Automated Quantification of QT-Intervals by an Algorithm: A Validation Study in Patients with Chronic Obstructive Pulmonary Disease

Dario Kohlbrenner, Maya Bisang, Sayaka S Aeschbacher, Emanuel Heusser, Silvia Ulrich, Konrad E Bloch, Michael Furian

1Department of Pulmonology, University Hospital Zurich, Zurich, Switzerland; 2Faculty of Medicine, University of Zurich, Zurich, Switzerland; 3Swiss University of Traditional Chinese Medicine, Bad Zurzach, Switzerland

Correspondence: Michael Furian, Department of Pulmonology, University Hospital Zurich, Raemistrasse 100, 8091, Zurich, Switzerland, Email michael.furian@usz.ch

Study Objectives: To assess the diagnostic accuracy of a purpose-designed QTc-scoring algorithm versus the established hand-scoring in patients with chronic obstructive pulmonary disease (COPD) undergoing sleep studies.

Methods: We collected 62 overnight electrocardiogram (ECG) recordings in 28 COPD patients. QT-intervals corrected for heart rate (QTc, Bazett) were averaged over 1-min periods and quantified, both by the algorithm and by cursor-assisted hand-scoring. Hand-scoring was done blinded to the algorithm-derived results. Bland-Altman statistics and confusion matrixes for three thresholds (460, 480, and 500ms) were calculated.

Results: A total of 32944 1-min periods and corresponding mean QTc-intervals were analysed manually and by computer. Mean difference between manual and algorithm-based QTc-intervals was −1ms, with limits of agreement of −18 to 16ms. Overall, 2587 (8%), 357 (1%), and 0 QTc-intervals exceeding the threshold 460, 480, and 500ms, respectively, were identified by hand-scoring. Of these, 2516, 357, and 0 were consistently identified by the algorithm. This resulted in a diagnostic classification accuracy of 0.98 (95% CI 0.98/0.98), 1.00 (1.00/1.00), and 1.00 (1.00/1.00) for 460, 480, and 500ms, respectively. Sensitivity was 0.97, 1.00, and NA for 460, 480, and 500ms, respectively. Specificity was 0.98, 1.00, and 1.00 for 460, 480, and 500ms, respectively.

Conclusion: Overall, 8% of nocturnal 1-min periods showed clinically relevant QTc prolongations in patients with stable COPD. The automated QTc-algorithm accurately identified clinically relevant QTc-prolongations with a very high sensitivity and specificity. Using this tool, hospital sleep laboratories may identify asymptomatic patients with QTc-prolongations at risk for malignant arrhythmia, allowing them to consult a cardiologist before an eventual cardiac event.

Keywords: QTc, long-QT syndrome, COPD, algorithm, validity, ECG

Introduction
Cardiovascular comorbidity is frequent in chronic obstructive pulmonary disease (COPD) and becomes more prevalent with progressing disease. In fact, individuals with COPD are more prone to die from cardiovascular disease than from COPD itself.

Estimating the risk of cardiovascular events, the heart rate corrected QT-interval (QTc) is widely used. The QT-interval describes the repolarization time of the heart from the beginning of the QRS-complex to the end of the T-wave. Corrected for heart rate, QTc is expressed. QT-intervals are recorded with an electrocardiogram (ECG), which patients with cardiopulmonary disorders undergo on a regular basis. Pathologically prolonged QTc are >460 ms in symptomatic and >480 ms in asymptomatic individuals. Values exceeding these thresholds reveal an increased probability for a long-QT syndrome (LQTS), which might manifest with life-threatening arrhythmia. A QTc >500 ms is considered highly pathological in both men and women and poses an increased risk for, potentially fatal, torsade de pointes tachycardies.
Usually, LQTS is acquired due to disease progression, decreased physical activity, the intake of QT-prolonging drugs, or electrolytic imbalance.\textsuperscript{8,9} Given this multifactorial development, a comprehensive clinical assessment is needed.

