Analyzing the Delineation Precision of Hannover ECG System (HES®): A Validation Study

A Khawaja[1], J Litwin[2], T Auzinger[2], W O'Rourke[2], T Devine[2], A Furlong[2], C Lehmann[1], R Fischer[1,3]

[1]Biosigna GmbH, Munich, Germany
[2]eResearchTechnology Inc., USA
[3]Department of Biometrics, Medical School of Hanover, Hanover, Germany

Abstract

There is a large interest in analyzing QT/QTc measurements (e.g. in clinical studies during drug safety evaluations), since prolonged QT intervals can cause ventricular tachyarrhythmias or other critical cardiac rhythm events. Furthermore a set of guidelines for clinical evaluation of the QT/QTc interval prolongation (for instance E14) is provided by the ICH. Due to the significance of the QT measurement for drug safety, the objective of this work is to compare the results of manual annotated QT and RR interval measurements with the results of the fully-automated HES® algorithm for validation purposes in clinical studies. In order to enhance the validity of the statistical test and to exclude adaptation between test data and algorithm, the number of test cases is set to more than 42000 annotated ECG signals. The statistical differences between QT and RR interval was observed to demonstrate the ability and reliability of automated methods using the HES® algorithm.

1. Introduction

The measurement of QT/QTc interval prolongation is considered as the standard surrogate biomarker for cardiac drug safety to the ICH clinical evaluation guideline E14 [1]. Manual, semi-automatic and automatic techniques are currently used in localizing the fiducial points of ECG signals and in measuring their corresponding ECG wave intervals including QT and RR intervals. The manual and semi-automatic methods have been employed so far to ensure reliable and precise detection and measurement of these ECG points and intervals, respectively. Accordingly, many efforts have been done to validate the automatic methods against the manual and semi-automatic ones [2–5] and furthermore to compare the output performance of different algorithms for T wave delineation [6]. One of the well-reputed ECG analysis and interpretation programs worldwide is the Hannover ECG System HES® [7]. In order to gain acceptance and trust of the drug safety authorities, validating automated ECG algorithms on individual basis is essential. In this work, the performance and behavior of HES® has been investigated using resting ECG data from real cardiac clinical studies. The validation process was based on comparing the output of HES® with the corresponding output of the manual detection and measurements on representative ECG data used in this study. Over forty-two thousand ECG signals, that were manually measured and segmented, were fed into HES® algorithm. Thereafter, the golden heart beat in each of these ECG signals was automatically delineated. During the whole process, no de-noising method, i.e. no filtering and no averaging, was carried out on the data. Choosing the golden beat in a given signal is based on the specific criteria provided by the sponsor of the corresponding clinical study. Furthermore, the output of HES® algorithm derived QT and RR intervals was compared with the corresponding output of the manual delineation. Finally, the statistical significance of the results derived was observed. The result shows that, the mean differences in QT interval and RR interval are -0.96 [ms] and 2.43 [ms], respectively. Whereas, the median differences are observed to be -2 [ms] and 0 [ms], respectively.

2. Methods

2.1. Used database

The database employed in this study contains 42566 12-lead resting ECG signals from Thorough QT trials. They were sampled with 500 [Hz] and 1000 [Hz]. The duration of each ECG signal is 10 seconds. Since the data have been recorded and annotated during several late-phase clinical studies, they are not available for public use. No ECG signal used for this validation test has been modified or manipulated by de-noising methods. For a given ECG record,

a corresponding set of measurement including QT interval and RR interval is annotated manually by cardiologists or by medical experts. The annotation rules are provided, for each of the cardiac studies, by the corresponding sponsor. The very large number of manually annotated records is a major benefit of this work in comparison to other databases, which had been used in the past for testing QT measurement (e.g. 105 records in [8]). The selection of the respective observed lead has been done during the clinical study according to the corresponding study protocol. The relative lead distribution can be presented as follows:

- Lead II: 94.2%
- V5: 4.5%
- All other leads: 1.3%

2.2. HES® delineation methodology

The delineation process of the Hannover ECG System (HES®) consists of four main steps. These stages are illustrated in figure 1.

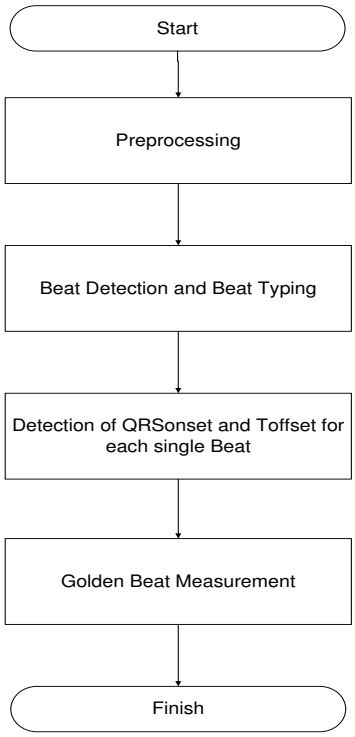

Figure 1. The workflow diagram for the delineation process of HES®.

