Accurate QT correction method from transfer entropy

Esa Räsänen, PhD,* Teemu Pukkila,* Matias Kanniainen,* Minna Miettinen,* Rostislav Duda,* Jiyeong Kim,* Janne Solanpää, PhD,* Katriina Aalto-Setälä, MD,†‡ Ilya Potapov, PhD*†‡

From the *Computational Physics Laboratory, Tampere University, Tampere, Finland, †Faculty of Medicine and Health Technology, BioMediTech, Tampere University, Tampere, Finland, and ‡Heart Hospital, Tampere University Hospital, Tampere, Finland.

BACKGROUND The QT interval in the electrocardiogram (ECG) is a fundamental risk measure for arrhythmic adverse cardiac events. However, the QT interval depends on the heart rate and must be corrected accordingly. The present QT correction (QTc) methods are either simple models leading to under- or overcorrection, or impractical in requiring long-term empirical data. In general, there is no consensus on the best QTc method.

OBJECTIVE We introduce a model-free QTc method—AccuQT—that computes QTc by minimizing the information transfer from R-R to QT intervals. The objective is to establish and validate a QTc method that provides superior stability and reliability without models or empirical data.

METHODS We tested AccuQT against the most commonly used QT correction methods by using long-term ECG recordings of more than 200 healthy subjects from PhysioNet and THEW databases.

RESULTS AccuQT overperforms the previously reported correction methods: the proportion of false-positives is reduced from 16% (Bazett) to 3% (AccuQT) for the PhysioNet data. In particular, the QTc variance is significantly reduced and thus the RR-QT stability is increased.

CONCLUSION AccuQT has significant potential to become the QTc method of choice in clinical studies and drug development. The method can be implemented in any device recording R-R and QT intervals.

KEYWORDS Electrocardiography; ECG; QTc interval; QT interval; Numerical techniques

Introduction

A prolonged or shortened QT interval in the electrocardiogram (ECG) is a significant risk factor for numerous cardiac diseases and adverse events such as torsades de pointes, ventricular fibrillation, and sudden cardiac death.1,2 To interpret the QT interval, it needs to be corrected for the heart rate (HR) determined by the R-R intervals. In clinical practice, the corrected QT (QTc) is routinely monitored and its assessment is a mandatory part of drug development.3

Despite the paramount importance of QTc, its use and impact are limited by the lack of universal and reliable correction methods. The commonly used power-law formulae of Bazett and Fridericia, as well as other population-based models such as Hodges and Framingham corrections, are known to under- or overcorrect the QT intervals in various situations and have thus received considerable criticism.1,8,9 On the other hand, subject-specific or individual QT corrections require a reliable long-term baseline of R-R and QT intervals. Moreover, QT hysteresis—that is, differences in the QT interval depending on the derivative of the HR (decreasing or increasing)—complicates the individual QT correction.

It is widely accepted that, on one hand, no universal formula can account for the QT/R-R dependence, and, on the other hand, individual corrections combined with, for example, hysteresis models are not well applicable to short-term QT assessment or clinical analysis.9 Furthermore, the intrasubject variability resulting from both biological and methodological sources limits the use of QTc monitoring in clinical applications. This factor should be taken into account in the development of future QT correction methods.

In this work we introduce a QT correction method called AccuQT, which employs the properties of transfer entropy to minimize the dependence of the QT intervals on the R-R intervals. The method is free of empirical data or models. We carry out a thorough testing of the method with 2 large datasets of long-term ECG recordings, and focus on the performance of AccuQT to reduce the individual QTc variation. Overall, AccuQT is found to have superior stability and consistency compared to the conventional QTc methods, and it is applicable also to short-data intervals.

Methods

The study has been approved by the ethical committee of Pirkanmaa Hospital District (R21067L).
AccuQT correction method

We assume that the QT dependence on the HR (or RR) is determined by the information flow from the R-R intervals to the QT intervals, as illustrated in Figure 1A. An effective measure for the information flow is transfer entropy (TE), originally introduced by Schreiber. Some of the present authors have applied TE to the RR-QT relationship and found that, first, the previous R-R intervals affect the present QT interval, whereas the previous QT intervals do not affect the present R-R interval. Secondly, the dependence of the present QT interval on the previous R-R intervals extends to the history by more than about 10 intervals. The key idea in the AccuQT method is to set TE to zero in order to reduce the information flow from R-R to QT, and thus to determine the corrected QT value (QTc). 16

