

# Clinical Severity of Noise in ECG

Estrella Everss-Villalba<sup>1</sup>, Francisco Melgarejo-Meseguer<sup>1</sup>, Francisco Javier Gimeno-Blanes<sup>2</sup>, Salvador Sala-Pla<sup>3</sup>, Manuel Blanco-Velasco<sup>4</sup>, José Luis Rojo-Álvarez<sup>5</sup>, Arcadio García-Alberola<sup>1</sup>

<sup>1</sup> Unidad de Arritmias. Hospital Clínico Universitario Virgen de la Arrixaca (Murcia), Spain

<sup>2</sup> Escuela Politécnica Superior de Elche. Universidad Miguel Hernández, Elche (Alicante), Spain

<sup>3</sup> Instituto de Neurociencias. CSIC. Universidad Miguel Hernández, Elche (Alicante), Spain

<sup>4</sup> Departamento de Teoría de la Señal y Comunicaciones. Universidad de Alcalá (Madrid), Spain

<sup>5</sup> Dep. Teoría de la Señal y Comunicaciones. Universidad Rey Juan Carlos (Madrid), Spain.

## Abstract

*Objective noise metrics such as the signal to noise ratio and the root mean squared value, are not always related to the actual impact that noise can have on the clinical evaluation of cardiac signals. This work is intended to design a database along with a set of criteria to be used as an initial solid gold-standard of noise severity within a scale created from the clinical point of view. Different from previous approaches, we used recordings from external cardiac event recorders, which have a signal morphology and characteristics typical for the usual Holter monitoring devices. These recorded events often correspond to noise segments, which makes simpler the retrieval of a variety of signal segments with real noise examples. We generated a database recordings from 8 patients (4 female and 4 male with a clinical indication for long-term monitoring,  $64.75 \pm 25.44$  years old). In a first attempt, a set of rules for clinical quality recording were initially defined by a cardiologist, so segments with clear clinical description were categorized into free, low, moderate, hard, or other noise. Segments with unclear description are set apart for revision. This process was iterated until convergence of expert opinion and labels. A specific software tool was created and modified for supporting the process for each iteration. The relative duration of the different kinds of noise in both leads according to their clinical severity was: 4h 14m 15s (free, 37.32%), 2h 54m 52s (moderated, 25.67%), 2h 21m 46s (hard, 20.81%), 1h 43m 13s (low, 15.15%) and 07m 07s -other- (1.05%). The generated database and criteria represent a valuable gold standard for clinical quality evaluation of noise impact on ECG signals.*

## 1. Introduction

Electrocardiogram (ECG) signals recorded in portable devices like Holter are highly sensitive to different types

of noise, and the corrupted segments need to be identified in order to avoid false diagnoses. The duration of the recorded signal in long-term monitoring devices (LM) may be several times higher than the obtained in usual 24h recordings, thus increasing the possibility of noisy records. Intensive signal processing with reasonable computational times is required in this setting for accurate noise detection. Many studies have evaluated noise detection algorithms with sets of ECGs from public repositories, such as the MIT-BIH Arrhythmia Database (MITDB) or the MIT-BIH Noise Stress Test Database (NSTDB) [1, 2].

In order to mimic real noisy recordings and to prove different preprocessing methods, clean ECGs are usually corrupted with usual types of noise, like baseline wander (BW), muscle artifact (MA), electrode motion (EM) or 50-60Hz powerline interference and with different magnitudes of them [3].

Previous works in the literature rely on mathematical and quantitative ECG noise measurements, such as root mean square (RMS) or signal to noise ratio (SNR) [4, 5]. However, these usual measurements may be insufficient in clinical environments, as far as sometimes a moderate SNR can barely affect the ECG morphology, whereas similar SNR in different conditions will dramatically decrease the signal quality and prevent a reliable diagnosis. Therefore, we focused on creating a noise classification based on its clinical impact, rather than power noise and signal magnitude descriptions, which may be more quantitative but often are not closely related to the ability of interpreting all or some characteristics of the distorted ECG segments from a clinical point of view.