Although specific pathophysiological pathways remain to be elucidated, COPD patients are at increased risk for prolonged QTc, especially during acute exacerbations.\textsuperscript{10,11} Impaired lung function, and thus the majority of cardiorespiratory disorders, is associated with increased cardiovascular comorbidity and mortality.\textsuperscript{12–14} Accordingly, QTc is an important marker regarding cardiovascular comorbidity risk in COPD and is widely acknowledged in clinical practice. Sleep studies are frequently required in patients with COPD. Thus, they would provide a widely available opportunity for LQTS screening in patients with COPD since continuous ECG is recorded alongside. In addition, QTc shows considerable variability over the course of a recording. Sampling for a longer duration (ie, overnight) will thus lead to a more accurate clinical picture.\textsuperscript{15}

A precise assessment of QTc requires scoring of all 12 ECG leads.\textsuperscript{7} Given that QTc quantification is currently done by hand-scoring and that in a majority of patients long-term ECGs (ie, Holter or sleep studies) are used, it is a very time-consuming marker.\textsuperscript{16,17} In addition, hand-scoring is assessor-dependent and error-prone.\textsuperscript{18} In contrast, algorithm-based scoring can be readily available, cost-efficient, and fast.\textsuperscript{19} Accordingly, an automated QTc scoring would be useful to increase availability, efficiency, and reliability of the diagnostic.

State-of-the-art ECG devices come with built-in automated calculation of QTc among other clinically important parameters. However, research has shown that these lack accuracy regarding the detection of the LQTS.\textsuperscript{20,21} The diagnosis of a LQTS has clinical impact, leading to treatment consequences such as drug prescription or withdrawal.\textsuperscript{17} Consequently, algorithms are expected to show very high validity and diagnostic accuracy when compared to the current gold-standard.

Thus, we aimed to assess the criterion validity and diagnostic accuracy of a purpose-designed QTc-scoring algorithm versus the established hand-scoring in patients with COPD undergoing sleep studies.

**Materials and Methods**

**Study Design and Patients**

This diagnostic accuracy study was performed within a prospective clinical trial in patients with COPD.\textsuperscript{22–24} All study participants had confirmed moderate to severe stable COPD; patients with hypoxemia, long-term oxygen therapy, or uncontrolled cardiovascular disease were excluded. For the present analysis, we included all patients who had at least one valid overnight ECG recording. All patients provided us with their current medical record. We screened the administered drugs and the diagnosed comorbidities according to their potential impact on QTc (\url{www.crediblemeds.org}). The responsible ethics committee (Kantonale Ethikkommission Zürich, 2013–0088) approved the study and the study adhered to the Declaration of Helsinki. The main study was registered on ClinicalTrials.gov (NCT02150590). All patients provided written informed consent.

**Electrocardiography**

All 4-lead ECGs (recorded at 200 Hz) were conducted using the same device for respiratory sleep studies (Alice 5, Philips Respironics, Zofingen, Switzerland). QRS-complexes in ECG Lead II were averaged for each one-minute epoch. To ensure valid measurements, patients refrained from consuming alcohol and caffeine for at least 12h before the procedure.

**QTc Scoring Algorithm**

We used a purpose-designed algorithm for QTc determination. The algorithm was programmed in MATLAB (MathWorks, Natick MA, USA), applies a graphical user interface and supports import of various data formats.\textsuperscript{22,25}

To compute QTc, the algorithm first detects the peaks of the R-waves. According to this peak, the ECG signal is split into RR-intervals. For all further calculations, multiple RR-intervals are superimposed and averaged. For averaging, the RR-intervals are synchronized according to the peaks of the R-wave. Averaging of the ECG signal results in a smoothing
of the signal. Smoothing is required because the definition of the additional metrics is mostly based on the first or second derivative of the signal.

