]. Therefore, analysis of fluctuations of time intervals between subsequent heart contractions, so-called heart rate variability (HRV), provides insight into controlling mechanisms by which the autonomic nervous system mediates between the heart action and the needs of the body. HRV analysis has been found to be a valuable tool in the evaluation of different pathological [5, 6] and physiological [7, 8] conditions.

Many methods have been devised to measure HRV from ECG recordings. They are based on signals containing the intervals between successive so-called QRS complexes with distinguished R-peaks. Various equipment is commercially available for storing ECG for later automatic detection of these R-peaks and then for assessment of the beat-to-beat time intervals. Such series are called RR-signals.

Short-term ECG recordings, *i.e.*, less than 30 minutes, which are performed in ambulatory conditions, usually provide signals of adequate quality. Moreover, the automatic extraction of time intervals can easily be visually verified by the specialist. Holter systems such as *e.g.*, Del Mar Reynolds Medical, UK, or Schiller, Switzerland, are portable equipment designed for long-term recording of ECG. Holter systems are also equipped with special algorithms for automatic detection of QRS complexes and the extraction of RR time intervals. However, it is often problematic to achieve a high quality of RR-signals with them [9, 10]. Standard HRV analysis involves methods based on RR-signals from three categories, *i.e.*, time domain, frequency domain and nonlinear measures [5]. Table I contains definitions of the HRV parameters studied by us.

There are two basic reasons for the imperfection of ECG recordings. First of all, long signals contain beats that are different from normal heart contractions, *i.e.* from contractions initiated by the sinus node — the heart's natural pacemaker. Especially ectopic beats, *i.e.* premature beats, which are known to be commonly present in healthy subjects, influence strongly indices of HRV [11]. Secondly, disturbances in ECG appear because of recording artifacts and imperfect automatic algorithms of detecting R-peaks [10]. Recording artifacts may result from badly adhering electrodes or from movement by the patient. The other source of misleading information present in long signals can be related to the nonstationarity which appears as a result of ordinary human activity. Many HRV estimators are established under the assumption that the signal is stationary. Several studies have shown that short-term measures of HRV quickly return to normal after gentle manipulation of the patient's condition. But there are few data on the stability of the HRV measures derived from 24-hour ECG recordings [5].

Many methods of editing have been considered, see, *e.g.* [10, 12, 13] for reviews. The deletion of wrong beats is the most popular method. However, each deletion destroys the continuity of the time. Frequency analysis, in particular, is known to be especially sensitive to the proper timing [14, 15].

TABLE I

List of 11 indices of HRV examined in this study, grouped into the two categories: time domain and nonlinear, with computational details of their significance [5].

Time domain	
Measure	Description [units]
$\overline{RR}$	The mean of normal RR-intervals [ms]
SDNN	Standard deviation of normal RR-intervals [ms]
RMSSD	Square root of the mean of the sum of the squares of differences between normal successive RR-intervals [ms]
pNN50	Percentage of pairs of normal successive RR-intervals that differ more than 50 ms in the entire recording [%]

Nonlinear	
Measure	Description [units]
α_1	Scaling exponent of detrended fluctuation analysis [16] for short-term fluctuations, *i.e.* $4 \leq n \leq 16$
α_2	Scaling exponent of detrended fluctuation analysis [16] for long-term fluctuations, *i.e.* $16 \leq n \leq 64$
SD1	Standard deviation of distances of points in the Poincaré plot (RR_i, RR_{i+1}) to the line-of-identity [ms]
SD2	Standard deviation of distances of points in the Poincaré plot (RR_i, RR_{i+1}) along the line-of-identity [ms]
ApEn	Approximate entropy [17] — the measure of repetitive m-patterns of fluctuations with r accuracy Our setting is $m = 2$, $r = 0.2$ SDNN
SampEn	Sample entropy — the negative logarithm of the conditional probability that two sequences of m successive RR-intervals matching with accuracy r will also match after including the $m+1^{\mathrm{st}}$ RR-interval Our setting is $m = 2$, $r = 0.2$ SDNN
ShEn	Shannon entropy of probability distribution of the line length in recurrence plot

For this reason, many methods of interpolation have been designed. Among them the insertion of the median (zero-line interpolation), linear interpolation, or cubic spline interpolation, as well as more sophisticated numerical methods like piecewise cubic Hermite interpolation [15] or statistical methods based on bootstrapping [18] are used. After experiments with many methods, the HRV indices in the time domain were found to be more robust to editing than those in the frequency domain [10]. The same has been found in the case of nonlinear indices [18].

Following these observations, we also consider that it is crucial to maintain the segment of time which is perturbed independently of the origin of perturbations: physiological or technical. Therefore, in our proposed methods, we make an effort to retain the equality between the length of disturbance and the length of substituted intervals. Moreover, the two new methods proposed by us are motivated by our present knowledge about the short-term dynamics of heart contractions rather than numerical analysis of the curve shape. We propose to replace abnormal RR-intervals by a value chosen at random from the nearby proper intervals (thermal approach), or from a set of similar situations that are found in the whole signal (dynamical approach).

Furthermore, abnormal events might occur in specific patterns, groups of abnormal RR-intervals. Therefore, we also investigate the role of these patterns. The effect of editing them on the popular time-domain and nonlinear HRV indices is investigated through Monte Carlo simulations. In these simulations, signals which are originally free of any disturbances are changed — edited following some real patterns of abnormal beats. Since frequency-domain indices demand very specific methods of preprocessing, different from those proposed in this paper, they are not considered here.