For this purpose, our first step, presented in this work, was to generate a clinical gold-standard for evaluating noise severity in ECG Holter recordings, independently of previous quality measurements, such as the SNR or RMS based ones. The problem was approached in two steps: first, a library of noise examples was compiled, using a set

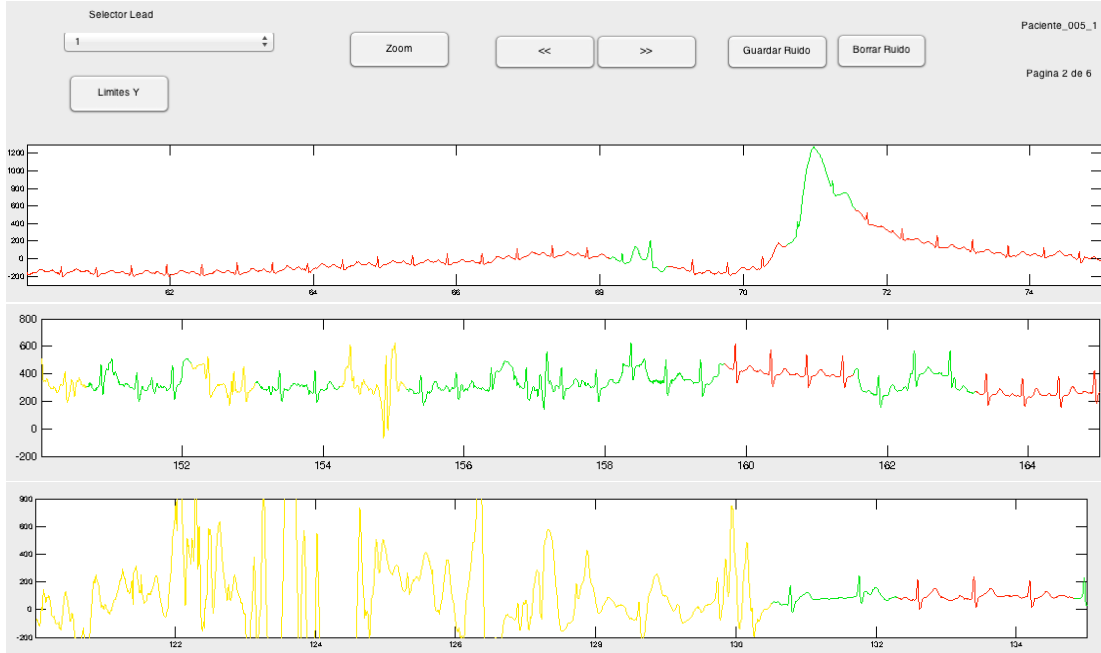


Figure 1. Example of the application used for the iterative process for labeling the event Holter recordings from 2 patients in terms of its noise and impact on the diagnosis quality of the ECG. In red, low noise (type 1) where P and T waves and the QRS complexes are readable; in green, moderated noise (type 2) where only the QRS complexes are reliably identified, and in yellow, hard noise (type 3, QRS complexes were not clearly recognizable). ECG Segments from Patient 5 (upper panel) and from Patient 8 (middle and lower panel).

of real ECG signals obtained from a long-term Holter event recorder; second, a set of criteria were defined and validated in this database, which can be used both for knowledge extraction and to be extended it in the future.

The scheme of the paper is as follows. In the next section, we describe our patient database and the generated software to visualize and to label the Holter recordings. Then, in Section 3, the results are presented and the achieved noise classification criteria are summarized. Section 4 includes the discussion of the results and the conclusions of the present work.

