Q-onset and T-end delineation: assessment of the performance of an automated method with the use of a reference database

I Christov\(^1\) and I Simova\(^2\)

\(^1\) Centre of Biomedical Engineering, Bulgarian Academy of Sciences, Sofia, Bulgaria
\(^2\) University Hospital ‘Aleksandrovskα’, Clinic of Cardiology, Sofia, Bulgaria

E-mail: Ivaylo.Christov@clbme.bas.bg

Received 16 November 2006, accepted for publication 19 December 2006
Published 12 January 2007
Online at stacks.iop.org/PM/28/213

Abstract
The aim of the study is to assess the performance of an automated method for Q-onset and T-end delineation, as well as QT measurement with the use of a ‘gold standard’ of a manually created reference database. The data to be used comprise 548 recordings of the PTB Diagnostic ECG Database. The ECG signals are preprocessed, suppressing power-line interference, electromyographic noise and baseline drift according to our previously published investigations, guaranteeing accurate Q-onset and T-end locality preservation. Our method for automatic detection of Q-onset and T-end is based on the minimum value of the angle between two segments having a common mid point and equal lengths of 10 ms. The minimum of the angle is searched in defined time intervals delimited separately for the Q-onset and T-end. All measurements are performed on lead II only. Mean ± standard deviations are obtained for 95\% of the recordings: \(-0.08 \pm 2.71\) for Q-onset, \(5.10 \pm 9.22\) for T-end and \(4.40 \pm 9.93\) for QT interval, as well as for 100\% of the recordings: \(0.46 \pm 4.84\) for Q-onset, \(1.28 \pm 16.75\) for T-end and \(0.83 \pm 16.67\) for QT interval.

Keywords: electrocardiography, Q-onset and T-end delineation, QT measurement, PTB Diagnostic ECG database.

(Some figures in this article are in colour only in the electronic version)

1. Introduction

Many studies were conducted linking the changes in QT interval and QT interval dispersion to the risk or prognosis of patients for developing ventricular arrhythmias and fibrillation as well as other cardiac diseases, and eventually for sudden cardiac death. New pre-clinical safety
studies and pharmacological therapy are also associated with the accurate QT prolongation assessments.

Delineation of Q-onset and T-end alone and as part of the measurement of the QT interval is a ‘classical’ problem in quantitative ECG, approached in many different ways, especially after the automation of the process. The investigators have to rely on a reference dataset of markings manually assessed by cardiologists, as suggested and used by the Common Standards of Electrocardiography (CSE) Working Group (The CSE Working Party 1985, Willems et al 1985). An international project consisting of active participants from 20 Institutions of the EC was initiated to overcome the lack of standards, to achieve agreement on wave definitions and measuring protocol equalization (Kors et al 1986). The CSE Working Group used repeated assessments in four rounds: the first three to correct the inter-observers differences and the fourth to correct the common referees’ median with respect to a program derived one.

A dataset of manually measured Q-onsets and T-wave ends fully accessible by the Internet was created by Christov et al (2006) for a selected heartbeat for all of the 458 recordings in the PTB Diagnostic ECG database. More than 6000 manual markings done by 5 experts were collected and analyzed and a reference library was therefore established through a comprehensive, interactive review process, following all the recommendations of the CSE Working Group. The available standard can be used for the development of automated methods for the detection of Q-onsets, T-wave ends and for QT interval measurements.

In comparison with manual methods, the automated ones offer advantages in terms of absolute repeatability of measurements, immunity from errors related to observer fatigue, lapses of attention, as well as efficiency and cost that permit either more extensive and rigorous testing for the same cost as manual methods, or more rapid testing at lower cost (Moody et al 2006). But a question arises: ‘Can the QT interval be measured by fully automated methods with accuracy acceptable for clinical evaluations?’ In an attempt to answer the question PhysioNet/Computers in Cardiology forwarded a challenge in 2006 (Moody et al 2006). Fifteen entrees were submitted in the manual measurements division, and used as a ‘gold standard’ QT reference. There were 25 entries in the automatic measurements division. Typical results presented as the RMS error of the manual measurements were 10–20 ms, while those of the automatic measurements were 20–30 ms.

The challenge did not require that all measurements were performed on one and the same previously selected P-QRT-T interval, which gave way to the following two disadvantages: (i) blocking the opportunity for individual assessment of the Q-onset and T-end delineations accuracy, and (ii) uncertainty of the QT accuracy reported in the challenge due to the interbeat QT interval variations in randomly selected beats. Figure 1 shows two recordings of one and the same patient having an interbeat QT variation of 20 ms. The first ECG trace is at the 6th s and the second one at the 37th s of lead II.