The paper is organized as follows. In Sec. 2, we give the motivation for editing and provide an analysis of patterns of disturbances that have been observed in the ECG signals analyzed. Then in Sec. 3, we define four methods for dealing with these abnormal RR-intervals. The two editing methods are well-known and widely used, *i.e.* deletion of abnormal RR-intervals and replacement of disturbance by the median of the surrounding normal values. In Sec. 4, we describe a Monte Carlo experiment which allows us to study the effects of editing when specific patterns of abnormal sequences are applied. The results of the simulations are presented in Sec. 5. The last section briefly concludes our investigations.

2. Patterns of artifacts

2.1. Motivation for signal editing

Perturbations in ECG, and then, in consequence, in time intervals between subsequent R-peaks, can change the values of indices of HRV, as exemplified in Fig. 1 and Table II. Figure 1 contains two tachograms with ECG signals. The top tachogram shows a signal consisting of 300 beats. The abnormal beats are labeled (annotated) as A. All proper heart contractions are marked by N. Since for HRV analysis only normal-to-normal time intervals can be considered (so-called NN-intervals), the first N-interval after the sequence of abnormal beats must also be excluded from further analysis. This first N value, after an ectopic beat, is called a compensatory pause [5].

Similarly, the first N is skipped in the case of technical artifacts. Because of these rules, 6 beats of the 300 beats considered (*i.e.*, 2%) have to be edited. The bottom tachogram shows the same RR-signal with the abnormal beats removed. Table II presents values of parameters of HRV for both RR-signals.

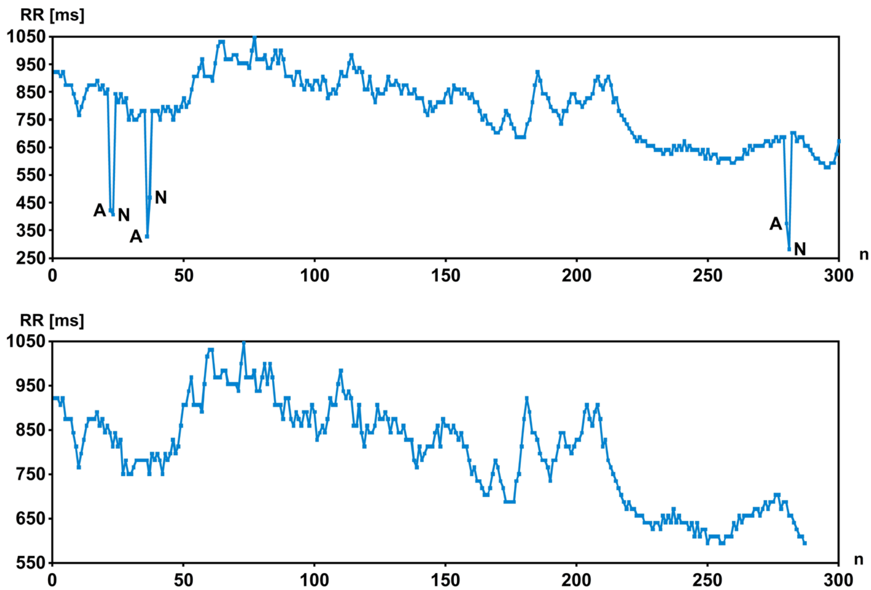

Fig. 1. Example of tachograms with RR-signals: the top signal — with artifacts, the bottom signal — edited. The top signal has three pairs of perturbations consisting of intervals annotated as A and the first following N-interval. The bottom signal has these six intervals removed, which results in a change in timing of subsequent NN-intervals, shortening the bottom signal.

TABLE II

Measures of HRV for both RR-signals from Fig. 1.

Time domain	$\overline{RR}$	SDNN	RMSSD	pNN50
Top signal	789.29	127.19	63.51	7.69%
Bottom signal	795.60	113.88	27.17	5.12%

Nonlinear	α_1^*	α_2^*	SD1	SD2	ApEn*	SampEn	ShEn
Top signal	1.07	1.12	44.98	174.17	0.74	0.71	3.44
Bottom signal	1.44	1.27	19.24	159.84	0.73	0.69	4.03

*Remark: index does not provide relevant information for short data.

The huge differences observed between values obtained for the top signal and values calculated for the bottom signal provide convincing evidence that abnormal beats change the HRV message substantially.

2.2. Patterns of artifacts

We studied ECG signals recorded from 196 healthy volunteers. These persons were adults, of different ages (from 18 to 94 years) and both genders. None of them suffered from chronic or acute diseases. All the subjects underwent 24-hour Holter monitoring during a normal sleep–wake rhythm. The Holter recordings were preliminarily analyzed by Del Mar Reynolds Impresario software for premature, supraventricular and ventricular beats, missed beats and pauses. The QRS complexes were detected and classified automatically by the software. In all the signals, the sampling rate was 128 Hz, which enabled the detection of R-peaks with an approximately 8 ms accuracy. The first part (1000 beats) and the last part (1000 beats) of each of the 196 recordings were not included in the considerations. These parts were assumed to correspond to the transient stage of each subject. Namely, a person is always aroused after the equipment is put on the body, and also when waiting for the equipment to be removed from the body. This arousal often results in perturbations in ECG recordings.

Since normal heart rhythm is defined by the rate of the sinus node depolarization, represented by the onset of the P-wave rather than the R-peak, each signal was inspected visually by an experienced cardiologist to verify the normality of the heart contraction. Finally, we obtained series of time intervals between subsequent R-peaks, together with their annotations as normal (N), ventricular (V), supraventricular (S), unspecified (U), or artifact (A). These signals are called RR-intervals. Figure 2 shows mean numbers of technical and cardiac abnormal RR-intervals in one RR-signal. Figure 2 demonstrates the quality of our recordings and their structure. We see that the number of arrhythmias increases with age and the recordings contain many more technical artifacts than the arrhythmias. In the following, we distinguish only normal beats from all the other beats. Each beat which is not normal is called abnormal and is annotated as A. Based on two-line information from ECG signal: the RR-interval and its annotation, the signals of NN-intervals are constructed by four methods of editing.

