Automatic Classification Normal ECGs Based on Normal PathECG and WaveECG Features

Elzbieta Pociask[1]*, Krzysztof P Malinowski[2,3]*, MHD Jafar Mortada[4], Klaudia K Proniewska[2,3], Peter M van Dam[3,5]

[1] Department of Biocybernetics and Biomedical Engineering, AGH University of Krakow, Poland
[2]Center for Digital Medicine and Robotics, Jagiellonian University Medical College, Krakow, Poland
[3]Department of Bioinformatics and Telemedicine, Jagiellonian University Medical College, Krakow, Poland
[4] Department of Information Engineering, Università Politecnica delle Marche, 60121 Ancona, Italy
[5]Department of Cardiology, University Medical Center Utrecht, The Netherlands

Abstract

Classification of the ECG waveform to normal or abnormal is important to the non-experienced ECG-reader. We propose an algorithm to use solely the waveform of a single ECG beat to classify the ECG as normal or abnormal. In this study we used a subset of the normal classified ECGs from the PTB-XL database to create a normal distribution of the ECG waveform (WaveECG) and its PathECG positions. The aim of this study was to use these distributions to classify all human validated ECGs from the PTB-XL database as either normal or abnormal. Our initial results show an accuracy of 87% to determine whether an ECG is normal or abnormal, irrespective of the gender group used. Using solely the ECG waveform can detect the vast majority of abnormal ECGs, including conduction disorders, ischemia, and arrhythmias.

1. Introduction

ECG interpretation is crucial in cardiovascular disease, so distinguishing normal from abnormal ECG waveforms is required at the earliest possible stage of diagnosis. Despite being one of the oldest diagnostic techniques, it still poses many problems in its correct reading. A recent systematic review and meta-analysis studies show that the median accuracy of ECG interpretation was 54% [1]. This is shows how challenging and difficult task is, and this difficulty is partly due to the inter-individual variability of the ECG, which is influenced by gender, body build and the position of the electrodes during ECG recording. Errors in analysis can lead to misdiagnosis, and less common and more subtle abnormalities can be overlooked, delaying appropriate treatment. Consequently, there is a strong weight on developing tools to expand skills and knowledge in ECG interpretation by healthcare personnel.

To this purpose, the PathECG and WaveECG distribution have been used to facilitate the classification of a resting ECG as normal or abnormal through additional analysis and visualisation of the 12-lead ECG.

2. Materials and Methods

Study was based on open source available ECG database with human validated classifications of the ECG: the PTB-XL database [2] was used to determine the normal distribution of the PathECG and WaveECG parameters, and to determine the standalone performance. Scheme of study is presented on figure 1.

14,380 12-lead ECGs were selected from the Physionet PTB-XL database, of which 5 408 (3124 women, 2283 men) were labelled as normal controls. A normal control was considered when the SCP code was normal >= 80% and the signal quality was good enough to create a median rhythm from 3 similar waveforms.

 ISSN: 2325-887X DOI: 10.22489/CinC.2023.216

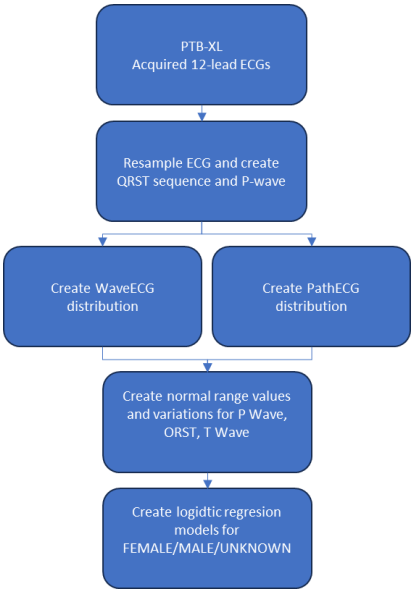

Figure 1 The scheme of data selecting and processing in this study

ECG

As a first step all ECG signals from each leads were resampled to create a QRST sequence of the same length. The length of the QRS complex shortens by only a few milliseconds as the heart rate increases, so the duration of the QRS complex was assumed to be constant for each resting ECG. Resampling was therefore applied only to the STT segment. The ECG was resampled to a normal QT interval. 400 ms was chosen as the reference QT interval length, and each resampled ECG contained three segments of standard duration, as shown in Figure 2.