The aim of this study is to assess the performance of an automated method for Q-onset and T-end delineation, as well as QT measurement with the use of a ‘gold standard’ of a manually created reference database.

2. PTB Diagnostic ECG Database

The data to be used comprise 549 recordings of the PTB Diagnostic ECG Database, which was contributed to PhysioNet in September 2004 by its creators Bousseljot et al (1995) and Kreiseler and Bousseljot (1995).

Each of the 549 recordings contains 15 simultaneously acquired signals: conventional 12 leads and 3 Frank (XYZ) leads. All of them are digitized at 1000 samples per second, with 16 bit resolution over a range of ±16.384 mV. The recordings come from
294 subjects (each represented with one to five recordings) within a broad range both of age and diagnosis. About 20% of the subjects are healthy controls. The recordings are typically about 2 min in length, with a small number of shorter recordings (none less than 30 s).

Each ECG recording is accompanied by a detailed clinical summary, including age, gender, diagnosis, and where applicable, data on medical history, medication and interventions, coronary artery pathology, ventriculography, echocardiography, and hemodynamics. Diagnostic classes of the subjects such as coronary artery diseases, heart failure, hypertensive heart disease, rhythm disturbances, etc, are also described.

No ECG-like tracings were observed in the recordings of patient 285/50544.re, and it was excluded from the study, as it has been done by the manually created reference database (Christov et al 2006).

3. Methods


The individual assessment of the Q-onset and T-end delineation accuracy gave the chance of some improvements to the reported method, decreasing the standard deviation for both the Q-onset and the T-end.

3.1. Signal preprocessing

The ECG signals are preprocessed suppressing power-line interference, EMG noise and baseline drift according to our previously published investigations of the Q-onset (Daskalov
and Christov 1999a) and T-end localization (Daskalov and Christov 1999b):

- Moving average of samples in one period of power-line interference. This filter is meant to eliminate the power-line interference. Its frequency response is having a first zero at interference frequency 50 Hz (60 Hz);
- A smoothing procedure for EMG noise suppression is applied (Daskalov and Christov 1999b, Christov and Daskalov 1999). It uses the least-squares approximation method, applied for defining the weighting coefficients. The mathematical description of the process is:

\[ Y_i = \frac{1}{N} \sum_{j=-n}^{j=n} C_j X_{i+j}, \]

where \( Y \) and \( X \) represent the signal after and before approximation respectively, \( n \) is the length of the approximation interval at both sides of a sample, \( C_j \) are weighted approximation coefficients and \( N \) is a normalization coefficient. The procedure is applied on \( 2n+1 \) samples. We are working with an approximation interval of 31 ms. The approximation coefficients are:

\[ C_j = 3n^2 + 3n - 1 - 5j^2, \]

and the normalization coefficient is

\[ N = \frac{(2n+1)(4n^2 + 4n - 3)}{3}. \]

The analysis of the approximation method for the EMG noise suppression shows that although the procedure is quite effective in obtaining high noise immunity, it expands in time-axis the QRS complex. The Q-onset and offset shifts depend on the frequency spectrum of the complex and on the choice of the approximation interval. The higher frequency spectrum of the complex and wider approximation interval result in a bigger shift, and vice versa. With \( n = 15 \) (an approximation interval of 31 ms) the Q-onset shift is 1–3 ms to the left (figure 2). These results were further used to correct the Q-onset detection. No T-end shifts were observed due to the lower frequency spectrum of this wave.

- A high-pass recursive filter for drift suppression (Daskalov et al 1998). The phase characteristic of this filter is constant and the phase distortions introduced in the forward
time direction are cancelled by a second-pass backward application. The high-pass recursive filter is given by the formula:

\[ Y_n = C_1(X_n - X_{n-1}) + C_2Y_{n-1}, \]

where \( Y_n \) is the filtered samples sequence, \( X_n \) is the samples sequence of the original signal and \( n \) is the consecutive number of samples. The constants \( C_1 \) and \( C_2 \) are calculated by the formulae:

\[ C_1 = \frac{1}{1 + \tan(Fc\pi T)}, \quad C_2 = \frac{1 - \tan(Fc\pi T)}{1 + \tan(Fc\pi T)}, \]

where \( T \) is the sampling period and \( Fc = 0.64 \) Hz is the chosen cut-off frequency.