Using the superimposed and averaged signal, the algorithm detects P-waves, beginning and ending of the QRS-complex, Q-waves, S-waves, and T-waves (ie, its peak and its ending) (Figure 1).

Isoelectric line
The isoelectric line is commonly defined as the TP-segment. However, since an approximation of the isoelectric line is required for the algorithm to define those points, the approximative isoelectric level is defined as the 45th percentile of the superimposed signal.

Peak of the P-wave
The P-peak is defined to be located at the time of the maximal value in the ECG signal before the Q-peak (Figure 1).

Peak of the Q-wave
Before the R-peak, the time of the last intersection of the ECG signal with the isoelectric line is searched. Within a predefined search window, Q is defined to be the local minima. If no local minima are available, Q is defined as the last elevation of the signal above the isoelectric line before the R-wave (Figure 2).

Figure 1 Single and averaged ECG waveforms as illustrated by Bisang et al. (A) Beat-by-beat ECG superimposed over a 1-minute period. (B) 1-minute average ECG waveform with algorithm-based ECG marker placement. P, P-wave, Q_start, the beginning of the Q-wave and start of the QT interval; T, T-wave peak; T_end, end of T-wave and end of the QT interval, set at the connection point between the tangent of the steepest downslope of the T-wave and the isoelectric line.
Peak of the R-wave
To detect the peak of the R-wave, the first derivative of the ECG signal is calculated. This first derivative is multiplied by the ECG signal itself, resulting in a signal that allows reliable detection of the R-wave.

Peak of the S-wave
The peak of the S-wave is defined to be located at the time of maximal (positive) curvature between the R-peak and the middle of the superimposed signal. The middle of the superimposed signal is approximately the middle of the RR-interval (Figure 3).

Start of the QRS-Complex
If the ECG signal level at the Q-peak is above the isoelectric line, the start of the QRS-complex is defined to be the Q-point (Figure 2B). Elsewise, it is defined to be at the time of the last local minimum of curvature in the ECG signal before the R-peak (Figure 2A).

Figure 2 Algorithm-based detection of P-waves (P), beginning of the QRS-complex (QRS_b), Q-waves (Q), and R-waves (R). (A) Shows the mode of Q-waves detection in cases where it is located below the isoelectric line (thin red line), (B) Shows the mode of Q-waves detection in cases where it is located above the isoelectric line (thin red line). This is achieved by detecting the maximal positive slope preceding the R-wave.

Figure 3 Algorithm-based detection of S-waves (S) and endings of the QRS-complex (QRS_end). (A) Shows the mode of QRS_end detection as intersection with the isoelectric line (thin red line), (B) Shows the mode of QRS_end detection as the first maximal negative slope after the S-wave.
End of the QRS-Complex
There are two situations, referred to as A and B, indicating the end of the QRS-complex. Situation A is the second intersection of the ECG signal with the isoelectric line after the R-peak (Figure 3A). If this does not happen, the end of the QRS-complex is defined as the first local maximum of curvature in the ECG signal after the S-peak (Figure 3B).

Peak of the T-wave
The peak of the T-wave is defined to be located at the time of the maximal ECG signal level after the S-wave. To prevent artefacts of the next P-interval from being detected as T-wave peaks, the search range is limited to the middle of the S-peak and the end of the superimposed ECG signal.

End of the T-wave
First, the first inflection point (minimum of the first derivative) after the T-peak is searched. At this point, the tangent to the signal is calculated. The time where this tangent intersects with the isoelectric line is defined as the end of the T-wave (Figure 1).

QT and QTc-interval
QT-interval is defined as the time between the beginning of the Q-wave and the end of the T-wave (Figure 1). QT-interval time is converted to QTc using Bazett’s transformation.\(^{26}\)

QTc Hand-Scoring
A single investigator, blinded to the results of the algorithm scoring, rated all ECGs. The assessor used cursor-assistance for scoring QTc intervals for each averaged 1-min QRS-complex in ECG Lead II. QT and QTc were defined as outlined above.