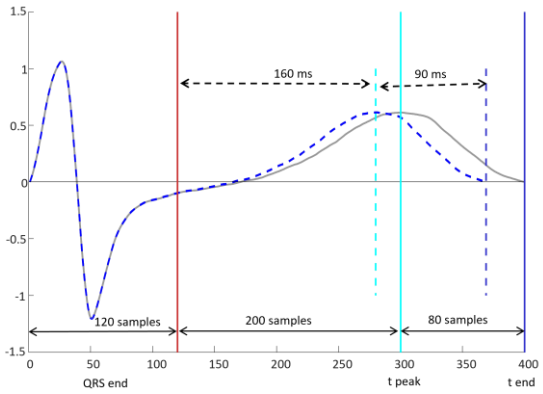

Figure 2 Construction of the ECG amplitude distribution from multiple ECGs. For the amplitude distribution construction all ECGs need to have the same QT time (400 sample of 1 ms).

WaveECG

For each resampled QRST, a distribution between minimum and maximum amplitude (lower and upper limits) was created by plotting each ECG recording one on top of the other.

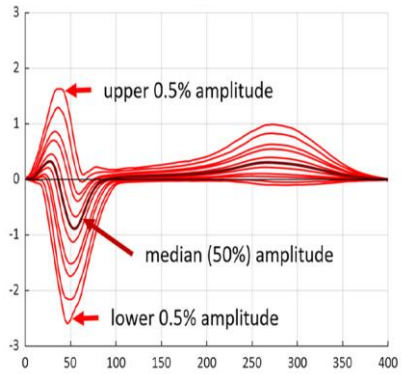

Figure 3 Construction of the ECG amplitude distribution from multiple ECGs [3].

A median beat was automatically constructed by the proposed algorithm [3]. For each median beat, the P-wave onset and end, QRS onset, QRS end, and T wave peak and end fiducial points were automatically determined.

The amplitude contour map thus created was used to determine whether or not the ECG falls within the normal amplitude distribution of the ECG leads (Figure 4).

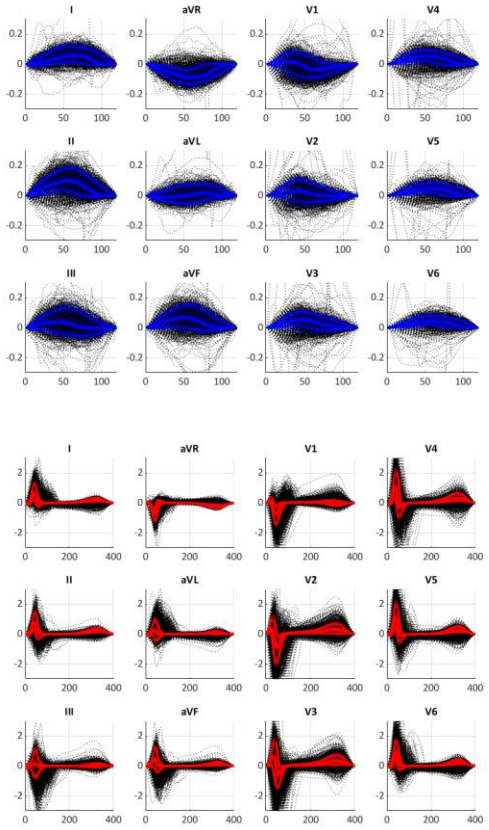

Figure 4 WaveECG Normal distribution on all abnormal P waves and QRST. In blue the distribution of normal males, in red the distribution of normal females, in black the abnormal ECGs.

The Normal ΔWaveECG is when ECG signal is inside the normal amplitude distribution.

Abnormal ECG is :

- when the amplitude is > than the upper limit normal distribution line of the respective ECG lead, in this study > upper 0.5% of the amplitude distribution
- when the amplitude is < than the lower limit normal distribution line of the respective ECG lead, in this study < lower 0.5% of the amplitude distribution

PathECG

The next step was to resampled ECG converted into the vectorcardiogram (VCG), representing the direction of cardiac activity through the heart beat. Furthermore it was used to estimate the mean temporo-spatial isochrone position for the QRST sequence what is well described in [4,5].

