PC-Based ECG Waveform Recognition—Validation of Novel Software Against a Reference ECG Database

Corina-Dana Dota, M.D.,*† Nils Edvardsson, M.D., Ph.D.,† Bo Skallefell,* and Gunnar Fager, M.D., Ph.D.*

From the *AstraZeneca R&D Mölndal, Mölndal, Sweden; and †Division of Heart and Lung Diseases, Sahlgrenska University Hospital, Göteborg, Sweden

Background: PC-based ECG measurements must cope with normal as well as pathological ECGs in a reliable manner. EClysis, a software for ECG measurements was tested against reference values from the Common Standards for Quantitative Electrocardiography (CSE) database.

Methods: Digital ECGs (12 leads, 500 Hz) were recorded by the CSE project. Data Set 3 contains reference values for 125 ECGs (33 normal and 92 pathological). Median values of measurements by 11 computer programs and by five cardiologists, respectively, refer to the earliest P and QRS onsets and to the latest P, QRS, and T offsets in any lead of a selected (index) beat. EClysis automatically measured all ECGs, without user interference.

Results: The PQRST points were correctly detected but in two ECGs with AV block II–III. The software was not designed to detect atrial activity in atrial fibrillation ($n = 9$) and flutter ($n = 1$). In one case of atrial fibrillation, atrial activity interfered with positioning of QRS and T offsets. Regression coefficients between EClysis[®] and CSE (software-generated and human) were above 0.95 ($P < 0.0001$). The confidence intervals were 95% for the slope and the intercept of the regression lines.

Conclusions: The PC-based detection and analysis of PQRST points showed a high level of agreement with the CSE database reference values.

Ann Noninvasive Electrocardiol 2009;14(Suppl. 1):S42–S47

ECG; computerized measurement; CSE database; QT interval

The measurement of ECG waveforms and intervals in drug development is gaining importance; such measurements are currently required for the assessment of safety and efficacy of new compounds. The current "golden standard" of ECG measurements states that they should be manually performed and reviewed by expert cardiologists, or that semiautomated methods should be used (computer-assisted measurements with manual adjudication).¹ However, there is significant intra- and interobserver variability in manual measurements of ECG intervals, in particular for the QT interval,^{2–4} and computer-assisted methods to analyze ECGs are evaluated.⁵ Compared with manual methods, the automated measurements are precise and objective, i.e., not subject to intra-

and interobserver variability⁶ and are cost effective, but they are not error-free, especially in "difficult" pathological ECGs. Reference libraries to assess the standardization and accuracy of multi-lead wave recognition programs and of computer-derived ECG measurements are required.⁷ An international cooperative project, Common Standards for Quantitative Electrocardiography (CSE) was initiated to establish a reference database,⁸ to allow the testing of accuracy of new analysis systems and "eventually improve multi-lead wave recognition programs."⁹ The present study aimed to compare and correlate the results obtained when testing new software against the reference values provided with the CSE ECG database.

The project was sponsored by AstraZeneca R&D Mölndal, Sweden.

Address for reprints: Corina-Dana Dota, M.D., AstraZeneca R&D Mölndal, S-431 83 Mölndal, Sweden. Fax: +46 31 776 37 28; E-mail: Corina.Dota@astrazeneca.com

METHODS

The CSE Reference Database

Data Set 3 of ECGs provided by the Commission of the European Communities in the CSE database has been described in detail elsewhere.⁸⁻¹⁰ This data set contains normal ($n = 33$) and pathological ($n = 92$) ECG recordings from 125 subjects. Some ECGs showed one or more abnormalities and in total 168 diagnoses were reported.⁹ In all ECGs, the eight independent standard leads (bipolar leads I and II and precordial leads V_1 – V_6) were simultaneously recorded as digital signals at 500 Hz. Leads III, aVL, -aVR, and aVF were mathematically derived from leads I and II.¹⁰ In each ECG, one given (index) beat was actually measured and a reference value was provided together with the recording. The index beat was identified by a specified interval of sampling points from start of the recording. The positions of PQRST points in each ECG were related to the starting point of the index beat. The CSE provided values for the earliest onset and the latest offset of P waves and QRS complexes, as well as for the latest offset of the T wave in any lead (defined as "global" onsets and offsets)¹¹ as measured by eleven different 12-lead and 6 XYZ computer programs. In the present study, comparison was made only with the first (12-lead) category. The global onsets and offsets provided by the 12-lead programs were ranked, and median values were given for each global onset and offset. These median values belong to an ideal, artificially created program, called the Combined Median Program, which provides the CSE reference value for each onset and offset, in all recordings.¹⁰ The multilead database together with the results of automatic and manual measurements were provided by CSE on CD-ROM.