Statistical Analysis
All results are shown as median (25\(^{th}\), 75\(^{th}\) percentiles) unless stated otherwise. Variable distributions were determined visually using quantile–quantile plots. A two-sided p-value <0.05 was considered statistically significant.

We pooled the QTc data if the same patient had several recordings. Criterion validity was calculated using Bland-Altman statistics. To quantify the accuracy of clinically relevant QTc prolongations, we used confusion matrixes on the three clinically established thresholds (460, 480, and 500 ms).\(^{6,7}\) For the interpretation of clinical usability, we compared the results to the minimal clinically relevant difference, exceeding within-individual variability, which is determined at 60 ms.\(^{27}\) We applied the following formula for calculations of test performance metrics:

\[
\text{Accuracy} = \frac{\text{true positives} + \text{true negatives}}{\text{N recordings}}
\]

\[
\text{Sensitivity} = \frac{\text{true positives}}{\text{true positives} + \text{false negatives}}
\]

\[
\text{Specificity} = \frac{\text{true negatives}}{\text{false positives} + \text{true negatives}}
\]

\[
\text{Positive predictive value} = \frac{\text{true positives}}{\text{true positives} + \text{false positives}}
\]

\[
\text{Negative predictive value} = \frac{\text{true negatives}}{\text{false negatives} + \text{true negatives}}
\]

In addition, absolute numbers are reported using 2×2 tables.

All statistical analyses were done using R 4.2.0 for Windows (R Core Team 2022, R Foundation for Statistical Computing, Vienna, Austria).
Study Patients and ECG Data
Of the 32 patients included in the main trial, we included 28 for the present analysis. The included patients were 66 years (62, 69) of age, mainly female (61%), were diagnosed with COPD GOLD grade 2 (79%) or GOLD grade 3 (21%), and a majority were non-smokers (68%). Additional patient characteristics are presented in Table 1. The patients provided 62 overnight ECGs from which a total of 32944 1-min intervals were available and were analysed. Eighteen patients showed prolonged QTc in at least one of the ECGs. Average QTc as determined by hand-scoring was 419 ms (399, 440), and as determined by the algorithm 421 ms (402, 441).

Criterion Validity of Algorithm Vs Hand-Scored QTc
The mean difference between the hand-scored and the algorithm-scored QTc intervals was −1 ms (95% confidence interval [95% CI] −2 to −1) with limits of agreement from −18 to 16 ms. The result is displayed as a Bland-Altman plot with a complementing histogram in Figure 4.

Diagnostic Accuracy of Algorithm Vs Hand-Scored QTc
Regarding the diagnostic accuracy, a total of 2587, 357, and 0 QTc-intervals exceeding the thresholds 460, 480, and 500 ms, respectively, were identified by hand-scoring. Of these, 2516 (97%), 357 (100%), and 0 (100%) were consistently identified by the algorithm. This resulted in a diagnostic classification accuracy of 0.98 (95% CI 0.98 to 0.98), 1.00 (1.00 to 1.00), and 1.00 (1.00 to 1.00) for 460, 480, and 500 ms, respectively. Sensitivity was 0.97, 1.00, and NA for 460, 480, and 500 ms, respectively. Specificity was 0.98, 1.00, and 1.00 for 460, 480, and 500 ms, respectively. Corresponding positive predictive value was 0.83, 0.91, and NA for 460, 480, and 500 ms, respectively; and negative predictive value was 1.00, 1.00, and NA. The absolute number of QT intervals is displayed as 2×2 tables stratified by diagnostic thresholds in Table 2.

