An Open-Source Standard T-Wave Alternans Detector for Benchmarking

A Khaustov[1], S Nemati[2], GD Clifford[2,3]

[1]Institute of Cardiological Technics (Incart), St. Petersburg, Russia
[2]Massachusetts Institute of Technology (MIT), Cambridge, MA 02139, USA
[3]Harvard-MIT Division of Health Sciences and Technology (HST), Cambridge, MA 02142, USA

Abstract

We describe an open source algorithm suite for T-Wave Alternans (TWA) detection and quantification. The software consists of Matlab implementations of the widely used Spectral Method and Modified Moving Average with libraries to read both WFDB and ASCII data under windows and Linux. The software suite can run in both batch mode and with a provided graphical user interface to aid waveform exploration. Our software suite was calibrated using an open source TWA model, described in a partner paper [1] by Clifford and Sameni. For the PhysioNet/CinC Challenge 2008 we obtained a score of 0.881 for the Spectral Method and 0.400 for the MMA method. However, our objective was not to provide the best TWA detector, but rather a basis for detailed discussion of algorithms.

1. Introduction

The ninth annual PhysioNet/Computers in Cardiology Challenge (PCinCC) [2] aims to improve understanding of methods for identification and analysis of T-wave alternans (TWA) in the ECG. The repeatability of commercial TWA detection methods depends on the ability of researchers to know exactly what version of software they are using, or the exact (proprietary) algorithm. Comparisons between clinical trials and different versions of software are therefore difficult. Patents, and often journal articles, do not always present complete details of how a technique is implemented, and reproduction of the technique can be time consuming or impossible. We have therefore produced an open source suite of algorithms that provide implementations of known standard TWA quantification techniques. We hope that this suite will provide a benchmark against which other TWA algorithms can be measured.

2. Methods

The algorithms and metrics chosen for inclusion in our open source suite are available in commercial equipment and used most often by clinicians. Our implementation is based on descriptions found in [3] and [4].

2.1. Pre-filtering

Simple baseline wander filtering was performed. The ECG was downsampled to 32Hz. For a given point, 25 percent of the median points in the window of 1.3 seconds were averaged. The obtained set of values determined the baseline level and were removed (after resampling to the original ECG sampling frequency).

2.2. QRS and T detection

Our software uses the QRS and T-wave detection algorithm of Laguna *et al* [5], available from PhysioNet. For the Challenge data we used both this detector and a commercial detector by St.-Petersburg Institute of Cardiological Technics. The annotations were also manually reviewed and a combined set of annotations was developed. The annotations included Q, S and T-end points.

2.3. Spectral methods

The analysis window uses 128 beats, starting from the first labelled beat in the file. If the analysis determines no significant TWA value was available, the window was moved forward in time by 20 beats and the analysis is repeated until a significant alternans value was found or the end of the recording was reached.

2.3.1. Beat alignment and rejection

Relative timing of points within different repolarization cycles (i.e. successive T waves) is crucial for any TWA algorithm to avoid introducing false TWA. The beat alignment algorithm consists of: 1) selection of the lead where the alignment was performed, 2) a search for basic fiducial points 3) template beat selection and fiducial point adjustment, and 4) validation of beats in each lead.

Peak-to-peak values on ST-T segments (from the S-point to the T-end) among the first eight first beats were calculated in each lead and every lead was assigned the median value of these. Leads were sorted according to these median values. Beat alignment was then performed on the lead with the maximum TWA value. If the alignment procedure described below fails for the chosen lead, the next lead in the sorted list was used until a TWA value is found.

The basic fiducial point for a beat was located on the front (i.e. a segment between a local maximum and minimum) with the maximum amplitude and specific orientation (i.e. up-down or down-up) on Q-S interval. The point between the maximum and minimum with amplitude closest to the middle of the front was picked. A template for alignment was chosen among the first 10% of beats in the analysis window. A beat was accepted as a template if it had acceptable optimal correlations (described below) with more than 90 percent of the beats in the analysis window.