The abnormal events occur in specific patterns, groups of abnormal RR-intervals. Let us call a sequence of RR-intervals *K-isolated* if there are at least $K+1$ normal R-peaks prior and K normal R-peaks after the given sequence.

For our editing purposes, three types of abnormal groups of RR-intervals have been identified in the signals studied. They are as follows:

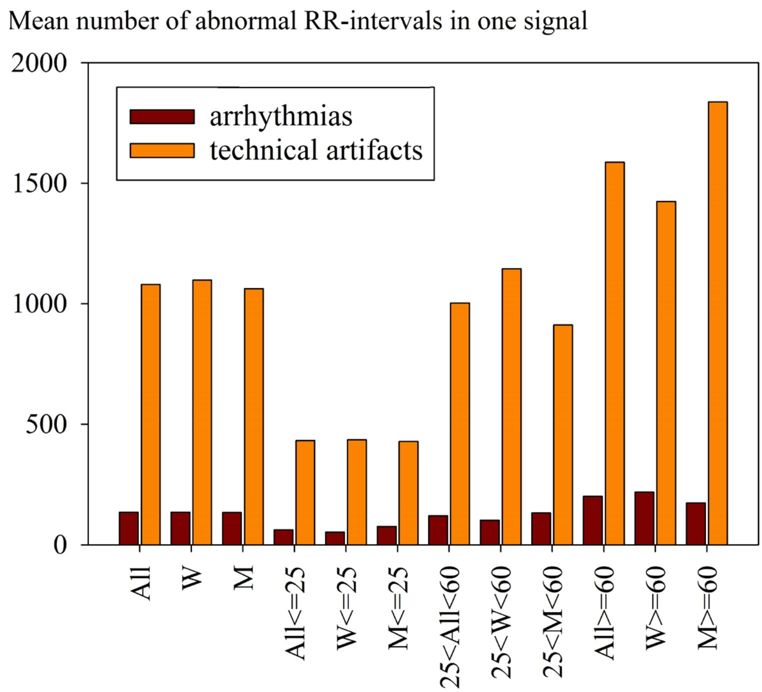

Fig. 2. Mean numbers of abnormal RR-intervals (arrhythmias and technical artifacts) in one 24 h Holter signal. The data are presented in 12 groups: All — all the signals (196), W — signals from women (97), M — signals from men (99), All $\leq$ 25, W $\leq$ 25, M $\leq$ 25 — signals from all (36), women (21), men (15) at the age at least 25 years, 25 < All < 60, 25 < W < 60, 25 < M < 60 — signal from all (99), women (39), men (60) between 25 and 60 years, All $\geq$ 60, W $\geq$ 60, M $\geq$ 60 — signal from all (61), women (37), men (24) over the age of 60.

— Type 1: *K-isolated abnormal beats*
  A sequence of abnormal beats of any length is isolated.
  In the following, we assume that $K = 4$. Hence the annotation sequence corresponding to type 4-isolated artifact of type 1 is (the group of beats that need editing in gray/red)

$$\dots NNNNNA\dots ANNNN\dots$$

— Type 2: *K-isolated group of two sequences with abnormal beats*
  A pattern contains two sequences with abnormal R-peaks. These sequences are separated by at most K normal R-peaks. For each possible number of beats annotated as N which are present between the two regions of artifacts, further subtypes can be considered.

In the case of $K = 4$, we have four (a)–(d) possible subtypes (with the beats that have to be edited in gray/red):

$$(a) : \ldots NNNNNA \ldots ANA \ldots ANNNN \ldots$$
$$(b) : \ldots NNNNNA \ldots ANNA \ldots ANNNN \ldots$$
$$(c) : \ldots NNNNNA \ldots ANNNA \ldots ANNNN \ldots$$
$$(d) : \ldots NNNNNA \ldots ANNNNA \ldots ANNNN \ldots$$

Note that this classification allows us to use the whole correct information from the original data.

— Type 3: *Other patterns of abnormal beats*
This type covers all other patterns of disturbance that we encounter in the series studied.

Figure 3 presents the total numbers of patterns with disorders classified according to the types listed above in an average recording. In the same figure, we provide information about the origin of the abnormalities, *i.e.* whether they are technical artifacts, or arrhythmias, or a mixture of both.

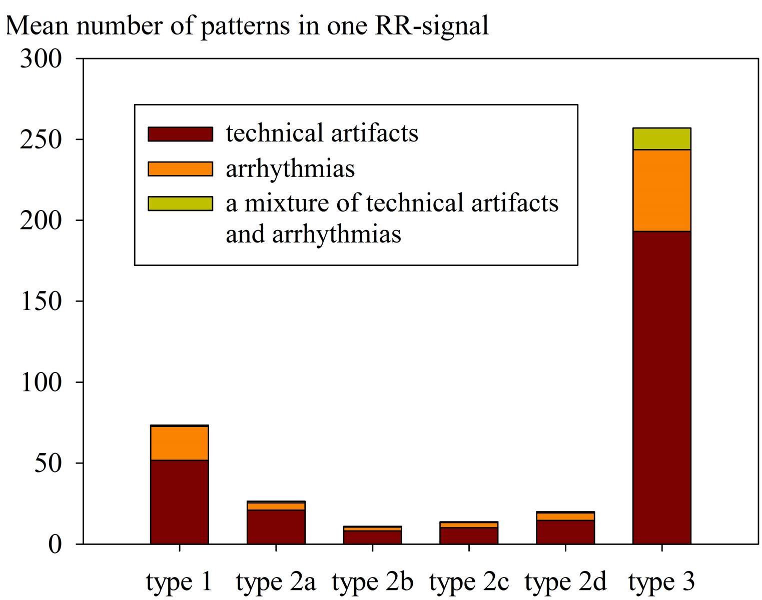

Fig. 3. Types of patterns of disorders with their origins.