Taking into account the results provided by all computer programs included in the CSE project and by the Combined Median Program, we constructed ranges for all CSE measurements.

Beside the automated measurements, 25 randomly selected (every fifth) ECGs were manually measured on highly amplified tracings by five blinded experienced cardiologists for quality control.⁸ The results of manual measurements were reported as medians with ranges for comparison.

The EClysis Software

The main features of EClysis (AstraZeneca R&D, Mölndal, Sweden) were described elsewhere.¹² Briefly, EClysis was primarily designed to automatically and accurately detect and measure ECG waves and intervals but not pacemaker rhythms or atrial activity in atrial fibrillation or flutter or P waves in AV block II–III.

Independent across all 12 leads, it automatically identifies the following fiducial points: P_{start} , P_{top} , Q_{start} , R_{start} , R_{top} , S_{top} , QRS offset (J point), T_{top} , T_{end} (defined as T_{tang} or $T_{90\%}$) and calculates their positions and relevant amplitudes. If a biphasic T wave or a U wave is present, the program identifies the $T1_{top}$, $T1_{end}$ and TU_{gap} , U_{top} and U_{end} , respectively.

Analysis is made from the original ECG signal without any derivations and uses a very large number of criteria to correctly identify normal and pathological waveforms. To find relevant P, T, and U fiducial points, minimally influenced by signal noise, a temporary, smoothed ECG curve is obtained as a floating mean of seven consecutive points. However, the reported amplitudes are the actual ECG amplitudes at these time points. QRS points are calculated without smoothing. In this comparison, T_{tang} was used to define the offset of the T wave. The T_{tang} is calculated as the tangent between 20% and 80% repolarization of the T wave calculated from the T_{top} to the isoelectric line. The T_{tang} is based on a best-fit regression line plotted through these points and the interception with the isoelectric line. The calculations are independent of positive or negative T waves.

To comply with the CSE way of measuring "global" ECG intervals,¹⁰ the software provided global onsets and offsets across the 12 leads, used in the present study.

EClysis is completely interactive for accredited users, allowing them to exclude disturbed beats, beat groups, or leads from calculations as well as to move or remove points that were incorrectly positioned. All such interventions are kept in write-protected files for audit purposes. In this study, no automatically positioned points were moved or removed.

EClysis was expected to correctly determine the wave onsets and offsets and to measure PQ, QRS and QT intervals in more than 95% of all normal and pathological ECGs during normal sinus rhythm and AV block I.

Comparison between EClysis and CSE Measurements

The positions of global onsets and offsets, determined by EClysis as the number of points sampled at 500 Hz from the given start position in the index beat, were compared with the medians of those in the CSE database, using linear regression analysis. First, the measurements were compared with those provided by the 12-lead computer programs ($n = 11$, plus the Combined Median Program) and secondly, to the results of manual measurements on the subsample of 25 subjects. In the

enclosed graphs (Figs. 1 and 2), error bars indicate the extreme values reported in the CSE database. Values missing in the CSE database were ignored. The mean values and standard deviations of the absolute and relative differences between EClysis-derived values and manual CSE values were calculated.

Statistical Procedures

Descriptive statistics were calculated using routine procedures. Experimental and reference values were compared using least squares linear

Figure 1. The earliest onset: of P wave (A) and of QRS complex (B), and the latest offset of QRS complex (J point) (C) and of T wave (D) in any lead. EClysis values are plotted on the y-axis, and the values provided by the CSE "Combined Median Program" on the x-axis ($n = 111$ in A, and $n = 122$ in B, C, and D). Error bars indicate the range of values generated by the 11 computer programs in the CSE database. The lines of identity (hatched) and regression (solid) are indicated.