The fiducial point for every beat was picked in the interval around the basic fiducial point so as to maximize the cross-correlation with the template beat. The fiducial point was optimized independently for correlation on Q-S and ST-T segment to account for QT variation. So, every beat had two optimal correlation values. A beat was considered acceptable if correlation on Q-S interval was higher than 0.96 and correlation on ST-T segment was higher than 0.8. Whenever more than 10% of beats were unacceptable the template was rejected and the procedure was repeated for another template candidate. If none of the template candidates succeeded, the procedure moved on to another lead in the sorted list. If the procedure failed for all of the leads the analysis window was considered too noisy or abnormal (in rhythm) and the analysis window was moved forward in time.

If the procedure was successful beat validation was performed independently in each lead (other than the alignment lead). Optimal fiducial points were transferred from the alignment lead to all other leads. In every lead other than the alignment lead validation was performed against all beat averages to account for the situation when the template beat was corrupted by noise in the alignment particular lead. Beats with correlations lower than the values mentioned above were marked as invalid. If there were more than 10 percent of invalid beats the whole lead was marked as invalid for the given analysis interval.

Amplitude- (as well as time-) alignment of the beats was also performed before computing the TWA series (see below) by subtracting the mean value of the Q-S segment. Figure 1 illustrates an example of time-aligned beats.

2.3.2. Alternans series

Alternans series were computed for every lead and every offset in the ST-T segment. The following computation
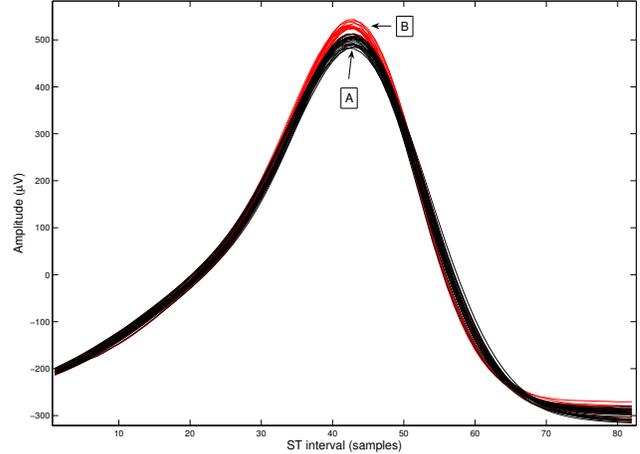
Figure 1. Time-aligned extracted ST-T segments for a single time series of 30 seconds. Note the two different beats types, separated by an average amplitude of 30 μV.

options were available:

• 'Standard' [4]: the alternans series was composed of amplitude values in points with the specified offset within ST-T segment in every beat in the analysis window; two beat averages were calculated - for even and odd valid beats and any invalid beat was replaced with the respective average to compute the amplitude; the series was subject to detrending via removal of best-fit linear trend before the periodogram estimation (see below).

• 'Differences': the alternans series was the sequence of the first differences of the series calculated in the 'standard' technique. In this way a natural detrending of the data is performed and lower frequencies are suppressed.

2.3.3. Periodogram and alternans values

The alternans series was subject to periodogram estimation via windowing with the Hamming window and taking squared absolute value of FFT ([6]).The periodogram was obtained for every offset in the ST-T segment and then averaged over the whole of this segment.

The mean noise level was calculated as the mean value of the squared absolute value of FFT on the interval 0.4-0.46 cycles per beat. (Note that this measure susceptible to intra- and inter-subject heart rate differences.) The alternans voltage was the square root of the difference between value at the frequency of 0.5 and noise mean level. The alternans ratio, A_r, was given as the squared alternans voltage divided by noise standard deviation calculated on squared absolute FFT. The alternans value was considered significant if $A_r > 3$.