TE corresponds to the difference in the information (or uncertainty) between (1) future observation QT(t+1) obtained from the past observations of QT and (2) future observation QT(t+1) obtained from the past observations of both R-R and QT. Formally, we can write TE from R-R to QT as in equation 1:

\[
TE_{RR\rightarrow QT} = \sum_i p\left(QT_i, QT_{i-1}^{(k)}, RR_{i-1}^{(n)}\right) 
\log_2 \frac{p(QT_i|QT_{i-1}^{(k)}, RR_{i-1}^{(n)})}{p(QT_i|QT_{i-1})}
\]

where \(p(x)\) and \(p(x|y)\) are probability and conditional probability distributions, respectively. In accordance with references 17 and 18, we refer to the numerator and denominator of eq. (1) as transition probabilities. The indices \(k\) and \(n\) correspond to the preceding QT and RR values, respectively, backwards from \(i-1\). In conjunction with the findings in reference 18 we set the QT history to \(k = 1\) and vary the RR history in the range \(n = 1 \ldots 50\). In practice, the RR history length of \(RR_h = 20\) is found to be sufficient, and it is used here as the default value. The effects of RRh are examined in the Supplemental Data (Appendix).

We apply the kernel density estimation with a Gaussian kernel to obtain smooth transition probabilities. For a given series of R-R and QT intervals, we estimate the joint probability distribution of events \(QT_i, QT_{i-1}, \ldots, QT_{i-k}\) and \(RR_{i-1}, \ldots, RR_{i-n}\). Then for an interval \(QT_i\) and its history \(QT_{i-1}\) and \(RR_{i-1}\), we take a slice of the joint distribution along the QT axis, as illustrated in Figure 1B. This slice is normalized according to the chain rule and it represents the 1-dimensional conditional probability density \(p(QT_i|QT_{i-1}^{(k)}, RR_{i-1}^{(n)})\) or \(p(QT_i|QT_{i-1})\) as a function of QT. According to eq. (1), the QT values where these 1-dimensional distributions intersect correspond to zero TE. Thus, these QT values can be considered as candidates for corrected QT (QTc) for each heartbeat as the corresponding information transfer from RR to QT is then zero—in principle—taking into account the statistical context of probability distributions.

As the final step, we select the target QTc candidate to represent the final QTc (for each heartbeat) as the output. The selection can be done in a nonempirical fashion by, eg, using the maximum probability (highest intersection point) or the mean of all the QTc candidates exceeding a minimum threshold in the probability density. Here, however, we determine QTc for each heartbeat as the QTc candidate that has the minimum distance to QT0; that is, the QT value that corresponds to 60 beats per minute (RR = 1000 ms) according

![Figure 1](image-url)
to a second-order polynomial fit to the RR-QT point cloud. In the Appendix we show how the results depend on the size of the point cloud, ie, the length of the RR-QT measurement.

Other QT corrections

We compare the AccuQT results with the 4 most commonly used QT correction methods discussed above and defined as follows.

- Bazett: $QTc = QT / \sqrt{RR}$
- Fridericia: $QTc = QT / \sqrt[3]{RR}$
- Hodges: $QTc = QT + 0.00175(60 - HR)$
- Framingham: $QTc = QT + 0.154(1 - RR)$

Here the QT, QTc, and R-R intervals are expressed in seconds and HR in beats per minute. In some cases we compare the results also with the individual QT correction method. In particular, we use the nonlinear regression model $QT_i = \chi - \phi (1 - (RR_i)^g)$ for individual RR-QT distributions, which leads to the QT correction formula of the form $QTc_i = QT_i + \phi (1 - (RR_i)^g)$. The obvious drawback of this method is the empirical approach requiring a large number of R-R and QT intervals to obtain reliable results.

Data

We focus on RR and QT data of healthy subjects and employ 2 databases, MIT-BIH Normal Sinus Rhythm Database in PhysioNet and THEW dataset.

PhysioNet data

The database contains 18 subjects without significant arrhythmias (5 men, aged 26–45 years; 13 women, aged 20–50 years). The time series vary between 500 and 22,000 beats, ie, between about 10 minutes and 5 hours. In the preprocessing, we removed the most obvious artifacts with a moving average method, while trying to keep the time series as contiguous as possible. We removed 3 samples completely owing to insufficient data quality to reliably extract the R-R and QT intervals. The basic measures of the remaining dataset are shown in Table 1.

THEW data

The THEW dataset (E-HOL-03-0202-003) contains 24-hour Holter ECG recordings with 3-lead pseudo-orthogonal lead configuration from 202 healthy subjects (100 men, aged 12–80 years; 100 women, aged 9–82 years; and 2 undefined sex). For the preprocessing we used our own ECG delineation algorithm. The dataset is subject to artifacts, that is, missing or erroneous signals or inaccuracies in the QT extraction, especially owing to the ambiguity of the T wave. In particular, some samples were found to contain a significant proportion of unusually long QT values for healthy subjects. Thus, we discarded 18 out of 202 samples (8.9% of the samples), which have more than 20% of QT intervals larger than 500 ms. The basic measures of the remaining THEW dataset are shown in Table 1.