We can see that patterns of type 3 dominate in the signals considered. These statistics strongly depend on the K value. If K is a large number, sequences of type 3 are more likely to occur. We decided to use $K = 4$ to balance in an optimal way the number of disturbances which is possible to treat in a systematic way and the size of the neighborhood acceptable for the editing purposes. The complicated structure of patterns of type 3 prevents us from automatic editing. The effective algorithm for the substitution is difficult to design. However, deletion of the whole disturbance pattern is always possible.

3. Editing methods

3.1. Editing of patterns of type 1

Let us assume that we have a K-isolated sequence of abnormal beats, *i.e.* we edit disturbance of type 1. Let this disturbance start at RR_i interval and consist of k abnormal RR-intervals.

Method I: Deletion of abnormal *RR*-intervals

Together with the deletion of abnormal intervals, the first N-interval is deleted. After the deletion, the two parts of the signal are concatenated.

Advantages: This is an easier method for implementation. It can be performed without any numerical instabilities.

Drawbacks: It strongly affects the temporal sequence. Consequently, a shorter signal is obtained.

Method II: Replacement of abnormal *RR*-intervals by the median estimated from neighboring NN-intervals

K subsequent NN-intervals prior to perturbation and $K-1$ NN-intervals after the disturbance are chosen in the calculation of the median. Each abnormal RR-interval is replaced by the median, however with the following constraint on the number of substitutions allowed: the length of all the abnormal RR-intervals must match the length of the substituted intervals, *i.e.*, k abnormal beats will be replaced by k' beats with median value me with $me/2$ accuracy, see Fig. 4,

$$\sum_{j=0}^{k} RR_{i+j} - \frac{me}{2} < k' \cdot me < \sum_{j=0}^{k} RR_{i+j} + \frac{me}{2} \,.$$

Figure 4 shows the application of the second editing method in the case of $K = 4$.

Advantages: Replacement of abnormal RR-intervals by the median maintains the temporal order. The NN-signal does not obtain new values, which might happen if the median were replaced by the mean of neighboring NN-intervals.

Drawbacks: The NN-signal is set to constant in edited fragments. The short-term correlations in the NN-signal might be affected.

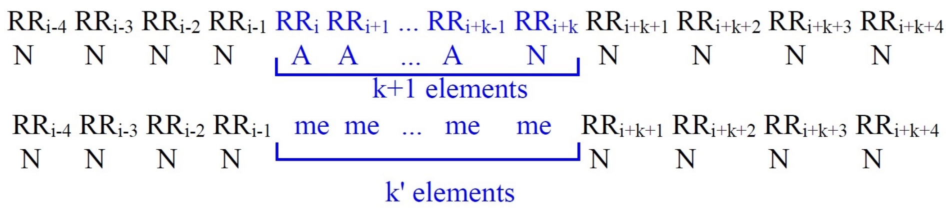

Fig. 4. The second editing method — replacing abnormal RR-intervals by the median.

Method III: Replacement of abnormal *RR*-intervals by values randomly chosen from neighboring NN-intervals

As before, K subsequent NN-intervals prior the perturbation and $K-1$ NN-intervals after the disturbance are used to form a set of intervals from which values are drawn. We assume equal probability for each value. Again, the constraint about the number of possible k' replacements is applied as follows, see Fig. 5,

$$\sum_{j=0}^{k} RR_{i+j} - \frac{m}{2} < \sum_{i=1}^{k'} v_i < \sum_{j=0}^{k} RR_{i+j} + \frac{m}{2},$$

where m is the mean of random values, v_i — random value drawn from the set of $2K-1$ values with neighboring NN-intervals, k' — number of substituted intervals. Figure 5 explains the algorithm described in the case of $K=4$.

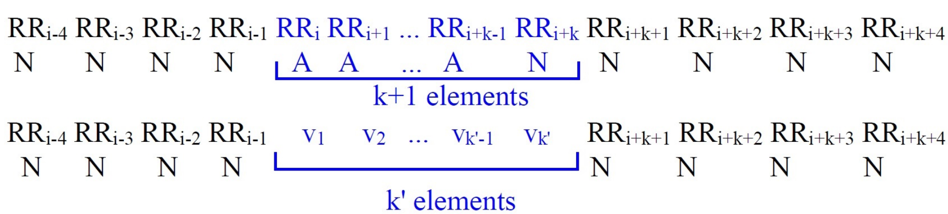

Fig. 5. The third editing method — replacement of abnormal RR-intervals by values randomly chosen from neighboring normal RR-intervals.

Advantages: The temporal order in this method is retained and there are no new values in the NN-signal after editing. The edited parts of the signal are not set to constant.

Drawbacks: This method destroys correlations in the edited part of NN-signal.

