Algorithmic testing for QT interval measurement

G Bortolan

Institute of Biomedical Engineering ISIB - CNR, Padova, Italy

Abstract

The PTB diagnostic ECG database has been used for testing the fully automated program for the identification and measurement of QT interval. The database consists of 549 ECG recordings of 294 subjects. The testing algorithm uses the Padova program, which identifies all QRS complexes analyzing a global spatial velocity and the ECG signal. Different strategies for the choice of the optimal QT interval have been compared, and the optimization of T end in the single leads was performed using the well annotated CSE database.

This paper indicates that the more reliable and accurate estimates of the QT interval is the median value of all the measurements from the analysis of the 12 ECG leads (QT global). In addition, the Challenge and the public PTB database reveal their potentiality for the improvement of ECG classifiers and in particular of QT estimators. However, for clinical evaluation it is necessary to have a more consistent "gold standard".

1. Introduction

The recent adoption of ICH E14 [1] by the US FDA, the EU's European Medicines Agency and other National Health Services has drawn attention to the problem of the acceptability for clinical evaluation of QT interval measurements by fully automated methods. In the framework of the facilities of PhysioNet [2], a public service of the Research Resource for Complex Physiologic Signal, the PTB ECG diagnostic database [3,4] is a valuable resource for testing computerized ECG systems. The PhysioNet/Computers in Cardiology Challenge 2006 has the merit of encouraging the investigation of the task of QT interval estimators with the use of the PTB Database [5].

The Challenge includes separate divisions for participants using manual and semi-automated methods, and fully automated methods. Quantitative comparison was achieved by a score system based on a golden standard defined as the medians of the QT measurements of all the Challenge participants.

In the present work, fully automated algorithms for QT interval measurement are investigated. Conditions for a stable identification of T waves are taken into account. The optimisation and validation process of particular threshold values were performed with the use of the well annotated CSE database.

2. Methods and material

2.1. PTB ECG database

The PTB (National Metrology Institute of Germany) diagnostic ECG database [3,4] has been used for the algorithmic testing and improvement for the fully automated identification and measurement of QT interval, in the framework of PhysioNet/Computers in Cardiology 2006 Challenge [5]. It was available in the PhysioNet Database, a public service of the Research Resource for Complex Physiologic Signals [2].

The database consists of 549 ECG recordings of 294 subjects. The ECGs were collected at the Department of Cardiology of University Clinic Benjamin Franklin in Berlin, Germany and it includes 54 healthy volunteers, 148 patients with myocardial infarction and 64 patients with other heart diseases.

Each record contains the 12 conventional leads and 3 frank leads ECGs, 1000 samples per second with 16 bit of resolution, for a period of time ranging from 30 seconds to 2 minutes.

2.2. The CSE database

The process of optimizing and testing the various algorithms for detecting the T-end in the single leads was performed using the CSE (Common Standards for Computerized Electrocardiography) ECG database [6,7]. This is a well annotated reference database for ECG measurement. The golden standard has been derived by an international group of cardiologists (referees), who have visually determined the onset and offset points of P

QRS and T waves. In particular they have determined the QRS indications for every lead, and the P and T wave indications for the four lead groups:

I : D1, D2, D3 III: V1, V2, V3
II: VR, VL, VF IV: V4, V5, V6

The performance of the different algorithms was validated using as reference points the lead group indications of T-end [7,8].

2.3. The score system

The PhysioNet/Computers in Cardiology Challenge 2006 provides a score system in order to valuate and compare the behavior of the various methods [5]. This score is computed by the coordinating center as the RMS QT error in milliseconds (with respect to the median of all partecipants) divided by the fraction of records measured by the particular program. After the submission deadline (September 4th) final scores were sent to all partecipants. In the same time the final reference QT measurements (the medians of the measurements contained in the final entries of the 15 participants in division 1) that were used as the basis for calculating the final scores were available in the PhysioNet public service. For this reason, all the scores reported in this paper are computed with respect to the final reference QT measurements.

3. Wave identification strategies

The fully automated algorithm for the identification and measurement of QT intervals is the Padova program [10] using an ECG signal of 500 Hz (even samples). It identifies all QRS complexes and measures all QT intervals analyzing a global spatial velocity and the ECG signal.

