Available online at www.sciencedirect.com
ScienceDirect
Journal of Electrocardiology 39 (2006) S152–S156
JOURNAL OF Electrocardiology
www.elsevier.com/locate/jelectrocard

Implications of methodological differences in digital electrocardiogram interval measurement☆

Fabio Badilini, PhD,[a,*] Nenad Sarapa, MD[b]
[a]Exec. VP, AMPS-LLC, New York, NY 10025, USA
[b]Exec. Dir, Translational Medicine, Daiichi-Sankyo Pharma Development, Edison, NJ 08837, USA
Received 10 May 2006; accepted 15 May 2006

Abstract Well-specified recommendations have yet to be established on how electrocardiogram (ECG) interval measurement should be performed by digital on-screen caliper systems to assess drug-induced effect on cardiac repolarization in pharmaceutical clinical trials with adequate precision and reproducibility. Since 1997, the industry has followed the European Committee for Proprietary Medicinal Products Points to Consider by using fully manual measurement of 3 consecutive sinus rhythm PQRST complexes in 1 lead only (typically limb lead II). More recently, semiautomatic measurement performed on representative (median) beats and based on the global leads has been considered. The International Conference on Harmonization E14 guidance (June 2005) advocates development of quality standards for centralized ECG interval measurement and allows all methods "whether or not assisted by computer" but includes no recommendations on how to perform the measurement. We provide an overview of the currently available methods for digital ECG interval measurement and the implications of between-method differences on quality of ECG interval measurements. We applied 4 methods most commonly used to assess QT prolongation (applied on 3 raw beats in limb lead II or by global measurement on 1 or 12 superimposed representative beats). QT, QTc Fridericia, and RR interval durations were measured on resting 12-lead digital ECGs obtained in 26 healthy volunteers predose and at 1, 2, and 3 hours after dosing with a single 160 mg oral dose of sotalol. Absolute interval durations and changes from baseline were compared between the 4 measurement methods. A better understanding of the implications from different measurement methodologies will facilitate more informed choice of the appropriate method for ECG interval measurement on clinical trials.

Keywords: QT interval; Clinical trials

Introduction

The Food and Drug Administration's Digital ECG Initiative from 2001 mandates that, for new drug approvals, digital electrocardiograms must be submitted from definitive ("thorough") QT studies and that the interval measurements be performed with annotations detailing exact offset and onset points on the ECG.[1] In consequence, digital ECG tracings and on-screen calipers have replaced paper ECG printouts and digitizing board as the primary tools for ECG acquisition and interval measurement in intensive QT assessment in clinical trials.[2,3] The new digital ECG environment has multiple important advantages over paper ECG for the investigating site, the core ECG laboratory, the sponsor, and the regulatory agencies, offering improvements of transmission, management, measurement, storage, and review of ECG data. Very important among these advantages is a completely new and improved approach to measure the intervals and evaluate waveform morphology on digital ECG. The only written recommendations for ECG interval measurement widely accepted before the digital era were published in 1997 by the European Committee for Proprietary Medicinal Products (CPMP) and were based on annotating 3 consecutive sinus complex, preferably from lead II.[4] At that time, detection of drug effects on cardiac repolarization was mostly exclusively based on paper ECG and was associated with considerable degree of variability and measurement errors.[5] Usage of

☆ Sources of funding: Grant to AMPS by Daiichi Medical Research, a legacy company of Daiichi Sankyo Pharma Development.
* Corresponding author. 25018 Montichiari (BS) Italy. Tel.: +39 030 9650745; fax: +39 030 9650572.

0022-0736/$ – see front matter © 2006 Elsevier Inc. All rights reserved.
doi:10.1016/j.jelectrocard.2006.05.030

manual digitizing board measurement methods was capable to detect measurement variations in the range of milliseconds[6,7]; yet, the precision of the measurements was still relatively poor.[8]

The introduction of on-screen methodologies based on digital ECGs has completely changed the measuring environment. For example, the potential advantages of implementing digital algorithms are now being considered.[9] Consequently, pharmaceutical sponsors nowadays commonly use semiautomated methods for centralized ECG interval measurement, where a trained human analyst decides whether the ECG interval annotations by the automated algorithm should be adjusted based on visual inspection of annotated waveforms on a computer screen. This approach potentially combines consistency of the automated interval measurement with the added precision of manual adjustment, although no data have been published thus far on the performance of semiautomated method.

Another important opportunity offered by digital ECGs is the possibility to perform measurements on the so-called representative beats, often simply referred to as median beats, generally available as part of the digital ECG source file. The concept of representative beats is well known in academic research for many years,[10] but it has been only recently considered as a viable alternative to generate reliable and reproducible interval measurements on ECG from clinical trials that may even have advantages over the traditional CPMP-recommended approach.[11]

From "HOW TO" to "WHERE TO" measure QT intervals

Great attention has been placed on the problem of "how to" measure the QT interval. Although a unique method to define where the end of the QT interval should be annotated is far from being accepted, a number of systems have been widely tested and validated and are commonly used in clinical trial practice.