Method IV: Replacement of abnormal RR-intervals by NN-intervals found in the set of the most similar patterns

We use 500 NN-intervals before the disturbance and 500 NN-intervals after the disturbance in the search for NN-intervals similar to the last two normal RR-intervals prior to the abnormal beat, see Fig. 6.

$$\begin{array}{c}RR_{i-4}\\ N\end{array}\ \begin{array}{c}RR_{i-3}\\ N\end{array}\ \begin{array}{c}RR_{i-2}\\ N\end{array}\ \begin{array}{c}RR_{i-1}\\ N\end{array}\ \underbrace{\begin{array}{cccc}RR_{i} & RR_{i+1}\ \ldots & RR_{i+k-1} & RR_{i+k}\\ A & A\ \ldots & A & N\end{array}}_{k+1\text{ elements}}\ \begin{array}{c}RR_{i+k+1}\\ N\end{array}\ \begin{array}{c}RR_{i+k+2}\\ N\end{array}\ \begin{array}{c}RR_{i+k+3}\\ N\end{array}\ \begin{array}{c}RR_{i+k+4}\\ N\end{array}$$

$$\begin{array}{c}RR_{i-4}\\ N\end{array}\ \begin{array}{c}RR_{i-3}\\ N\end{array}\ \begin{array}{c}RR_{i-2}\\ N\end{array}\ \begin{array}{c}RR_{i-1}\\ N\end{array}\ \underbrace{p_1\quad p_2\ \ldots\ p_{k'-1}\quad p_{k'}}_{k'\text{ elements}}\ \begin{array}{c}RR_{i+k+1}\\ N\end{array}\ \begin{array}{c}RR_{i+k+2}\\ N\end{array}\ \begin{array}{c}RR_{i+k+3}\\ N\end{array}\ \begin{array}{c}RR_{i+k+4}\\ N\end{array}$$

Fig. 6. The fourth editing method — entering values from the past and the future that occurred in a similar sequence.

We assume that the vector $\boldsymbol{v}_r = [RR_{r-1}, RR_{r-2}]$ is similar to the vector $\boldsymbol{v}_0 = [RR_{i-1}, RR_{i-2}]$ with $p \in [0, 0.05)$ accuracy if

$$\begin{aligned}(1-p)RR_{i-1} \ &< \ RR_{r-1} < (1+p)RR_{i-1}\,,\\ (1-p)RR_{i-2} \ &< \ RR_{r-2} < (1+p)RR_{i-2}\,.\end{aligned}$$

All vectors similar to a given $\boldsymbol{v}_0$ which are found in the described set of 1000 surrounding NN-intervals are selected to form a set of possible similar patterns. One of these patterns is chosen at random. The abnormal interval RR_i is replaced by RR_r and its annotation is set to N.

We set the size of the search sequence to 1000 RR-intervals around the disturbance, because we checked that this size is a minimal neighborhood in which we can find at least one proper sequence for each edited perturbation. Specifically, in the case of signals considered with $p = 0.01$, the set with the most similar patterns was usually found not to be empty (in 90% of the edited signals). However, if this set was empty, the p-value was increased by 0.01 until a vector similar to a given one was found. With our signals, only 3% of recordings needed $p \geq 0.03$.

After replacing the first abnormal interval, we repeat the algorithm for the next abnormal RR-interval from the pattern, *i.e.* we move to $\boldsymbol{v}_0 = [\widetilde{RR}_i, RR_{i-1}]$, where $\widetilde{RR}_i$ denotes the corrected RR_i interval.

Remembering the temporal sequence order, we again impose a similar constraint on the number k' of intervals entered

$$\sum_{j=0}^{k} RR_{i+j} - \frac{m}{2} < \sum_{i=1}^{k'} p_i < \sum_{j=0}^{k} RR_{i+j} + \frac{m}{2}\,,$$

where m — mean of value from the new NN-intervals entered, p_i — subsequent new NN-interval. Figure 6 presents the usage of method IV.

Advantages: The correlations and the temporal order are retained. There are no new values in the NN-signal after editing.

Drawbacks: This method is the most complicated of the methods presented. It uses additional parameters: the size of the interval scanned with NN-intervals and the accuracy of matching.

3.2. Editing of patterns of type 2

Direct application of the methods described above to a pattern of type 2, namely a K-isolated group of two sequences with abnormal beats, is not straightforward. Modifications are needed which allow the inclusion of all normal RR-intervals present between the groups of abnormal beats.

Method I — deletion can be applied without any correction. It is easy to see that any disturbance of type 2 (a) can be edited in the same way as the pattern of type 1. In the case of the last subtype on the list of possible disturbances of type 2 (in our case, this means type 2 (d)), the length of normal intervals between abnormal RR-intervals is large enough to apply any editing method considered for the first sequence of abnormal beats. So we edit the first sequence as it is a K-isolated disturbance, and then use the edited sequence in repairing the second sequence.

For all other subtypes, we adopt the following strategy. The set consisting of K NN-intervals prior to the whole disturbance and $K-1$ NN-intervals after the disturbance is enlarged by normal values extracted from the range between the two sequences with abnormal beats. In this way, the median of method II, as well as accessible NN-intervals in method III, is calculated in larger sets, though for each abnormal sequence the timing is treated separately. In particular in the case of $K = 4$, this means that the median is calculated/values are drawn (2^{nd}/3^{rd} methods, respectively) from a set of 8 (type 2 (b)) or 9 (type 2 (c)) NN-intervals.

Since method IV is applied iteratively and one beat is corrected after another, in this case only the timing must be controlled. We assume that each sequence of abnormal beats in the pattern is edited independently with its individual timeline.

4. Monte Carlo experiment on inserting abnormal beats

We thoroughly analyzed the 196 Holter signals to find the parts which were free of any disturbances. Moreover, because HRV during the night is known to be especially complex [19], we additionally assumed that these clean parts would correspond to the nocturnal rest of a subject. We found only 21 recordings which met our expectations. All these ECG recordings were obtained from healthy, young people (11 females, 10 men; mean age 22.4 ± 1.6).