2.2.1. Preprocessing

In this step, the ECG signal under study is converted to 500 [Hz], the default sampling frequency that HES® can analyze, without carrying out any kind of denoising processes. That is, no filter is used for power-line interference, muscle artifact, distortion from misconnected electrodes, moving artifact or baseline wandering. Afterward, the 12-channel ECG signal is fed into resting HES® for further analysis. The observed lead was determined by specialists during past clinical studies.

2.2.2. Beat detection and beat typing

All valid channels are used to detect the ventricular activities in the input ECG signal and their types. This process is carried out on the output signal of the preprocessing step. QRS complexes and their corresponding Rmax, the time stamp of maximal QRS amplitude, are localized by carrying out a so-called spacial velocity method and spacial velocity signal SV for QRS slope detection and employing a specific set of adaptive beat thresholds. SV is defined as the rate of change in an input signal (ECG signal in this case) with respect to time. In a given time window in the input signal under study $t_1, t_2, ..., t_n$ and $y_1, y_2, ..., y_n$ can be considered as a set of time instances and their corresponding sample values, respectively. In this case, SV for a specific sample k, $SV(k)$, can be derived as follows:

$$SV(k) = \frac{\frac{1}{n}\sum_{i=1}^{n}(t_i - \overline{t})(y_i - \overline{y})}{\frac{1}{n}(t_i - \overline{t})^2} \quad (1)$$

where $\overline{t}$ and $\overline{y}$ are the mean values of all time instances $t_1, t_2, ..., t_n$ and all the corresponding sample values $y_1, y_2, ..., y_n$, respectively. The classification algorithm for the heart beats implies a complex decision tree model with number of independent variables and morphological features as well as additional variables like RR interval and heart rate. thereafter, all dominant intrinsic and normal heart beats of each channel are identified for further analysis.

2.2.3. Detection of QRS onset for each beat

In this stage, the onset of each QRS complex, whose location is already detected in last step, is delineated. Searching for QRS onset is taking place within a specific time interval. The boundaries of this interval are chosen based on number of certain criteria, e.g. the age of the patient. QRS onset is localized by using spacial velocity method and number of adaptive thresholds.

2.2.4. Detection of T-wave offset for each beat

After delineating the QRS onset for each of the detected beats in the last step, the T-wave offset is localized in this stage. Searching for the T-wave offset is carried out within

a certain time interval and boundaries around the end of the T-wave. Like in QRS onset detection, the limits of the searching interval are calculated using a number of criteria. The detection procedure of T-wave offset is based on number of features and independent variables including the following:

- The variability of the special velocity signal within the searching interval
- The amplitude of QRS onset
- An initial estimation of QT interval
- The time stamp and amplitude of T-wave apex
- The duration of QRS complex
- The level of noise signal and baseline wander

2.2.5. Golden beat measurement

Choosing the golden beat measurement in a given signal is based on the specific criteria provided by the sponsor in a given clinical study. In this work, the golden measurement is considered as the average measurements for three consecutive beats with dominant type.

2.3. Validation

The measurements obtained from the fully-automated HES® ECG program, namely QT and RR intervals were compared with reference measurements annotated by the experts on the used database. The results of the validation process are represented in this work by the differences and absolute differences between the results of these two delineation techniques as well as the mean $Mean$, standard deviation STD and median $Median$ of these differences for all records under study.

3. Results

3.1. QT interval measurement

The $Mean$, standard deviation STD and $Median$ values for the differences and absolute differences between the results of the fully-automated HES® program and the manual reference technique for QT interval using the used databank are illustrated in table 1.

The mean of the QT difference between HES® and reference value is very low, since it is smaller than the mean signal digitization error, i.e 1 [ms] and since the median QT difference is identical to the sampling interval, i.e 2 [ms]. The distribution of the absolute differences for QT interval is shown in table 2, while figure 2 depicts the large range of the relation between measurement difference and reference QT interval which is represented in a Bland-Altman plot.

Table 1. The results for the differences and absolute differences between the results of the fully-automated HES® program and the manual reference technique for QT interval.

Mean (ms)	Median (ms)	STD (ms)
-0.96	-2	11.1
Absolute mean (ms)	**Absolute median (ms)**	**Absolute STD (ms)**
7.93	6	7.77

Table 2. The distribution of the absolute differences between HES® and reference value for QT interval.

Absolute difference (ms)	Number of records	Relative frequency (%)	Cumulative frequency (%)
0 - 5	18966	44.56	44.56
6 - 10	12740	29.93	74.50
11 - 15	6579	15.46	89.95
16 - 20	2429	5.71	95.66
21 - 30	1107	2.60	98.26
31 - 40	320	0.75	99.01
41 - 50	207	0.49	99.50
51 - 60	101	0.24	99.74
61 +	112	0.26	100.00

3.2. RR interval measurement

The $Mean$, standard deviation STD and median $Median$ values for the differences and absolute differences between the results of the fully-automated HES® program and the manual reference technique for RR interval using the used databank are illustrated in table 3.

Table 3. The results for the differences and absolute differences between the results of the fully-automated HES® program and the manual reference technique for RR interval.