First, the computation of the overall QT-duration considering all the 12 leads is performed, analyzing a filtered spatial velocity with adaptive threshold levels. The QT-global is computed considering the median of all QT values in the considered ECG interval.

Then, the algorithm performs a T-end adjustment in every single lead, considering a threshold level on the derivative of the ECG signal. The optimization of this threshold value was performed using the CSE ECG database. The optimal threshold was chosen as 20 µV/ms.

In Figure 1. this process is illustrated, where the end of the T wave of the lead group 1 given by the CSE referees, the global T-end and its refinement on the three leads D1, D2, D3 are reported.

Two different strategies for the choice of the optimal QT interval have been compared and tested:

a) identification of the “most stable” representative beat for the identification and measurement of a single QT interval (QT-single)
b) identification of the QT intervals in all beats and computation of the median value (QT-global).

In the first case the representative beat was selected considering the largest interval with smaller differences between two consecutive RR intervals. In the second

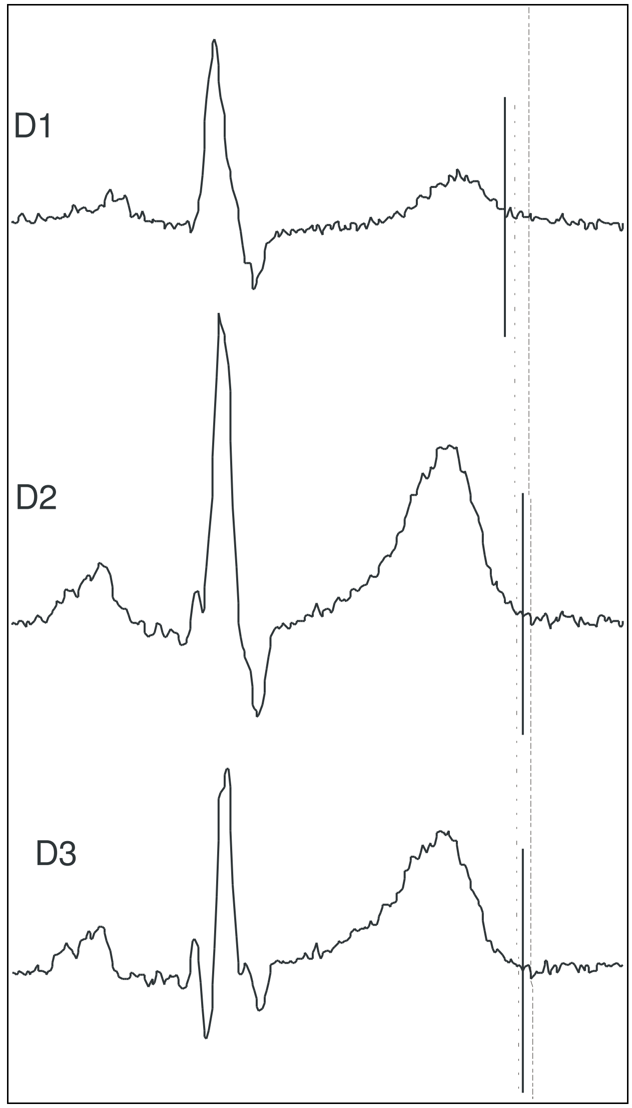

Fig 1. Example of the recognition process of T-end in one ECG signal of the CSE database: indication of the end of the T wave in the lead group I given by the referees (dashed line), the global T-end computed by the program (dotted line) and the refinement T-end in the single leads (solid line).

case, the global QT interval was computed and reported in the representative beat previously identified.

These two independent computation methods are then compared in the successive tests, and they are both submitted to the Challenge as independent entries.

A further test was performed using the information obtained in single leads, and in particular the indications of D2 were considered.

The influence of the duration of the ECG analysis was tested considering the intervals of the first 10 and 30 seconds.

4. Results

The fully automated program analyzed all the 549 records of the PTB ECG diagnostic database, obtaining QT measurements in 522 records.