Each accepted signal consisted of 3500 RR-intervals classified as normal and related to sleeping time, namely corresponding to 2 a.m.–4 a.m. The first 500 RR-intervals and the last 500 RR-intervals were used only for the proper execution of method IV, and were not included in the calculation of HRV parameters.

We scanned our ECG resources again to find natural patterns of disturbances that might occur during nocturnal rest. Our interest was focused on patterns satisfying the following criteria: all disturbances in a sequence of 2500 RR-intervals were of type 1 or type 2, and the total number of abnormal beats was approximately 2%. Only five signals were found which fulfilled these demands.

Annotations of these five patterns were applied to originally perfectly clean signals. Then we cyclically attached the annotations starting at a randomly chosen place in the clean signal, see Fig. 7.

N	N	N	N	...	N	N	N	...	N	N	N	N
844	836	840	848	...	904	896	892	...	852	864	868	856
1	2	3	4	...	k-1	k	k+1	...	2498	2498	2499	2500

random location: k

P_{2502-k}	P_{2503-k}	P_{2504-k}	P_{2505-k}	...	P_{2500}	P_1	P_2	...	P_{2498-k}	P_{2499-k}	P_{2500-k}	P_{2501-k}
844	836	840	848	...	904	896	892	...	852	864	868	856
1	2	3	4	...	k-1	k	k+1	...	2498	2498	2499	2500

Fig. 7. Cyclical attachment of a pattern with disorders to a clean signal at a random location. The top line describes the original signal with all annotations set to N. The bottom line shows the same RR-intervals but annotated according to a pattern P cyclically shifted to a randomly chosen position k.

For comparison, we edited signals for which patterns of annotations were constructed at random. Fifty random patterns with 2% abnormal beats in total, arranged in type 1 or 2 disturbances, were generated for each of 21 clean signals.

In this way, we obtained two groups of synthetic signals with disturbances for which the effects of editing could be tested. Each group considered consisted of 1050 files. The main group provided the opportunity to study the influence of editing if the perturbation was of natural origin, while the second group allowed the verification of whether natural disturbances did have specific patterns and because of them, the editing provided significantly different results.

5. Results

Every synthetic signal was edited by each of the methods described earlier. Then, all the HRV parameters listed in Table I were calculated for each edited signal. For each value of the HRV index calculated, we found its percentage error, *i.e.*

$$\mathrm{PE}_X = \frac{|X_0 - X_{\mathrm{ed}}|}{X_0} \times 100\% \,,$$

where X_0 is the value of the HRV parameter of the original signal, and X_{ed} is the value of the same parameter obtained from the synthetic signal after editing.

The values of PE_X obtained were collected in groups corresponding to the HRV index (described in Table I), method of editing (I–IV), and origin of disturbances (natural or random). In each group the distribution of PE_X values was investigated by counting the number of values which fell in the following intervals:

$$\begin{array}{llll} 0 \le \mathrm{PE}_X \le 1\,, & 1 < \mathrm{PE}_X \le 3\,, & 3 < \mathrm{PE}_X \le 5\,, & 5 < \mathrm{PE}_X \le 10\,, \\ 10 < \mathrm{PE}_X \le 30\,, & 30 < \mathrm{PE}_X \le 50\,, & 50 < \mathrm{PE}_X \le 100\,. & \end{array}$$

If a given method of editing signals provides results falling into the first interval, we can say that the method correctly reconstructs the signal properties. If the values are found within the second interval, we should see that the method corresponds approximately linearly to 2% changes of the basic signal. Results that are obtained with errors larger than 5% measure how divergent the given editing procedure could be for a given HRV index.

The stacked bar charts in Figs. 8 and 9 present distributions of PE_X values (in percentage) for the subsequent HRV parameters and the given method of editing. The acceptable results, *i.e.*, when $\mathrm{PE}_X \le 3\%$ with probability greater than 0.95, and not acceptable, divergent results, when $\mathrm{PE}_X > 5\%$, were obtained with the probabilities enumerated in Table III.

Thus, we see that

— there is a HRV parameter, namely the mean value of $\overline{RR}$, which is insensitive to the method of editing. Also standard deviation of the mean, SDNN, provides values close to the original values in unedited signals[1];

[1] These observations do not contradict the results in Table II, because here we compare indices of edited synthetic signals to an unperturbed real signal, whereas, in the introductory example, the comparison is made between a signal with perturbations and an edited signal.