Results

PhysioNet dataset

The recordings of the PhysioNet dataset are relatively stable and have low HRs with combined mean R-R of 921 ms. Table 2 shows the mean values and standard deviations (SDs) of all the R-R and QT intervals combined from all the subjects, as well as the QTc values computed with the correction methods introduced in the Methods section. It is noteworthy that, as expected, the Bazett values are higher than the rest owing to likely overcorrection, whereas the AccuQT values are relatively low and close to the raw QT values. However, the SD is considerably lower in AccuQT than in the other methods.

The difference between AccuQT and the other QTc methods is highlighted in the last column showing the pooled SD. This measure takes into account the different mean values of the subjects and thus yields a better measure for the QT(c) variability of the dataset than the SD of all the available subject-independent intervals. In AccuQT the pooled SD is as low as 13 ms, whereas in the other correction methods the variation ranges from 24 ms (individual QTc) up to 32 ms (Bazett).

Table 2 Mean values and standard deviations of the R-R, QT, and computed QTc intervals collected from all 15 subjects of the PhysioNet dataset included in the study

<table>
<thead>
<tr>
<th>Interval</th>
<th>Mean (ms)</th>
<th>SD (ms)</th>
<th>Pooled SD (ms) *</th>
</tr>
</thead>
<tbody>
<tr>
<td>R-R</td>
<td>921</td>
<td>156</td>
<td>116</td>
</tr>
<tr>
<td>QT</td>
<td>402</td>
<td>32</td>
<td>26</td>
</tr>
<tr>
<td>QTc-Bazett</td>
<td>422</td>
<td>37</td>
<td>32</td>
</tr>
<tr>
<td>QTc-Fridericia</td>
<td>415</td>
<td>32</td>
<td>28</td>
</tr>
<tr>
<td>QTc-Hodges</td>
<td>414</td>
<td>30</td>
<td>26</td>
</tr>
<tr>
<td>QTc-Framingham</td>
<td>414</td>
<td>32</td>
<td>28</td>
</tr>
<tr>
<td>QTc-individual</td>
<td>410</td>
<td>30</td>
<td>24</td>
</tr>
<tr>
<td>QTc-AccuQT</td>
<td>403</td>
<td>23</td>
<td>13</td>
</tr>
</tbody>
</table>

*Pooled SD values, where subject-specific SDs are combined.
In Figure 2, we visualize the QT(c) dependence on the R-R intervals for the (a) raw QT values, (b) Bazett, (c) Fridericia, and (d) AccuQT method computed for the complete PhysioNet dataset without subject-specific separation (>120,000 RR-QT(c) pairs in total). The point clouds on the left are smoothened to a normalized (0…1) density plot according to reference 27, using 200 × 200 bins and smoothing parameter \( \lambda = 5 \). The box plots on the right have 50-ms-wide R-R interval bins. The lines in the boxes represent the first, second, and third quartiles, and the whiskers extend up to 1.5 times of the interquartile range. The gray dots denote outliers falling outside the whiskers.

The trends of the point clouds in Figure 2 are also visible in the box plots on the right: the Fridericia and AccuQT methods yield the most stable behavior, but in AccuQT the deviations within the bins are significantly smaller than in the Fridericia method, especially regarding the outliers shown in gray dots. Overall, the QTc variability is significantly reduced in AccuQT, compared to the other methods across the R-R interval range. This superior stability would allow effective QTc assessment in clinical practice and in drug development. The bright horizontal fringe in the AccuQT density cloud in Figure 2(D) is owing to the limited number of subject-specific point clouds (15), which are combined into 1 figure.

Next, we consider in more detail the subject-specific RR-QT(c) point clouds of the PhysioNet dataset. Figure 3 shows the point clouds for the 6 longest recordings of the set compared between raw QT (upper row), Bazett (middle row), and AccuQT (lower row). The color scale indicates the difference between QTc and QT, so that negative values (blue) correspond to reduced intervals in the correction and vice versa. The tendency of overcorrection of the Bazett method, that is, too-large QTc values at small R-R intervals
and—in some cases—too-small QTc values at large R-R intervals, is evident. In contrast, AccuQT gives relatively stable QTc values as a function of the R-R intervals, and the overall variability is also significantly reduced, so that a large majority of the points fit between 350 and 450 ms. As the AccuQT values are calculated based on the histories, there is no evident pattern in the coloring. This is in contrast with the Bazett point clouds that display an evident pattern that follows from the simple QTc formula. However, we should bear in mind that the Bazett formula was originally designed not for long ECG recordings, but for relatively stable and short measurements.