The first experiment considers the measurements from the analysis of all 12 ECG leads. Two QT estimators are evaluated:

- QT-single
- QT-global

and two durations of the ECG signal have been considered:

- the first 10 seconds
- the first 30 seconds

Considering 30 seconds, the QT global showed 8.9% lower score than the QT-single, and 7.8% in the case of 10 seconds. On the other hand, the use of 30 seconds has a better performance with respect to 10 seconds of 8.5% and 7.4% in the case of QT-global and QT-single respectively.

Then these experiments were performed considering the T-end refinement in lead D2, the suggested lead for submitting the measurements for the scoring system. In this case the QT global shows a 5.4% and 5.1% lower score than the QT-single with 30 and 10 seconds respectively. Similarly 30 s has a better performance than 10 s: 6.4% and 6.2% in the case of QT-global and QT-single respectively. But the analysis of lead D2 produced higher scores. In fact the QT estimates with 12 ECG leads has a lower score from 27.0% to 30.7% with respect to lead D2. These experiments are reported in Fig. 2.

Then, the T-end estimates in various single leads were considered. In particular the leads D1, D2, VF, V1, V2 and V5 are studied, and the results are reported in Fig. 3. From this figure, it is clear that the QT-global performs better than QT-single in all the 6 considered leads. In addition, lead V1 is the one with the lowest score.

From these results the following conclusions may be drawn:

- the increasing of the duration of the analysis of the ECG signal improves the scores
- the better results are obtained by considering all the 12 ECG leads in the QT computation
- the QT-global estimate is always better than the a single measurement in the most "stable" RR interval.

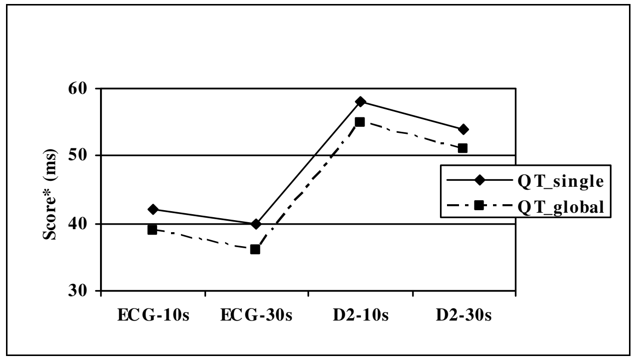

Fig. 2 Scores of different QT estimators, considering 12 leads (ECG) and one lead (D2) analyzing an interval of ECG signal of 10 or 30 seconds, and considering QT-global and QT-single.

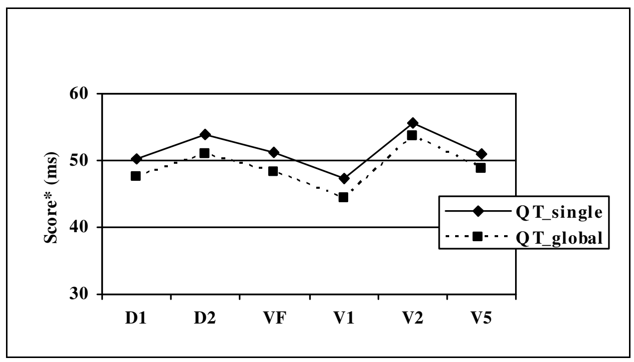

Fig. 3. Scores of different QT estimators, considering the T-end refinement in the single leads D1, D2, VF, V1, V2, V5 , and considering QT-global and QT-single.

The best results are obtained considering all 12 leads from the analysis of 30 seconds of ECG signal. In this case the results of the score system are the following:

Score(QT-global) = 36.5

Score(QT-single) = 40.9

In the case of the best result, the histogram of the differences between the program and the median of all the programs participating to the Challenge is considered and reported in Fig. 4. This histogram suggests the presence of a systematic error of about 25 ms. In fact computing the score with an artificial correction of the QT measurements of 25 ms, a significant improvement is observed:

Score (program) = 36.5

Score (program -25 ms) = 23.3

This fact may suggest the necessity of a more deep analysis of the differences with the median values of all the participants to the Challenge.