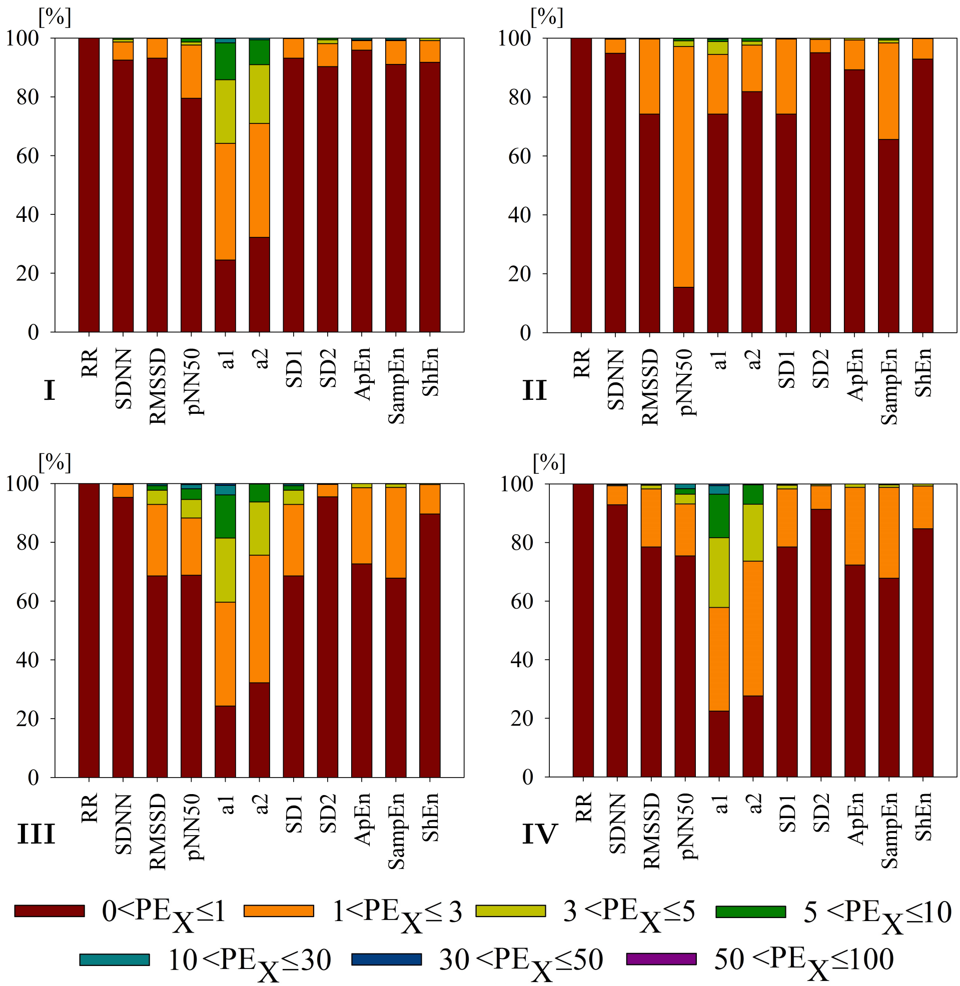

Fig. 8. Results of editing for the group of signals with random patterns of disturbances.

— the following two nonlinear indices: SD2 index of Poincaré plot and Shannon entropy can be considered not to be dependent on the editing method;

— pNN50 and the nonlinear parameters of fractality α_1 and α_2 are the most sensitive to editing, independently of the method applied;

— the best stability of our results was achieved with the second editing method, by replacing abnormal *RR*-intervals by the median from the close neighborhood (similar results were obtained in [20] and [18]);

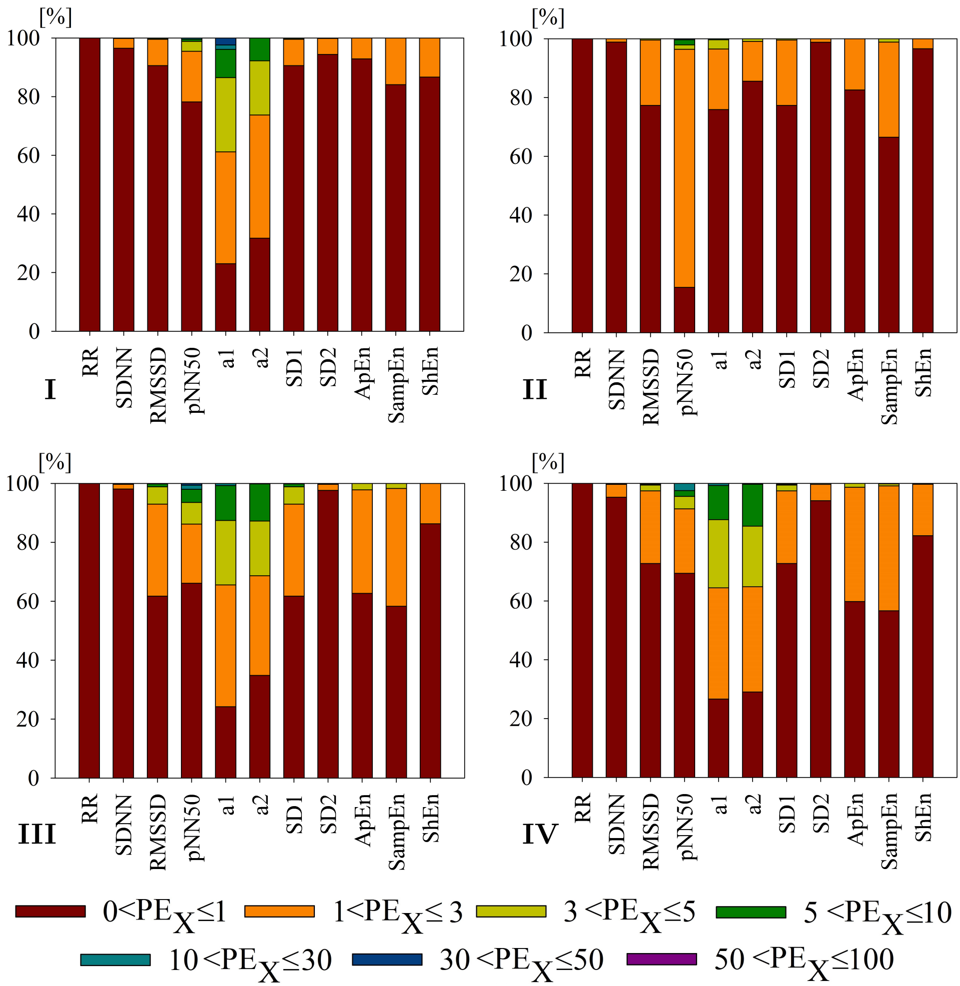

Fig. 9. Results of editing for the group of signals with random patterns of disturbances.

— editing of random patterns of perturbations, in general, provides similar results to those obtained from natural patterns. Only for α_2, we observed significantly larger number of strongly different results;

— a comparison of effects of editing by methods III and IV shows us that some indices responsible for the short-term variability (RMSSD, SD1) more frequently produce correct values when method IV is applied, although, to our surprise, neither ApEn nor SampEn are sensitive to these methods.