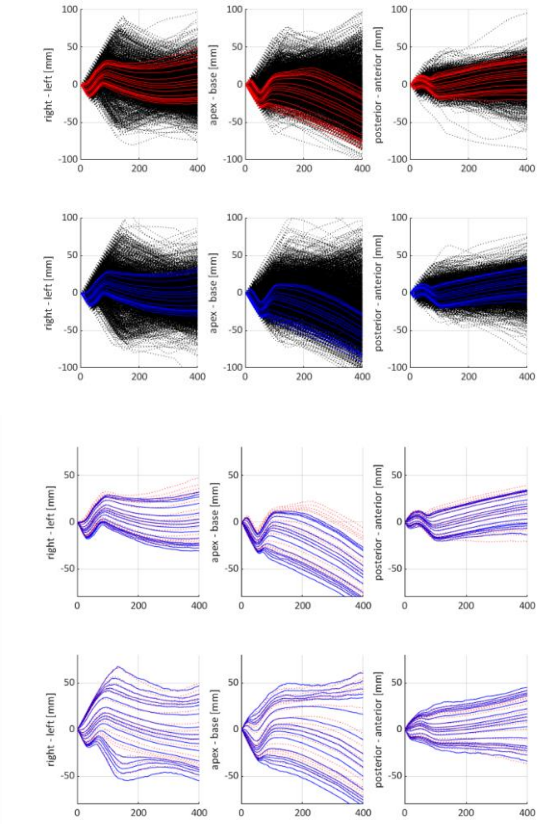

Figure 5 The PathECG distribution of all used PTB-XL ECGs between 0.5-99.5%: in red the distribution of normal females, in blue of normal males , in black the abnormal ECGs.

The X,Y,Z components of the PathECG (Figure 5) are plotted relative to the initial point in the heart in our case in trans-septal position, then move to the right, subsequently back to the left LV and finally the T-wave position moves towards the apex.

Definition of the normal range for WaveECG and PathECG

The *In Normal Range* (*INR*) for either the PathECG (INRP) or WaveECG (INRW) is defined as:

$$INRX(t_{seg}) = \frac{N(xECG(t_{seg})>Normal_{-limit}\ and\ xECG(t_{seg})<Normal_{+limit})}{t_{seg}} 100\% \quad (1)$$

where, $xECG(t_{seg})$ is either the PathECG or Wave ECG of the segment $(t_{seg}), i.e.$ P-wave, QRS, ST segment or T-wave. The subset of the normal range is defined as the full normal range minus a predefined symmetrical applied outlier percentage.

For the P-wave, QRS and T-wave two additional parameters were computed, measuring the variation between the median PathECG or WaveECG.

3. Statistical Analysis

Univariate and multivariate logistic regression was used to evaluate discrimination between normal and abnormal ECG signals for each model, with ROC analysis.

The best threshold was selected as point closest to the top-left corner on the ROC plot, then sensitivity (Se) and specificity (Sp) at this specific threshold were calculated.

Positive and Negative Predictive Values (PPV; NPV) were calculated using bayes theorem to account for prevalence of ECG abnormalities in target population, based on following formulas:

- Positive predictive value is defined as

$$PPV = \frac{Se*\gamma}{Se*\gamma + (1\text{-}Sp)(1\text{-}\gamma)} \quad (2)$$

- Negative predictive value is defined as

$$NPV = \frac{Sp(1-\gamma)}{Sp(1-\gamma)+(1\text{-}Se)*\gamma} \quad (3)$$

where γ is the prevalence of ECG abnormality in target population.

Final regression model was calibrated and validated (internal validation) using bootstrap resampling.

The sample size calculation and data analysis were performed in R version 4.3.1 (R Foundation for Statistical Computing, Vienna, Austria, 2023) with package 'rms' version 6.7-0 and 'pROC' version 1.18.2.

4. Results

The best multivariable predictive models for the QRST and P-wave were created for each combination of patient gender and criteria including all analyzed parameters as covariates. The male QRST ECGs with male criteria combination performs best. For the Female and Unknown

QRST ECGs the Male and Unknown criteria perform rather similar.