**THEW dataset**

In Table 3, we present the basic statistical measures of the THEW dataset comprising over 18 million R-R and QT intervals of 184 subjects. Compared to the PhysioNet dataset (Table 2), the R-R and QT intervals are significantly shorter with higher SD, which is due to the (approximately) 24-hour measurement protocol that includes various daily activities, compared to the shorter and possibly more controlled measurements behind the PhysioNet data. Also for this dataset, AccuQT yields the smallest mean for the QTc values and the smallest pooled SD. However, the difference is not as drastic as with the PhysioNet dataset.

![Figure 3](image-url) **Figure 3** R-R interval (RRI)-QT(c) point clouds for the 6 longest recordings of the PhysioNet dataset. In the Bazett and AccuQT clouds the colors indicate the shift from the raw QT values (QTc-QT).

The proportion of QTc values satisfying the normal limits is the highest in AccuQT, shown in Figure 4(F) with 75%, whereas in other QTc methods the proportion varies between 51% (Bazett) and 73% (Framingham), compared to 66% in raw QT values shown in Figure 4(A). Hence, the results are in line with the PhysioNet dataset in Figure 2 in the sense that AccuQT is the most and Bazett the least accurate. In particular, the obvious overcorrection of the Bazett method at low R-R intervals is clearly visible.

However, here for the THEW dataset the proportions in the normal regime are significantly lower than in the case of PhysioNet data. This is owing to the 24-hour recordings in uncontrolled conditions, which are prone to large variations in the intervals as well as to detection errors. We suspect that especially the leftmost linear “tail” of the point cloud results from measurement and/or QT extraction errors.

Regarding the apparent RR-QT trend or slope in the AccuQT data in Figure 4(F), we remind that the point clouds represent combinations of data from several subjects, whose individual trends may be significantly smaller. This is analyzed in detail below.

**Table 3** Mean values and standard deviations of the R-R, QT, and computed QTc intervals collected from all the 184 subjects of the THEW dataset included in the study

<table>
<thead>
<tr>
<th>Interval</th>
<th>Mean (ms)</th>
<th>SD (ms)</th>
<th>Pooled SD (ms)*</th>
</tr>
</thead>
<tbody>
<tr>
<td>RR</td>
<td>775</td>
<td>173</td>
<td>150</td>
</tr>
<tr>
<td>QT</td>
<td>398</td>
<td>54</td>
<td>48</td>
</tr>
<tr>
<td>QTc-Bazett</td>
<td>456</td>
<td>49</td>
<td>46</td>
</tr>
<tr>
<td>QTc-Fridericia</td>
<td>435</td>
<td>46</td>
<td>42</td>
</tr>
<tr>
<td>QTc-Hodges</td>
<td>435</td>
<td>43</td>
<td>40</td>
</tr>
<tr>
<td>QTc-Framingham</td>
<td>432</td>
<td>43</td>
<td>39</td>
</tr>
<tr>
<td>QTc-Individual</td>
<td>441</td>
<td>41</td>
<td>37</td>
</tr>
<tr>
<td>QTc-AccuQT</td>
<td>408</td>
<td>42</td>
<td>35</td>
</tr>
</tbody>
</table>

*Pooled SD values, where subject-specific SDs are combined.
Slopes and stability of the RR-QT cloud

The RR-QT(c) point clouds shown in Figures 2 and 4 indicate relatively large differences between the methods regarding the general RR-QTc trends. In Table 4, we show the slopes for the RR-QT(c) and HR-QT(c) clouds of both the PhysioNet and THEW datasets, respectively. The slopes have been computed for the full set of RR-QT(c) interval pairs in the dataset, i.e., not as means of subject-specific slopes. For comparison, we include the slopes computed from 2 other studies using different databases and measurement protocols, the RR-QTc slopes of Vandenberk and colleagues and the HR-QTc slopes of Luo and colleagues.

For the THEW data the AccuQT slope is noticeably higher than with any of the conventional methods, whereas for the PhysioNet data the magnitude of the AccuQT slope is smaller than for most of the methods and similar to that of the Hodges correction. Interestingly, the raw QT slope is also considerably smaller for the PhysioNet data, compared to THEW data or to that of Luo and colleagues. As expected, the individual QTc shows the smallest slopes, resulting from the aims and the definition of the method to minimize the RR-QTc dependency for each subject.