TABLE III

Lists of acceptable editing methods for a given index, namely, methods giving errors less than 3% at Prob > 0.95, together with lists of editing methods providing results with values significantly different from the original. In brackets, the corresponding characterization is given for the random patterns with disturbances.

HRV index name	Methods providing $PE_X \leq 3\%$	Methods providing $PE_X > 5\%$	Probality to get $PE_X > 5\%$
RR	I, II, III, IV (I, II, III, IV)		
SDNN	I, II, III, IV (I, II, III, IV)	I II III, IV	0.5% (0.1%) 0.3% (0.1%) 0.2% (0.3%)
RMSSD	I, II, IV (I, II, IV)	I, II III IV	0.2% (0.4%) 0.2% (0.1%) 2.3% (1.1%) 0.5%
pNN50	I, II (I, II)	I II III IV	1.3% (1.2%) 1.0% (2.1%) 5.4% (6.5%) 3.5% (4.5%)
α_1	none (II)	I II III, IV	14.2% (13.5%) 1.1% (0.4%) 19% (13%)
α_2	II (II)	I II III IV	9.0% (7.8%) 1.0% (0.2%) 6.9% (13%) 6.9% (15%)
SD1	I, II, IV (I, II, IV)	I II III IV	0.2% (0.4%) 0.2% (0.1%) 2.3% (1.1%) 0.5% (0.6%)
SD2	I, II, III, IV (I, II, III, IV)	I II (III) IV	0.6% (0.1%) 1.0% (0.1%) 0.2%
ApEn	II, III, IV (I, II, III, IV)	I II, III	0.8% (0.1%) 0.1%
SampEn	II, III, IV (I, II, III, IV)	I II III, IV	0.9% (0.1%) 0.7% (0.1%) 0.2% (0.2%)
ShEn	I, II, III, IV (I, II, III, IV)	I (IV)	0.1% (0.1%)

6. Summary

Deleting improper contents is the most straightforward approach to correcting missing or wrong beats. Replacing these beats by the median is in agreement with the idea of homeostatic equilibrium, which, if properly maintained, allows the cardiovascular system, controlled by the autonomic nervous system, to deliver blood to the organs most adequately and efficiently. Our proposition of editing expressed in method III, namely, entering a kind of thermal noise around the median, instead of a single median value, follows the idea of the thermal equilibrium. In contrast to the thermal noise, method IV tests the existence of a possible dynamical dependence in the short term dynamics. The almost negligible differences in results provided by the third and fourth methods give a further indication that our understanding of the regulation of the cardiovascular system is still unsatisfactory.

Our considerations have been limited to patterns of disturbances of specific types. Moreover, we do not discuss the problem of the stationarity of a signal. However, since method II of editing proved to be the most stable when compared to other methods under analysis, its usage is recommended in preprocessing of ECG recordings for HRV analysis.

We thank the Polish National Science Center for financial support: UMO: 2012/06/M/ST2/00480.

REFERENCES

[1] W. Einthoven, *Über die Form des menschlichen Electrocardiogramms*, *Pflügers Arch. fd ges. Physiol.* **60**, 101 (1895).

[2] M.E. Josephson, *Clinical Cardiac Electrophysiology*, Lippincott Williams & Wilkins, 2008.

[3] M.M. Wolf *et al.*, *Med. J. Australia* **2**, 52 (1978).

[4] M. Malik, A.J. Camm, Eds. *Heart Rate Variability*, Armonk, Futura, 1995.

[5] M. Malik, *Eur. Heart J.* **17**, 354 (1996).

[6] M. Malik, A.J. Camm, eds. *Dynamic Electrocardiography*, Blackwell, Futura, 2004.

[7] D.F. Dietrich *et al.*, *Eurospace* **8**, 521 (2006).

[8] M. Reardson, M. Malik, *Pacing Clin. Electrophysiol.* **19**, 1863 (1996); S.M. Pikkujämsä *et al.*, *Circulation* **100**, 393 (1999); F. Beckers, B. Verheyden, A.E. Aubert, *Am. J. Physiol. Heart Circ. Physiol.* **290**, H2560 (2006); D. Makowiec *et al.*, *Eur. Phys. Lett.* **94**, 68005 (2011).

[9] M. Malik, *Standard Measurement of Heart Rate Variability*, in: [6], pp. 13–21.

[10] M.A. Peltola, *Front. Physiol.* **3**, 148 (2012).

[11] M.V. Kamath, E.L. Fallen, *Correction of Heart Rate Variability Signal for Ectopics and Missing Beats*, in: [4], pp. 75–85.

[12] A.L. Goldberg *et al.*, *Circulation* **101**, e25 (2000).

[13] M. Peltola, *Analysis of Heart Rate Variability from 24-hour Ambulatory Electrocardiographic Recordings*, *Acta Universitatis Ouluensis D Medica*, **1087** (2010).

[14] G.B. Moody, *Comp. Cardiology* **20**, 715 (1993).

[15] K.K. Kim *et al.*, *Physiol. Meas.* **30**, 1039 (2009).

[16] C.K. Peng *et al.*, *Chaos* **5**, 82 (1995).

[17] S.M. Pincus, *PNAS* **88**, 2297 (1991).

[18] K.K. Kim *et al.*, *Comp. Meth. Programs Biomedicine* **106**, 210 (2012).

[19] E. Vanoli *et al.*, *Circulation* **91**, 1918 (1995).

[20] K.K. Kim *et al.*, *Physiol. Meas.* **28**, 1485 (2007).