Figure 2. The earliest onset of P wave (A) and QRS complex (B), and the latest offset of QRS complex (J point) (C) and T wave (D) in any lead. Values obtained with EClalysis are plotted on the y-axis, against the median values provided by five expert cardiologists in the CSE database on the x-axis ($n = 24$ for P onset, and $n = 25$ for QRS and T). The lines of identity (hatched) and regression (solid) are indicated.

regression analyzes. The 95% confidence intervals of slopes and intercepts were calculated, and values outside this interval were regarded as being significantly different. P values < 0.05 (2-sided tests) were regarded as being statistically significant.

RESULTS

Correlations between EClalysis and the Combined Median Program values were very close ($r > 0.98$, $P < 0.0001$) for "global" onsets and offsets of waves (Figs. 1A–D). Furthermore, the slopes and intercepts of the regression lines were very close to identity. Indeed, the 95% confidence intervals for intercepts and slopes included zero and

1.0, respectively. In a few cases, EClalysis provided values outside the range of the other 12-lead programs, and there were high as well as low extreme values.

Likewise, correlations with the medians of manual measurements by expert cardiologists were very good ($r > 0.95$, $P < 0.0001$) for these global values (Figs. 2A–D). The slopes and intercepts of the regression lines were close to identity. However, the y intercept of the regression line for the P onset was not zero. The intercept was 5 acquisition points above zero (range 1–9), which corresponds to 2–18 ms. Occasionally, EClalysis provided values above as well as below the range of referee values.

Table 1. Mean Values and 95% Confidence Intervals for Absolute and Relative Differences between PQRSTU Positions Determined by EClysis and by Five Experienced Cardiologists (Median Manual CSE Values)

	Subjects (n)	Absolute Difference		Relative Difference ^a	
		Mean (ms)	CI	Mean	CI
P _{onset}	24	3.5	18.4	0.12	0.22
QRS _{onset}	25	−3.4	9.9	−0.10	0.04
QRS _{offset}	25	4.4	12.9	0.02	0.03
T _{end}	25	−20.2	44.4	−0.06	0.06

^a(EClysis – CSE)/CSE

The mean values and confidence intervals of absolute and relative differences between the PQRSTU positions provided by the two methodologies are presented in Table 1. These results indicate that the values provided by the novel software did not differ significantly from the corresponding median values derived from manual measurements of five experienced cardiologists.

None of the 12-lead programs included in the CSE evaluation could detect/measure atrial activity in 11 recordings (132 instances) with atrial fibrillation or flutter and AV-block > I degree. Apart from these, values for some additional point estimates were missing. In some instances, one or more 12-lead computer programs failed to detect the P-wave onset, so that another 21 global P-wave onsets remained unidentified. Since these programs were expected to generate P onsets in 112 ECGs (1344 instances), this corresponded to failure in 1.5% of all cases. Also, these programs failed to identify T offset in 1.6% and QRS onsets and offsets in 0.5% of all cases.

As a consequence of EClysis design, P wave onsets were identified according to specifications in all 112 cases but one. In this one case of atrial fibrillation/flutter, atrial activity disturbed the QRS complex and the T wave and made appropriate identification of the true QRST points impossible. These values are reported as missing and correspond to 0.8% of the expected number of values.

DISCUSSION

EClysis accurately detected PQRST points in more than 95% of all ECGs in accordance with

program acceptance criteria and identified fiducial points at least as often as comparator programs. No program performed any measurements in the two cases (1.6%) with pacemaker rhythm. Atrial activity was not at all or incorrectly identified in another 12 cases (9.6%) with atrial fibrillation/flutter or AV blocks > degree I. None of the programs was designed to handle these P-wave abnormalities. In ECGs without such pathology, EClysis correctly identified P waves, whereas the comparator programs failed to detect the onset of the P in another 1.5% of all instances.