The tendency of AccuQT to leave a positive RR-QTc slope and correspondingly a negative HR-QTc slope, especially for the THEW data, requires further examination. As mentioned above, the leftmost part of the density clouds toward R-R intervals ~500 ms is suspicious regarding the quality of the data and QT extraction. At R-R intervals above ~800 ms the clouds are more consistent and would yield smaller slopes than obtained through global fitting. In fact, restricting the analysis to the stable segments of the THEW data—as defined in the following section—leads to a slope of 0.077 ± 0.087.

For practical applications, it is informative to examine individual RR-QTc behavior (see Figure 3), which we here demonstrate as a function of time. Figure 5 shows a 200-beat segment of QT(c) values (Figure 5A) and the corresponding R-R intervals (Figure 5B) for a healthy subject in the THEW dataset. Compared to the raw QT, Bazett and Fridericia corrections show similar or even increased variability, whereas AccuQT displays significantly reduced variation throughout the segment. This tendency is characteristic of the AccuQT method, as discussed above in the context of reduced (pooled) variances for the full datasets.
For this individual example, AccuQT also has the smallest linear slope for the RR-QT(c) clouds, as seen in Figure 5C. Therefore, the global slopes presented in Table 4 need to be interpreted with care, since in a practical application of a QT correction the individual performance regarding the RR-QTc consistency and stability is crucial.

Discussion

The QT correction, ie, the correction of the QT interval in the ECG for the heart rate, has remained as one of the most important problems in electrocardiography owing to the paramount importance of the corrected QT values in clinical practice, in cardiotoxicological assessments, in drug development, and, in the future, also in preventive care through QT measurements with wearable devices.30 It is generally agreed that no universal formulas of the form QTc = f(QT, RR) can be developed. On the other hand, the empirical individual QT correction methods, which are useful for long measurements providing a lot of RR-QT data, have limitations in practical applications with shorter segments.

In contrast with the conventional QTc corrections, AccuQT has been developed from the first principles without external models by using transfer entropy between the R-R and QT intervals. This approach provides superior performance, but the method has some practical limitations. First, R-R history of preferably 20 intervals needs to be collected for the transfer entropy, alongside with a sufficient number—at least a few dozen—of RR-QT intervals for the probability densities (see the Appendix for details). Secondly, a reliable estimation for the QT at 60 beats per minute (QT0) requires preferably ~100 RR-QT intervals, although accurate determination of QTc is not very sensitive to QT0. Thirdly, it still needs to be evaluated with high-quality (preferably manually inspected) data how the method works for short segments with highly varying R-R intervals. And finally, it is still to be examined how AccuQT deals with the hysteresis effects that have a range above several minutes, thus exceeding the history length included in the method. It is also noteworthy that not all the QT variations are due to the R-R, even though the QT correction—by definition—solely focuses on the removal of the R-R effect from the QT. Despite these limitations we find that AccuQT is well applicable to various QTc assessments apart from very short recordings below about 30 seconds.

We remind that in this work our focus is on healthy subjects and regular rhythms. Therefore, an obvious limitation of the present study is the absence of patient data with heart diseases and the lack of clinical evaluation with statistical measures of the performance of our QT correction compared with the other methods. We have left these examinations as follow-up studies, where we use the existing THEW data for subjects under QT-prolonging medication, as well as subjects under cardiac disease such as long QT syndrome.

Conclusion

To conclude, we have developed a model-free QT correction method—AccuQT—and tested it for healthy subjects in PhysioNet and THEW databases. AccuQT provides exceptionally small intrasubject variation, especially for the PhysioNet database representing more stable measurements than the 24-hour Holter data of the THEW database. The RR-QTc point clouds of the AccuQT method are well condensed in the 350–450 ms range; in the PhysioNet case the result is 97%, compared to that of 84% in Bazett and 90% in Fridericia.

We find significant potential in the AccuQT method to improve the reliability of QT correction. For clinical applications, thorough validation under appropriate measurement conditions compliant with the regulatory pathway are required. Studies of the method on the performance for different diseases and drug effects are already underway.

Funding Sources

This study was supported by Business Finland TUTL (R2B) Project No. 1608/31/2019 (2019-2021) and Novo Nordisk Fonden Pioneer Innovator Grant No. 0066148 (2020-2022).

Disclosures

The authors have no conflicts to disclose.

Authorship

All authors attest they meet the current ICMJE criteria for authorship.
Appendix

Supplementary data

Supplementary data associated with this article can be found in the online version at https://doi.org/10.1016/j.cvdhj.2022.10.006.

References