## 2. Methods

Instead of using conventional 24h ambulatory Holter monitoring devices, which perform a continuous, long-term recording of the ECG signal, we used data from patients with a long-term external event recorder (EER). These devices perform a continuous ECG ambulatory monitoring and analyze the signal in real time looking for QRS complexes. Arrhythmic events detection is based on the rate and regularity of the QRS series and a short ECG strip is recorded when such an event is suspected [6]. Muscle or motion artifacts are often misinterpreted as QRS complexes by the device and detected as arrhythmias.

Therefore the recorded strips are more likely to contain and show noisy segments than the obtained in other kinds of monitoring devices and provide an ideal tool for assembling a database to configure a noise classification. We used the standard configuration in EER devices, in which every detected event (correctly or incorrectly interpreted as arrhythmic) triggers an automatic recording for 30 to 300 seconds, and occasionally longer. EER recorders were used for the ECG signal registering, namely, the Spider-FlashT from Sorin Group. Sampling frequency was 200 Hz. The system has 3 electrodes, yielding 2 signals for subsequent analysis. Our database consisted of data from 8 patients (4 women and 4 men,  $64.75 \pm 25.44$  years) who had been referred to our Arrhythmia Department in the University Hospital Virgen de la Arrixaca (Murcia, Spain) for palpitation, syncope or presyncope evaluation.

The following methodology was followed for providing with the description and labelling of the noise severity. First, an expert cardiologist proposed an initial qualitative description of noise degrees in Holter recordings in terms of its impact on the clinical diagnosis of basic parameters, such as waveforms distortion on the ECG. Then, a trained expert followed these criteria for manually labeling 200 ECGs randomly selected from the available in our patient database. Complicated cases were subsequently re-

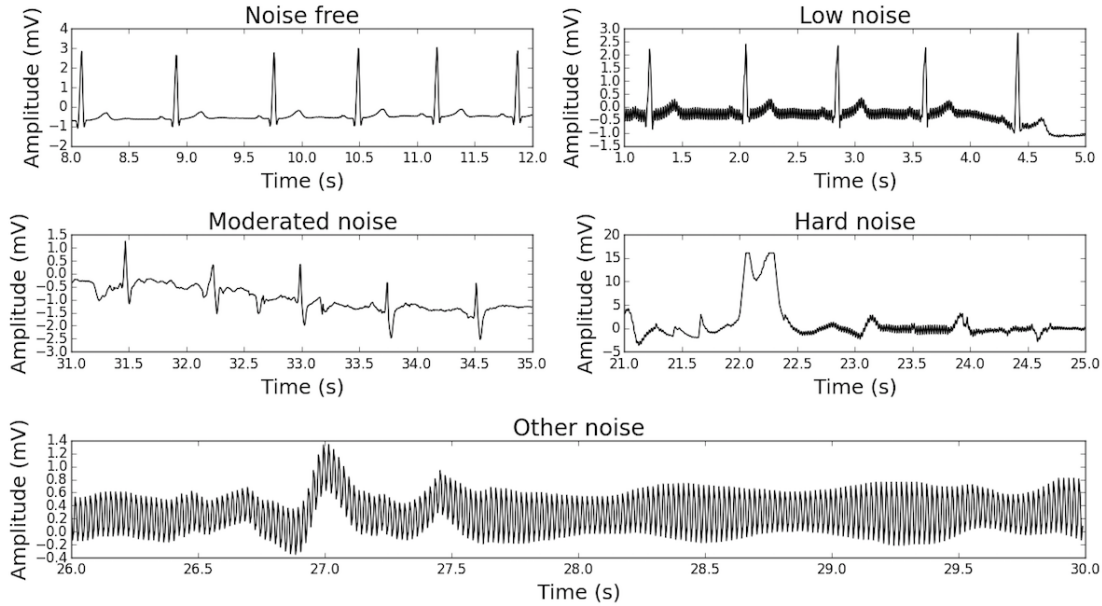


Figure 2. Isolated examples of types of noise according to the initial classification: noise free (type 0), low noise (type 1), moderate noise (type 2), hard noise (type 3) or other noise (type 9).

vised and several criteria were adjusted. After some few iterations and feedback among the experts and by using a specifically developed interface, the final set of criteria were settled and the complicated cases were relabelled.