Mean (ms)	Median (ms)	STD (ms)
2.43	0	33.1
Absolute mean (ms)	**Absolute median (ms)**	**Absolute STD (ms)**
16.6	6	28.72

The mean of the RR difference between HES® and reference value is to be considered low as well, since it is around the sampling interval of 2 [ms]. The distribution of the absolute differences for RR interval is shown in table

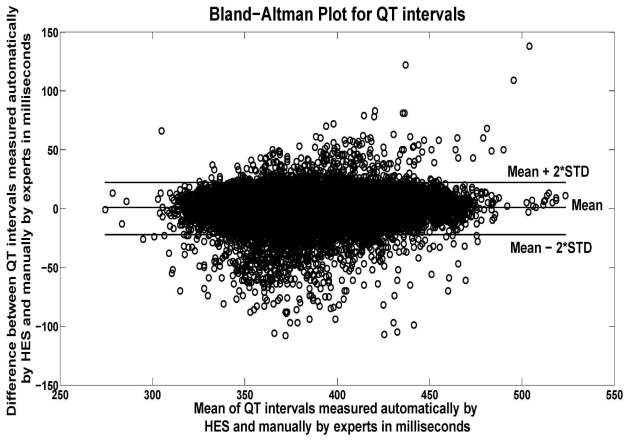

Figure 2. The Bland-Altman plot for QT interval.

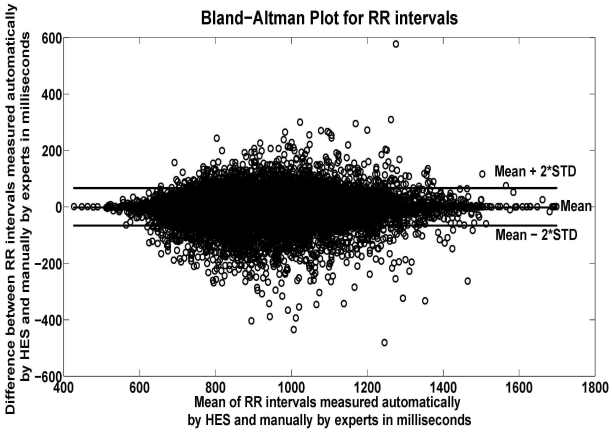

Figure 3. The Bland-Altman plot for RR interval.

4, while figure 3 depicts the large range of the relation between measurement difference and reference RR interval which is represented in a Bland-Altman plot.

Table 4. The distribution of the absolute differences between HES® and reference value for RR interval.

Absolute difference (ms)	Number of records	Relative frequency (%)	Cumulative frequency (%)
0 - 5	21063	49.48	49.48
6 - 10	4657	10.94	60.42
11 - 15	3478	8.17	68.59
16 - 20	2737	6.43	75.02
21 - 30	3570	8.39	83.41
31 - 40	2106	4.95	88.36
41 - 50	1377	3.23	91.59
51 - 60	872	2.05	93.64

4. Discussion and conclusions

The delineation precision for single-lead measurements of the QT and RR interval in the HES® algorithm has been investigated and validated in this work. It is carried out using a large internal database containing more than forty-two thousand manually-annotated ECG records from two late-phase cardiac safety studies. The measurement obtained by the experts in these studies has been considered as golden reference for this work. The low mean differences between the reference and automatic measurements of QT and RR intervals expresses that HES® is a useful tool in the evaluation of cardiac safety in clinical trials. We interpret the results and findings of this work as indicator of the reliability of the fully-automatic HES® ECG analysis algorithm.

References

[1] Darpo B, Nebout T, Sager PT. Clinical evaluation of qt/qtc prolongation and proarrhythmic potential for nonantiarrhythmic drugs: the international conference on harmonization of technical requirements for registration of pharmaceuticals for human use e14 guideline. J Clin Pharmacol 2006;46:498–507.

[2] Hughes N. Probabilistic Models for Automated ECG Interval Analysis. University of Oxford, 2006.

[3] Meij SH, Klootwijk P, Arends J, Roelandt JRTC. An algorithm for automatic beat-to-beat measurement of the qt-interval. Computers in Cardiology 1994;597–600.

[4] McLaughlin NB, Campbell RWF, Murray A. Accuracy of automatic qt measurement techniques. Computers in Cardiology 1993;863–866.

[5] Sahambi JS, Tandon SN, Bhatt RKP. An automated approach to beat-by-beat qt-interval analysis 2000;19:97–101.

[6] Baas T, Gravenhorst F, Fischer R, Khawaja A, Dossel O. Comparison of three t-wave delineation algorithms based on wavelet filterbank, correlation and pca. Computing in Cardiology 2010;361–364.

[7] Zywietz C, Borovsky D, Faltinat D, Klusmeier S, Schiemann W. The hannover ekg system hes. Trend in Computer Processed Electrocardiograms 1977;159–167.

[8] Laguna P, Mark RG, Goldberg A, Moody GB. A database for evaluation of algorithms for measurement of qt and other waveform intervals in the ecg. Computers in Cardiology 1997;673–676.

Address for correspondence:

Dr. Antoun Khawaja
Lindwurmstrasse 109, 80337, Munich, Germany.
antoun.khawaja@cordiagpro.de