Discussion
We report on the criterion validity and diagnostic accuracy of a QTc scoring algorithm in overnight ECGs from individuals with stable COPD. The algorithm shows very high validity and accuracy as compared to the gold standard of hand-scoring. These results may reduce time needed for QTc determination dramatically. Therefore, this algorithm could serve as a simple and cheap tool to screen patients with elevated risk for cardiac repolarization disturbances (eg, obstructive sleep apnoea, COPD, and obesity) undergoing sleep assessments in sleep laboratories.

Table 1 Demographic Characteristics (n = 28)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, years</td>
<td>66 (62, 69)</td>
</tr>
<tr>
<td>Sex, female/male (%)</td>
<td>17/11 (61/39)</td>
</tr>
<tr>
<td>Body-mass index, kg/m²</td>
<td>25.1 (22.4, 28.3)</td>
</tr>
<tr>
<td>GOLD stage, 2/3 (%)</td>
<td>22/6 (79/21)</td>
</tr>
<tr>
<td>Forced expiratory volume in one second, % predicted</td>
<td>56 (51, 63)</td>
</tr>
<tr>
<td>Forced expiratory volume in one second, liters</td>
<td>1.6 (1.3, 2.1)</td>
</tr>
<tr>
<td>Forced vital capacity, % predicted</td>
<td>89 (78, 97)</td>
</tr>
<tr>
<td>Forced vital capacity, liters</td>
<td>3.0 (2.7, 3.4)</td>
</tr>
<tr>
<td>Smoking status, yes/no (%)</td>
<td>9/19 (32/68)</td>
</tr>
<tr>
<td>Ex-smoker, yes/no (%)</td>
<td>12/17 (43/57)</td>
</tr>
<tr>
<td>Smoking, pack years</td>
<td>42 (30, 60)</td>
</tr>
<tr>
<td>QT prolonging drug, yes/no (%)</td>
<td>3/25 (11/89)</td>
</tr>
<tr>
<td>QT prolonging comorbidity, yes/no (%)</td>
<td>4/24 (14/86)</td>
</tr>
</tbody>
</table>

Notes: Data are median (25th, 75th percentiles) or numbers (proportion). GOLD, Global Initiative for Chronic Obstructive Lung Disease. Stage 2, moderate COPD; Stage 3, severe COPD.
Regarding criterion validity, the algorithm has a bias of −1 ms. In other words, the algorithm overestimates QTc with a mean deviation of 1 ms. This is a very slight error, considering that clinically relevant QTc change exceeding within-subject variability is 60 ms.27 Consulting the Bland-Altman plot, we conclude that bias remains constant across QTc spectra and the histogram points out that a majority of measurements agree between the methods. However, at the bottom right of the Bland-Altman plot there is a cluster of QTc that shows no agreement. These QTc are from a single individual and exhibit an overestimation of QTc by the algorithm of >120 ms, which is unacceptable. In this case, the ECG recording was of very poor quality, as noted by the investigator performing the hand-scoring. Accordingly, recordings with poor signal quality (eg, due to sweating, body hair, movement artefacts) should still be evaluated manually to avoid misclassifications. We performed a sensitivity analysis, investigating the effect of this outlier on the overall performance.

Table 2 Diagnostic Accuracy of the QTc Algorithm Vs Hand-Scoring

<table>
<thead>
<tr>
<th>Hand-scoring</th>
<th>Computer-based algorithm</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>&lt;460 ms</td>
</tr>
<tr>
<td>&lt;460 ms</td>
<td>29,829</td>
</tr>
<tr>
<td>≥480 ms</td>
<td>32,552</td>
</tr>
<tr>
<td>&lt;500 ms</td>
<td>32,910</td>
</tr>
</tbody>
</table>

Notes: Date are numbers, describing the number of averaged QTc intervals over 1-minute periods.
Excluding the outlier did not change overall results, but narrowed the 95% confidence interval slightly, −1 ms (−1 to −1). The limits of agreement also narrowed slightly, ranging from −16 to 13 ms.