A specific software tool was developed using Matlab<sup>R</sup>, and based on the iterative process and on the clinical experts suggestions during its use. The visualization interface provided the selected record, divided in 4 panels of 15 seconds (time in  $x$ -axis, amplitude in  $y$ -axis) and indicating the label and time of that record. In every example, the operator was able to save the noise type of the selected segments. The interface (see Fig. 1) also allowed: (1) to select the desired lead to identify types of noise; (2) to zoom the detailed beat morphology in a new window of selected segment; (3) to move for and backward inside the same recordings; (4) to select and remove the noise labels in each segment, and (5) to re-scale the  $y$  axis to high amplitude noise.

### 3. Results

The initial description of noise classification in terms of clinical severity was the following:

- Noise free (type 0): segment without noise.
- Low-noise (type 1): noisy segment but P and T waves and the QRS complexes are readable.
- Moderated-noise (type 2): noisy segment in which only the QRS complexes are reliably identified.
- Hard-noise (type 3): noisy segment where the QRS com-

plexes were not clearly recognizable.

- Other-noise (type 9): calibration pulses and straight lines due to complete absence of signal or amplifier saturation.

Figure 2 shows several typical examples of these noise types. In the type 1 example, the SNR is probably low, but P and T waves can be readily recognized. In the type 2, the SNR is likely high if we quantify it in terms of signal and noise power, however, it has an impact on the waveform which makes hard to retrieve P and T waves, specially in some of the beats. In the type 3 example, the noise is strong enough to hide even the QRS complexes.

Changes from one to other type of noise are frequent, may be smooth or sharp, and the transitions are not always easy to be automatically identified. Figure 1 shows three segments of recorded ECGs from 2 patients with different types of noise one after the other. Again, in the type 1 noise at the upper panel (in red), the ECG morphology (P and T waves, PR and QT segments, QRS complexes) can be reasonably evaluated, even though there is some BW (which in this case could be easily filtered out). In the middle panel, we can observe a signal segment with different noise types, which in general should be described by a quantitatively high SNR, whereas we find type 3 hard-noise (in yellow) that prevents QRS recognition, as well as type 1 and type 2 noise. At the bottom panel, a signal segment is shown with continuous type 3 noise, followed by type 2 noise (in green, only QRS readable), and by a low-noise segment (type 1). All the signals were continuously

Table 1. Duration (in seconds) of every type of noise in Lead1 (left) and Lead 2 (right).

Nº pat	Total duration	Free	Low	Moderated	Hard	Other	Free	Low	Moderated	Hard	Other
Pat1	2843,00	1925,46	690,07	178,01	44,44	5,03	2203,58	366,54	240,62	27,24	5,03
Pat2	2433,00	651,64	608,67	687,60	466,37	18,73	655,46	587,58	699,83	471,40	18,73
Pat3	2854,00	441,89	636,78	851,18	911,04	13,13	511,72	314,92	856,54	1127,66	43,17
Pat4	1673,00	790,85	73,93	385,82	417,76	4,65	607,39	70,29	600,52	390,36	4,45
Pat5	4214,06	2142,64	2065,23	0,75	0,24	5,21	1975,19	2232,69	0,75	0,24	5,21
Pat6	2134,00	808,96	504,64	556,36	139,64	124,41	789,07	562,82	516,11	138,53	127,48
Pat7	1172,00	365,05	391,64	363,09	38,36	13,88	413,63	381,29	342,08	21,33	13,68
Pat8	3114,00	500,75	591,52	1023,30	986,23	12,21	471,88	413,52	1203,89	1012,32	12,41

labeled by determining which type of noise was present at each signal segment. Some specific rules were documented, but they are not included in the present work.