3.2. Delineation of the time interval for Q-onset search

An 'isoelectric' (flat or of low slope) segment is searched in the interval from the biggest peak of the complex (QRS_p, figure 3(a)) to 120 ms backwards on the time axis. The segment is found if all successive differences in 20 ms interval between adjacent samples are less than a preset value Crit and the difference between the end-samples of the 20 ms interval is less than \( 4 \times \text{Crit} \). The value of the Crit is dependent on the QRS magnitude:

\[ \text{Crit} = 0.02(\max\text{QRS} - \min\text{QRS}). \]

The leftmost sample of this segment (\( Q_L \), figure 3(a)) is set as the leftmost point of the searched time interval.

The rightmost point of the searched interval (\( Q_R \)) is found if a peak or a slope (whichever occurs first) is detected to the right of \( Q_L \). Looking for a peak we analyze three samples separated by 10 ms. Differences between the middle and the two adjacent ones are made. A peak is found if both differences have same sign and if they are greater than \( 3 \times \text{Crit} \). Detection of a slope is done by analysis of nine samples, separated by 2 ms. Differences between successive samples are formed. A slope is found if the eight differences have same sign and their absolute values are greater than \( 4 \times \text{Crit} \). The midpoint of the slope or the peak is set as the rightmost point of the searched interval.

3.3. Delineation of the time interval for T-wave end search

The QRS-offset point (\( J \), figure 3(b)) is searched to the right of the QRS_p, repeating the described above criteria for \( Q_L \) search.
Two adjacent segments forming ‘wings’ are defined, each segment being of 40 ms length:

\[ W_1 = D_i_{-40\, \text{ms}} - D_i \quad W_2 = D_i - D_i_{+40\, \text{ms}} \]

where \( D \) are the corresponding signal samples.

The ‘wings’ function \( W = W_1 \times W_2 \) in the interval from \( J \) to \( J + QTC - 100 \, \text{ms} \) is shown in figure 3(b) (lower trace). \( QTC \) is calculated by the well-known equation of Bazett. The minimum of ‘wings’ corresponds to the T-wave peak \( T_p \), no matter if the T-wave has a positive or negative direction.

The steepest slope is searched as a maximum of the \( W \) in the interval from \( T_p \) to \( T_p + 0.2 \times QTC \).

The right sample of the search interval \( T_R \) (figure 3(b)) is sought as an absolute minimum of the \( W \) in the interval from the point of the steepest slope to \( T_R + 0.2 \times QTC \).

The left sample of the search interval \( T_L \) (figure 3(b)) is obtained as a point where the amplitude of the T-wave is \( 0.8(T_p - T_R) \).

3.4. Q-onset and T-wave end detection

Our method for automatic detection of Q-onset and T-end (figures 3(a) and (b)) is based on the minimum value of the angle between two segments having a common mid point and equal lengths of 10 ms. The minimum of the angle is searched in the defined time intervals delineated separately for the Q-onset and T-end.

All Q-onsets were corrected with 2 ms in the rightward direction following the approximation procedure analysis made in chapter 3.1.

In our algorithm if no T-wave in lead II can be observed or its magnitude is less than 0.06 mV (126 cases, or 22.95% of all the recordings), we localize the search interval in the precordial lead V2, and then perform the T-end measurement in lead II. The reference dataset (Christov et al 2006) was approached the same way. In cases where no T-wave could be observed or its amplitude was very small, the referees of the working team were instructed to mark a ‘group-T-wave-end’ taking into consideration leads with better manifested T-wave.

4. Results

548 recordings of the PTB Diagnostic ECG Database have been processed. The automatic Q-onset and T-end delineation was performed only on lead II, and on the same heart beat, as chosen by the reference dataset (Christov et al 2006).

Hnatkova et al (2006) are excluding ‘nonmeasurable’ Q-offset and T-end detections in the cases of too noise-polluted ECGs, flat T-waves and T-waves with fussing U wave patterns. Moody et al (2006) are approaching the same way, requiring only 95% of the recordings for a steady-working automatic algorithm. For that reason mean and standard deviation of the automated method, calculated according to the ‘gold standard’ of the reference dataset of Q-offsets, T-ends and QT intervals (Christov et al 2006) are given in table 1 for 95% and for 100% of all the recordings.

Histograms of deviations between the markings of the automatic algorithm compared with the ‘gold standard’ of manually measured Q-onsets, T-ends and QT intervals are presented in figures 4, 5 and 6, respectively. Subplots (a) for all of the figures are for 95% of the recordings, while subplots (b) are for 100% of the recordings. The mean deviation of the algorithm is depicted by a small vertical line, and the horizontal bar presents the ±standard deviation.
Figure 4. Histograms of Q-onset deviations between the markings of the automatic algorithm compared with the ‘gold standard’ of manually measured Q-onsents (a) for 95% of the recordings and (b) for 100% of the recordings. The mean deviation of the algorithm \( Q_M \) is depicted by a small vertical line (not seen in the figure, because it is much closer to zero), and the horizontal bar presents the ± standard deviation \( Q_{SD} \).