EClysis performed successful measurements of QRST points in all but one recording (0.8%) that showed atrial fibrillation/flutter. In that same recording, one of the comparator programs also failed to determine QRST points and another program failed to identify T offset.

There was good agreement between measurements by EClysis, the Combined Median Program, and the five cardiologists, with respect to the positioning of the global onset and offsets of waves. In fact, all correlations but one had zero intercept and slope 1.0 (identity) within its 95% confidence intervals. The exception (an y intercept of 5 sampling points, or 10 ms) refers to the comparison of global P onset by EClysis and the cardiologists.

The criteria imposed at the beginning of this validation study of EClysis against the measurements provided in the CSE database were: accurate determination of global onsets and offsets of ECG waves in normal and pathological ECGs (excepting cases with pacemaker rhythms, multiple P waves in AV blocks > degree I, or atrial activity in atrial fibrillation or flutter). The fulfillment of these criteria suggests that automatic measurement by the novel software is reliable in the definition and measurement of PQRST values in normal as well as pathological ECGs.

Limitations of this study include the relative small size of the reference database.

In summary, the EClysis software accurately detects the fiducial points in a variety of pathological as well as in normal ECGs from the CSE reference database.

Acknowledgments: We are grateful to the Common Standards for Quantitative Electrocardiography Project (Leuven, Belgium) for their permission to use the Data Set 3 as reference in this comparative study. The authors acknowledge Tomas Morsing, Ph.D. (former Biostatistics, AstraZeneca R&D Mölndal, Sweden) for statistical advice.

Conflicts of Interest: All authors were employed by AstraZeneca R&D, Mölndal, Sweden at the time of the research and writing of this article.

REFERENCES

1. International Conference on Harmonization (ICH). 2005. Guidance on E14 Clinical Evaluation of QT/QTc Interval Prolongation and Proarrhythmic Potential for Non-antiarrhythmic Drugs. Available at: <http://www.fda.gov/der/guidance/6922f1.pdf>.
2. Lepeschkin E, Surawicz B. The measurement of the Q-T interval of the electrocardiogram. *Circulation* 1952;6:378.
3. Kautzner J, Gang Y, Camm AJ, et al. Short- and long term reproducibility of QT, QTc and QT dispersion measurements in healthy subjects. *PACE* 1994;17:928-937.
4. Azie NE, Adams G, Darpo B, et al. Comparing methods of measurement for detecting drug-induced changes of the QT interval: Implications for thoroughly conducted ECG studies. *An Noninv Electrocardiol* 2004;9:166-174.
5. Goldberger AL, Amaral LAN, Glass L, et al. PhysioBank, PhysioToolkit, and PhysioNet components of a new research resource for complex physiologic signals. *Circulation* 2000;101:e215-e220.
6. Puddu PE, Bernard PM, Chaitman BR, et al. QT interval measurement by a computer assisted program: A potentially useful clinical parameter. *J Electrocardiol* 1982;15:15-21.
7. Bailey JJ, Horton M, Itscoitz SB. A method for evaluating computer programs for electrocardiographic interpretation III Reproducibility testing and the sources of program errors. *Circulation* 1974;50:88-93.
8. Willems JL, Arnaud P, van Bommel JH, et al. Establishment of a Reference Library for Evaluating Computer ECG Measurement Programs. *Comp Biomed Res* 1985;18:439-457.
9. Commission of the European Communities- Medical and Public Health Research. Common Standards for Quantitative Electrocardiography: CSE Multilead Atlas. Measurement Results- Data Set 3, Ref. Nr. CSE 88-04-15, 1988.
10. Willems JL, Arnaud P, van Bommel JH, et al. Common standards for quantitative electrocardiography: Goals and main results. *Meth Inform Med* 1990;29:263.
11. The CSE Working Party. Recommendations for measurement standards in quantitative electrocardiography. *Eur Heart J* 1985;6:815.
12. Dota CD, Edvardsson N, Skallefäll B, et al. Computer-based analysis of dynamic QT changes. Toward high precision and individual rate correction. *An Noninv Electrocardiol* 2002;7:289-301.