With the progress of the analysis the previous definitions were refined and modified to better describe the clinical impact of the noise. First, most arrhythmic events must last a minimum of several seconds to be clinically significant. Thus, at least 3 consecutive beats with the same type of noise are required for classifying a noisy segment. In addition, when doubts between 2 types of noise were present, the general rule was to classify the segment as having the higher level of noise. According to the specific rules and to the previous modifications, we finally defined 4 types of noise, changing the following definitions:

- Type 1: some noise is present, but P and T waves and QRS complexes are readable and their morphology can be ascertained.
- Type 2: Only the QRS complexes can be reliably identified, at least in 3 consecutive beats.

We note that this final classification is closely related to the provided in other references (e.g. [2]). However, in contrast to other work, our database includes a continuous running labeling of all the signals. Therefore, after achieving a uniform criteria and classifying all the EER segments in these types of noise, we obtained a running Noise Typology Gold-Standard. Table 1 shows with detail the duration of each type of noise in each patient and for each lead in the EER. Overall, the duration of different types of noise for Lead 1 was (in hours:minutes:seconds): 2:07:07 (37,32%) for *noise free* (type 0), 1:32:43 (27,22%) for *moderate noise* (type 2), 1:07:26 (19,80%) for *hard noise* (type 3), 50:04 (14,69%) for *low noise* (type 1) and 3:17 (0,97%) for *other-noise* (type 9). In the case of Lead 2 the results were: 2:07:08 (37,32%) for *noise free* (type 0), 1:22:10 (24,12%) for *moderate noise* (type 2), 1:14:20 (21,82%) for *hard noise* (type 3), 53:09 (15,60%) for *low noise* (type 1) and 3:50 (1,14%) for *other-noise* (type 9).

#### 4. Conclusions

Ambulatory Holter recordings are commonly corrupted by noise ECG segments that can interfere with accurate clinical classification. Different to other studies, which

make the analysis by adding artificial noise to a baseline-clean ECG for working with know signal to noise ratio, our approach used a real EER database with noisy and clean recorded segments. The annotated database and the proposed classification are based on the clinical needs for ECG interpretation and may be more useful for testing automatic classifiers than other noise descriptions based just on the level of the SNR estimated by mathematical algorithms.

#### Acknowledgements

This work was partly supported by PRINCIPIAS (TEC2013-48439-C4-2-R, supported with FEDER fundings) and FINALE project (TEC2016-75161-C2-2-R).

#### References

- [1] Agrawal S, Gupta A. Fractal and EMD based removal of baseline wander and powerline interference from ECG signals. *Comput Bio Med* 2013;43(11):1889–1899.
- [2] Li Q, Rajagopalan C, Clifford G. A machine learning approach to multi-level ECG signal quality classification. *Comput Meth Prog Bio* 2014;117:435–447.
- [3] Luo S, Johnston P. A review of electrocardiogram filtering. *J Electrocardiol* 2010;43(6):486–496.
- [4] He H, Tan Y, Wang Y. Optimal base wavelet selection for ECG noise reduction using a comprehensive entropy criterion. *Entropy* 2015;17(9):6093–6109.
- [5] Sivakumar R, Tamilselvi R, Abinaya S. Noise analysis and QRS detection in ECG signals. In *12nd Int Conf Comput Technol Science*. 2012; 141–146.
- [6] Palma Gámiz J, Arribas A, González J, Marin E, Simarro E. Guidelines of the spanish society of cardiology on dynamic electrocardiography and ambulatory blood pressure monitoring. *Rev Esp Cardiol* 2000;53(1):91–109.

Address for correspondence:

Dr. Arcadio García-Alberola  
 Cardiac Service, Arrhythmia Unit.  
 Hospital Clínico Universitario Virgen de la Arrixaca, Ctra. Madrid-Cartagena s/n, 30120 El Palmar, Murcia, Spain.  
 email: arcadi@secardiologia.es