If the analysis failed on a data file or there was no significant alternans, the record was assigned a value of 0,

otherwise the record was assigned the value of the maximum TWA voltage among leads where significant alternans were found.

2.4. Modified Moving Average

2.4.1. Initial beat selection

The alignment procedure described in section 2.3.1 was performed for the first 50 beats in the file. The template beat was selected as the initial beat for even beat averaging. The first valid beat (i.e. with a correlations above the thresholds given in section 2.3.1) having an index of different parity was selected as the initial beat for odd beat averaging. The alignment procedure ensured the selection of a normal beat with an acceptable noise level.

2.4.2. Beat averaging

Beat averages were separately calculated for Q-S and ST-T segments. Each beat after the initial beat was subject to fiducial point adjustment described in section 2.3.1. The adjustment versus current average of the corresponding index parity on Q-S and ST-T segments was performed. Whenever any of the two correlation levels obtained after beat adjustment was lower than the given thresholds, the beat was marked invalid. If the beat was valid the respective averages were updated using optimal fiducial points for respective segments and a weighting procedure as follows. For every point on Q-S or ST-T segment the value of the average was altered towards the value of the current beat. The amount by which the average was altered was equal to the difference between the average and the current beat at this point unless this difference was smaller than 1 μV or larger than 32 μV. In the latter case, the TWA amplitude was set to 1 μV or 32 μV respectively. (See [4] for more details.) If there were more than 10% of points on ST-T segment where the difference exceeded 32 μV, the current beat was marked invalid.

2.4.3. Alternans value

The alternans value was calculated every 15 seconds. The number of invalid beats during these 15 seconds was also noted. If more than 10% of beats were marked as invalid, the segment was considered invalid. Otherwise, the alternans value was computed as the maximum absolute difference between odd and even averages at the end of the interval on ST-T segment in all valid leads. The alternans values were then combined on 1 minute intervals (i.e. four 15 second intervals) by taking a minimum value out of the four intervals (with zero corresponding to no valid alternans value). The maximum one-minute alternans value

across the whole file was reported as the alternans value for the file.

3. Data and results

In order to evaluate our algorithms before submitting entries to the challenge, we chose to use a artificial multi-lead ECG model which can exhibit realistic TWA, described in an accompanying paper [1]. Records with TWA amplitudes of 2, 4, 6, 8, 10, 16, 22, 28, 34, 40 and 60 μV were generated. For each TWA amplitude a 'clean' file, files with added white noise (with standard deviation of 5, 10, 20, 30 and 40 μV) and baseline wander (obtained from the MIT-BIH Noise Stress Test Database [7]), were separately generated. Since the TWA normalization issue is unclear (see section 4) the results of all methods were scaled by the ratio of 60 divided by the result obtained for the clean file with TWA level of 60 μV.

All four methods identified TWA level on clean records with accuracy better than 2 μV (maximum value of difference).

On files with white noise SM 'standard' and 'differences' determined TWA with accuracy better than 6 μV and 5 μV respectively if TWA was greater than 0.35κ (where κ is white noise standard deviation) and failed otherwise. On records where SM was successful MMA had accuracy better than 5 μV, however it failed to reject records with lower TWA to noise ratio and overall accuracy was significantly worse. This suggests that TWA to noise level ratio estimation might improve the implementation of MMA.

On files with real noise added all methods had an accuracy better than 7 μV with a general tendency of improving as the TWA amplitude increased.

The real patient data for the challenge was selected by the competition organizers from amongst the databases available on PhysioNet [2], [8]. For the spectral methods we obtained the following scores on the challenge data: 0.880 for the 'standard' option (see section 2.3.2), 0.881 for the 'differences' option (which ranked third in the competition)and 0.400 for the MMA method. When SM was used as a noise detector for MMA (i.e. records with SM zero entries were assigned zero in MMA too) the MMA score improved to 0.834.