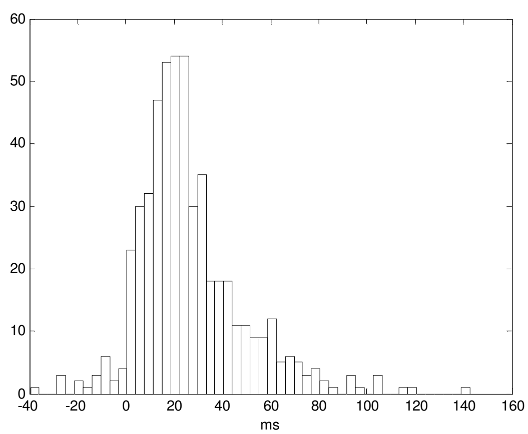

Fig. 4. Histogram of the differences (milliseconds) between the best QT estimator (ECG 30s) and the median of all the Challenge participants.

5. Discussion and conclusions

The PTB diagnostic ECG database has been used for testing the fully automated program for the identification and measurement of the QT interval. All the 549 ECG recordings were analyzed by the Padova Program. Different strategies for the choice of the optimal QT interval have been compared.

This paper indicates that the more reliable and accurate estimates of the QT interval is the median value of all the measurements from the analysis of the 12 ECG leads (QT global), with a score of 36.5 ms. From this study, it is evident the potentiality of the public PTB database and the Challenge for the improvement of QT estimators. However, for clinical evaluation it is necessary to have a more consistent “gold standard”.

References

[1] ICH. E14 Clinical Evaluation of QT/QTc Interval Prolongation and proarrhythmic Potential for Non-Antiarrhythmic Drugs. 2005. http:www.fda.gov/cberg/ gdlns/ice14qtc.html

[2] Goldberger AL, Amaral LAN, Glass L, Hausdorff JM, Ivanov PC, Mark RG, Mietus JE, Moody GB, Peng CK, Stanley HE. PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals. Circulation 2000 (June 13);101(23):e215–e220. Circulation Electronic Pages: http://circ.ahajournals.org/c gi/ content/ full/101/23/e215

[3] Bousseljot R, Kreiseler D, Schnabel A. Nutzung der EKGSignaldatenbank CARDIODAT der PTB ¨uber das Internet. Biomedizinische Technik 1995;40 Erg¨anzungsband I:S 317.

[4] Kreiseler D, Bousseljot R. Automatisierte EKG-Auswertung mit Hilfe der EKG-Signaldatenbank CARDIODAT der PTB. Biomedizinische Technik 1995;40 Erg¨anzungsband I:S 319.

[5] Moody GB, Koch H, Steinhoff U. The PhysioNet / Computers in Cardiology Challenge 2006: QT Interval Measurement. Computers in Cardiology (33) 2006.

[6] Willems J. L., Arnaud P., van Bemmel J. H., Bourdillon P. J., Degani R., et al. Establishment of a reference library for evaluating computer ECG measurement programs. Comput. Biomed. Research 18:439-457, 1985.

[7] Willems J. L., Arnaud P., van Bemmel J. H., Degani R., Macfarlane P. W., Zywietz C. Common standards for quantitative electrocardiography: goals and main results. Methods of Information in Medicine 29:263-271, 1990.

[8] Bortolan G., Bressan M., Cavaggion C., Fusaro S., for the ILSA Research Group Validation of QT dispersion algorithms and some clinical investigations. In Computers in Cardiology 96, Murray A., Arzbaecher R. (Eds.), IEEE Computer Society, Los Alamitos, CA, 1996, 665-668.

[9] Bortolan G., Bressan M., Cavaggion C., Fusaro S., for the ILSA Research Group Different combination techniques in the computation of the QT dispersion index. In Computers in Cardiology 97, Murray A., Swiryn S. (Eds.), IEEE Computer Society, Los Alamitos, CA, 1997, 669-672.

[10] Degani R., Bortolan G. Methodology of ECG interpretation in the Padova Program. Methods of Information in Medicine 29:386-392, 1990.

Address for correspondence
Giovanni Bortolan
Institute of Biomedical Engineering ISIB - CNR
Corso Stai Uniti, 4
35127 Padova, Italy
E-mail: bortolan@isib.cnr.it