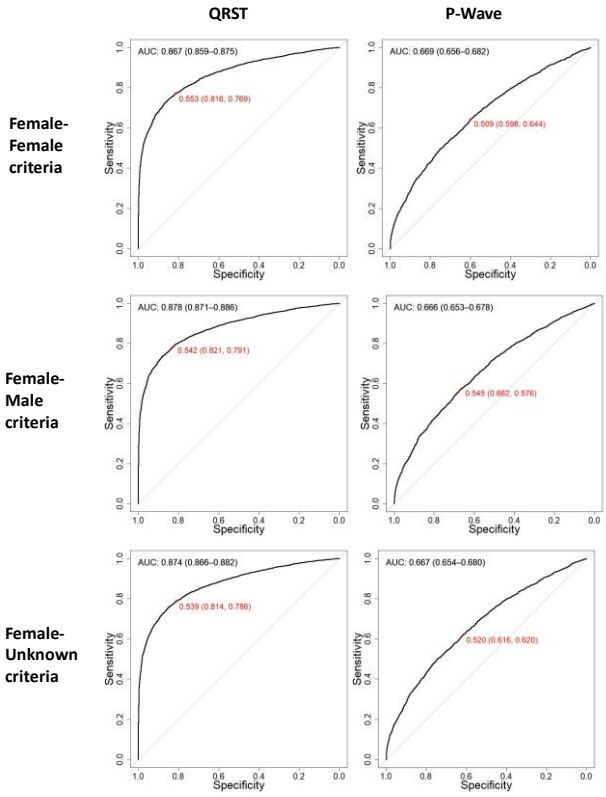

Figure 6 Area under the curve for the discrimination of abnormal / normal QRST waveforms (left) and P-wave waveforms (right) using the logistic model

5. Discussion and Conclusions

Results show that the Wave/PathECG distributions can be used to distinguish between normal and abnormal amplitudes in different ECG segments and detect abnormalities that may not be easily identifiable by the non-ECG expert. Based on results we recommended the use of :

-Male ECGs the male criteria

- QRST: AUC 87.2 (CI: 86.4-88.0)
- P-wave: AUC 65.7 (64.3-67.1)

-Female ECGs the unknown criteria as these criteria also include females, although these female ECGs perform slightly better with male criteria.

- QRST: AUC 87.8 (CI: 87.1-88.6)
- P-wave: AUC 66.6 (CI: 65.3-67.8)

- Unknown (undefined gender) ECGs the unknown criteria as these criteria include both males and females, although these unknown ECGs perform slightly better with male criteria.

- QRST: AUC 87.2 (CI: 86.6-87.7)
- P-wave: AUC 66.2 (CI: 65.3-67.2)

The performance of for the P-wave is significantly lower compared to the QRST. This is majorly caused by the fact that the abnormal ECGs were related to ventricular related heart diseases, like conduction disorders, ischemia, or premature ventricular complexes. For many of these ventricular related abnormal ECGs the atrial P-wave was still normal. Future work will focus on the collection and use of a qualified normal atrial database.

Acknowledgments

Scientific work was supported by the AGH Research Grant 16.16.120.773, and UJCM Research Grant: LIDER/17/0064/L-11/NCBR/2020

References

[1] JAMA Internal Medicine November 2020 Volume 180, Number 11, David A. Cook, MD, MHPE; So-Young Oh, MS, MA; Martin V. Pusic,MD, PhD, Accuracy of Physicians' Electrocardiogram Interpretations A Systematic Review and Meta-analysis, doi: 10.1001/jamainternmed.2020.3989.

[2] Wagner, P., Strodthoff, N., Bousseljot, R., Samek, W., & Schaeffter, T. (2022). PTB-XL, a large publicly available electrocardiography dataset (version 1.0.3). PhysioNet. https://doi.org/10.13026/kfzx-aw45.

[3] Klaudia K. Proniewska, Roger Abächerli, Peter M. van Dam, "The ΔWaveECG: The differences to the normal 12-lead ECG amplitudes", Journal of Electrocardiology, Volume 76, 2023, Pages 45-54, ISSN 0022-0736, https://doi.org/10.1016/j.jelectrocard.2022.10.014.

[4] Boonstra MJ, Brooks DH, Loh P, van Dam PM. CineECG: A novel method to image the average activation sequence in the heart from the 12-lead ECG. Comput Biol Med. 2022 Feb;141:105128. doi: 10.1016/j.compbiomed.2021.105128. Epub 2021 Dec 11. PMID: 34973587.

[5] Peter M. van Dam, Machteld Boonstra, Emanuela T. Locati, Peter Loh, "The relation of 12 lead ECG to the cardiac anatomy: The normal CineECG", Journal of Electrocardiology, Volume 69, Supplement, 2021, Pages 67-74,ISSN0022-0736, https://doi.org/10.1016/j.jelectrocard.2021.07.014.

Address for correspondence:

Elzbieta Pociask
Av. Mickiewicza 30
30-059 Krakow
epociask@agh.edu.pl