Validity and diagnostic accuracy of our algorithm outperforms commercial device-integrated ones.20,21 We identified the strength of the presented algorithm in the robust detection of QRS-complex beginnings and endings. It is frequent that QRS-complexes do not match the isoelectric line optimally. Therefore, our algorithm implemented alternative automated detections in these cases, see Figures 1 and 3.

Our findings allow the conclusion that the algorithm performs with very high accuracy compared to hand-scoring. This in both numeric comparison and diagnostic accuracy. However, caution is advised in highly pathological (ie, >500 ms) QTc. In our sample, no patient exhibited such a high QTc as determined by hand-scoring, see Table 2. Thus, no sensitivity, positive and negative predictive values were calculated. Accordingly, we would still encourage confirmation by hand-scoring, if the algorithm indicates QTc exceeding 500 ms. Nevertheless, the Bland-Altman plot shows very constant results across QTc spectra, and we therefore expect the algorithm to perform equally in highly pathological ranges. A confirmation of this expectation with a study on individuals with suspected highly pathological QTc is encouraged.

The routine implementation of QTc screening in hospital sleep laboratories may improve the identification of cardiovascular comorbidity in individuals with COPD. Since LQTS is associated with increased mortality in patients with COPD,2 routine screening might identify asymptomatic individuals with LQTS and enable early treatment. Patients with COPD usually present with several internal comorbidities and are prone to ingest various pharmacologic agents.28 Thus, the underlying cause of a diagnosed LQTS must be closely evaluated. In addition, the research-based application of the algorithm may contribute to a better understanding of LQTS in numerous cardiopulmonary disorders. Most of the studies in these populations were performed on selected, moderately sized samples and applied varying methods of QTc quantification.

In our sample, 18 patients were found to have LQTS, 3 were taking drugs known to prolong QTc (ie, amiodarone in one patient and escitalopram in two patients), and 4 were diagnosed with a comorbidity known to prolong QTc (ie, rheumatoid arthritis in one patient and type-2 diabetes in three patients). Although the number of patients in our study taking a drug or having a comorbidity prolonging QTc is relatively low, this emphasises on the multifactorial nature of LQTS in COPD and the importance of systematic screening and clinical evaluation, eg, when obtaining a sleep exam.

We applied the algorithm to overnight, sleep-study related ECGs. These are generally of high quality, suffering rarely from movement-related noise. We validated each QTc against the corresponding one in the gold standard hand scores. Therefore, we conclude that our results would remain unchanged when applied to standard resting or Holter ECGs.

This study has some limitations. First, we included individuals with a relatively homogenous disease status (ie, GOLD stage 2 or 3). This results in limited generalizability on more severe COPD and other cardiopulmonary disease populations with potentially higher prevalence of LQTS. Additional studies validating our algorithm in distinct populations are encouraged. Second, as discussed in the preceding paragraph, values exceeding 500 ms would still need hand-scored confirmation. Finally, although largely similar and compatible, there may be ECG devices that output curves in a format that is not readable by our algorithm.

Conclusions

In conclusion, the presented automated algorithm accurately identified clinically relevant QTc-prolongations in a relatively large sample of individuals with COPD undergoing sleep studies. Accuracy is very high, and we therefore conclude that the algorithm can replace hand-scoring in the vast majority of cases. The implementation of this tool in hospital sleep laboratories may identify asymptomatic patients with long-QT at risk for malignant arrhythmia, allowing them to consult a cardiologist before an eventual cardiac event.

Data Sharing Statement

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.
Acknowledgments

This paper abstract was presented at the 2023 ERS International Congress as a poster presentation with interim findings. The poster’s abstract was published as a supplement in the European Respiratory Journal: https://erj.ersjournals.com/content/62/suppl_67/PA2432.

Funding

The study was supported by the Swiss National Science Foundation (143875) and Lunge Zurich. Siemens Health Engineers provided some equipment for the study.

Disclosure

The authors report no conflicts of interest in relation to this work.

References