4. Discussion

Although the method of the PCinCC allows for an independent voting scheme on the performance of any algorithm, there might exist a global bias in the 'voting' algorithms. Therefore a standard database of TWA data is still required. To avoid this *chicken-and-egg* scenario we developed a model of TWA activity [1]. In future papers we will evaluate the performance of TWA algorithms at vary-

ing noise levels and determine the minimum noise floor for TWA detection. It will be interesting to determine if the reported 'natural' TWA activity of normal subjects of up to $10\,\mu V$ [9] is a real phenomenon, or simply just noise.

The PCinCC provides evidence to support the idea that different algorithmic TWA evaluation implementations give largely different results. Therefore results of a clinical trial can only be safely compared to the results obtained on the exactly same equipment, with similar noise levels. Publication of the actual algorithm's source code will facilitate repeatability.

However, we should note that for now the described algorithms lack important processes which must be applied as pre-processing or post-processing steps:

• The analysis should only be performed within a certain heart rate interval (to be predictive). There are different reports on what this interval should be [4], but it can be anywhere between 90 and 110 BPM [9].
• The periodogram used in the spectral method to compute the alternans values can be normalized in several ways [6], which influences the estimated value of TWA amplitude and therefore interpretation and it is unclear which normalization method should be used.
• There are different criteria for TWA value interpretation; criteria for negative, positive and indeterminate results are necessary.
• Criteria for TWA interpretation often include lead-specific requirements.
• Since TWA manifests with a different magnitude on different leads it is likely that multi-lead data should be rotated to resolve the axis of maximum TWA activity [3].

5.   Conclusions

This article details an open source suite of Matlab algorithms that provides implementations of known standard TWA quantification techniques to provide a benchmark against which other TWA algorithms can be measured. The software suite can run on labelled or unlabelled ECGs in both batch mode and with a provided graphical user interface to aid waveform exploration. The accuracies of the algorithms have been evaluated on both artificial and real ECG data.

The algorithms have been released under the GNU Public License (GPL) in the hope that others will improve and build upon these routines. In particular, we hope that inventors, manufacturers and other researchers will notice misinterpretations of the commercially available algorithms we attempted to mimic and suggest/make modifications for/to our code.

References

[1] Clifford GD, Nemati S, Sameni R. An artificial multi-channel model for generating abnormal electrocardiographic rhythms. Comput in Cardiol September 2008;35. In Press.

[2] Moody GB. The Physionet / Computers in Cardiology Challenge 2008: T-wave alternans. Comput in Cardiol September 2008;35. In Press.

[3] Martínez JP, Olmos S. Methodological principles of T wave alternans analysis: a unified framework. Biomedical Engineering IEEE Transactions on April 2005;52(4):599–613. ISSN 0018-9294.

[4] Narayan SM. Ch 7: Pathophysiology Guided T-Wave Alternans Measurement, volume 1 of Engineering in Medicine and Biology. 1 edition. Norwood, MA, USA: Artech House, October 2006; 196–214. ISBN 1-58053-966-1.

[5] Martinez J, Olmos S, Laguna P. Evaluation of a wavelet-based ECG waveform detector on the QT database. Computers in Cardiology 2000 2000;27:81–84.

[6] Harris F. On the use of windows for harmonic analysis with the discrete fourier transform. Proc IEEEE Jan 1978;66(1).

[7] Moody G, Muldrow W, Mark R. A noise stress test for arrhythmia detectors. In Computers in Cardiology. 1984; 381–384.

[8] 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;101:e215.

[9] Kaufman ES, Mackall JA, Julka B, Drabek C, Rosenbaum DS. Influence of heart rate and sympathetic stimulation on arrhythmogenic T wave alternans. Am J Physiol Heart Circ Physiol 2000;279(3):H1248–1255.

Address for correspondence:

Alexander Khaustov, St. Petersburg Institute of
Cardiological Technics, St. Petersburg, Russia.
email: alexander.khaustov@gmail.com