Figure 5. Histograms of T-end deviations between the markings of the automatic algorithm compared with the ‘gold standard’ of manually measured T-ends (a) for 95% of the recordings and (b) for 100% of the recordings. The mean deviation of the algorithm \( T_M \) is depicted by a small vertical line, and the horizontal bar presents the ± standard deviation \( T_{SD} \).

Table 1. Mean ± standard deviation of the automatic algorithm compared with the ‘gold standard’ of manually measured Q-onsents, T-ends and QT intervals. Results are presented for 95% and for 100% of all the recordings.

<table>
<thead>
<tr>
<th></th>
<th>95% of recordings</th>
<th>100% of recordings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q-onset</td>
<td>-0.08 ± 2.71</td>
<td>0.46 ± 4.84</td>
</tr>
<tr>
<td>T-end</td>
<td>5.10 ± 9.22</td>
<td>1.28 ± 16.75</td>
</tr>
<tr>
<td>QT interval</td>
<td>4.40 ± 9.93</td>
<td>0.83 ± 16.67</td>
</tr>
</tbody>
</table>

5. Discussion

The deviations of the method working with 95% of the recordings are almost two times better compared with the results obtained with 100% of the recordings (table 1). This demonstrates
that a common standard for assessment and report of the methods’ efficiency is required. Different algorithms are facing common difficulties, such as the noise accompanying the ECGs, the low magnitude of the T-wave, T having bidirectional waveform, fusing U waves, etc. But the algorithms’ performances diverge on the level of these difficulties. Our opinion is that a more standardized approach must be applied requiring certain best performing percentage of all recordings. 95% is an acceptable choice due to the fact that the QT measurement Challenge (Moody et al 2006) has already been a fact and the attending researchers have already reported their best 95%.

The presented method is performing better with QT mean ± standard deviation of 4.40 ± 9.93 ms than the reported version Christov and Simova (2006) of 3.86 ± 12.52 ms. This is partly due to the improvement of the algorithm, the rest comes from the direct comparison with the same P-QRS-T interval of the reference ‘golden standard’, thus eliminating the interbeat QT interval variations.

Hnatakova et al (2006) have tested the ‘new’ and ‘old’ version of the ‘12SL ECG’ algorithm by the GE Healthcare (Milwaukee, WI, USA) with two sets of recordings A and B consisting of 15 194 and 29 866 10 s ECGs respectively. The two sets differ significantly in quality with set B being substantially more noise polluted. Our method performs far better than the ‘old’ version of 12SL ECG 0.51 ± 12.41 ms for set A, and −0.17 ± 14.89 ms for set B, and slightly worse than their ‘new’ 3.95 ± 5.50 ms for set A, and −2.41 ± 9.47 ms for set B. It must be mentioned that the 12SL ECG algorithm measures a ‘group QT’, using all the 8 preliminary leads from which the Standard 12-leads can be obtained, which is far more easier than a single lead QT measurement (Davey 2000, Denes 1992). Furthermore, both A and B sets consist of ‘ECGs of healthy subjects, uninfluenced by drug treatment and by drugs that are known to prolong QT interval’.

6. Conclusion

It can definitely be said that some automated methods possess acceptable accuracy for clinical evaluations. Furthermore, combining objectively the strengths of varied approaches, as is done in the ‘Meta-6’ algorithm (Moody et al 2006), accuracy close to the experts’ measurements can be obtained.
References


Chesnokov Y C, Nerukh D and Glen R C 2006 Individually adaptable automatic QT detector Comp. Card. 33 337–40


Christov I I and Simova I I 2006 Fully automated method for QT interval measurement in ECG Comp. Card. 33 321–4


Davey P P 2000 Which lead for Q-T interval measurements? Cardiology 94 159–64

Denes P 1992 The importance of derived 12-lead electrocardiography in the interpretation of arrhythmias detected by Holter recording Am. Heart J. 124 905–11

Hayn D, Kollmann A and Schreier G 2006 Automated QT interval measurement from multilead ECG signals Comp. Card. 33 381–4

Hnatkova K, Gang Y, Batchvarov V N and Malik M 2006 Precision of QT interval measurement by advanced electrocardiographic equipment Pacing Clin. Electrophysiol. 29 1277–84


Willems J L et al 1985 Assessment of the performance of electrocardiographic computer programs with the use of a reference data base Circulation 71 523–34

Xue J Q 2006 QT interval measurement: what can we really expect? Comp. Card. 33 385–8