An equally important aspect that has not received proper attention is "where to" measure the QT interval. The only official guideline on this matter is the CPMP Points to Consider from 1997 that recommend centralized manual measurement of 3 consecutive sinus rhythm complexes in only 1 lead (typically limb lead II). This document addressed the measuring context of a standard paper printout generated by ECG machines (according to different display formats), where the analyst was manually placing calipers with or without the use of a digitizing board, and in some cases with assistance of a magnifying lens.[4]

Today, the measuring context for which the CPMP guidelines had been written has completely changed and virtually all core laboratories use computer systems and deal with digital ECG. However, because of the absence of new guidelines (the recent International Conference on Harmonization E14 document does not address this specific point[12]), the ECG interval measurement in the majority of studies is still being carried out with the "three consecutive beats from lead II approach." No data are available on the comparison between the global QT measurement on median beats and the single lead-based measurement.

On-screen environment points to consider

Display organization

Unlike paper ECG, digital environment offers multiple modalities for display and organization of waveforms on a computer screen with individually specific implications on the precision of interval measurement. These include simple but important factors like the size and resolution of computer screen, background ECG grid features (width between the thick and thin lines), number of leads and beats per lead displayed on screen during measurement, and the pixel-to-sample ratio. Although no data are available, common sense indicates that the bigger diagonal width (eg, 21 vs 17 in) and higher resolution of the computer screen would allow for greater precision of on-screen interval measurement, and that displaying just 1 lead instead of all 12 would facilitate better measurement. Because core ECG laboratories use different equipment with diverse default settings, sponsors should define their own standard requirements and request that they be used by all analysts performing on-screen interval measurement on a given study. Ideally, the same settings will be used across all studies with intensive QT assessment on a drug development program.

Another important aspect very often underestimated is the display relationship between the computer screen display elements (the pixel) and the digital samples actually drawn at any given time. The pixel-to-sample (PS) ratio is a property inherent to digital ECG environment that can significantly affect the outcome of the on-screen interval measurement in manual or semiautomated methods.

Indeed, if PS ratio is less than 1.0, then the analyst performing on-screen measurement cannot reach the *intrinsic* resolution of the digital ECG (2 milliseconds for 500 Hz sampling rate typical of standard electrocardiographs). If PS ratio is more than 1.0, the analyst can match the intrinsic resolution because all of the digital samples are displayed on screen during the measurement. Modern on-screen caliper tools (eg, CalECG-2 from AMPS-LLC, New York, NY) allow the analyst to be continuously aware if PS ratio is maintained above 1.0 during measurement. As a general rule, a larger PS ratio is more favorable, but the exact optimal value is uncertain. With PS greater than 1.0, the analyst can theoretically place on-screen calipers between samples and achieve resolution smaller than 2 milliseconds. To ensure consistency between analysts, sponsors should prescribe to core ECG laboratories if the analyst is allowed to do this.

Global vs lead-based measurement

One of the most important features of the on-screen interval measurement is the type of measurement (global vs single-lead) and the properties of the ECG signal that either type is applied on (representative/median vs raw beats).

Global measurement is typically performed on PQRST complexes from all 12 leads superimposed onto the joint isoelectric line. Such global measurement by the electrocardiograph's automated algorithm applied on a single representative beat (ie, the mathematical model of the electrical activity during the whole cardiac cycle, also called

the "median" beat) is the standard approach to interval measurement in routine health care. Global approach can also be applied to the superimposed raw beats, although the higher noise of the electrical signal would decrease the precision of the interval measurement. Single-lead interval measurement was historically most commonly performed in the limb lead II, where the morphology of the T wave is expected to be most clearly demarcated. Fully manual interval measurement on 3 consecutive raw sinus beats in lead II has been used in the majority of clinical trials to date that characterized drug effects on cardiac repolarization. Interval measurement could also be done on 1 or more representative/median beats from a chosen single lead, but it has not been used in clinical trials to date. By means of reducing the noise, this approach might improve the precision of single-lead measurement. Clearly, the outcome of QT assessment might be different when each approach is applied to the same ECG data. Global and single-lead measurements have hitherto been compared only within the context of fully automated and fully manual techniques.[13] Whereas both global and single-lead measurement on representative or raw beats can be used for the semi-automated interval measurement, the best approach to lead selection and display of beats during interval measurement in the intensive QT assessment has yet to be determined.

Comparing methods

We report here the summary results of a comparative methodological analysis performed on 104 ECGs from 26 normal subjects from a previously reported study.[14] Each subject had 1 resting supine 10-second ECG taken at baseline and at 2, 3, and 4 hours after the intake of a single dose of 160 mg of sotalol. Electrocardiograms were obtained by ELI200 electrocardiograph (Mortara Instrument, Inc., Milwaukee, Wis) after 5 minutes of quiet rest in fully supine position. All ECGs consisted of 10-second raw data plus the set of representative beats, as computed by internal ELI200 algorithm. The ECGs were analyzed by a single reader using CalECG2.1.0 on-screen caliper system (AMPS-LLC) using the following 4 measurement methods (Fig. 1):

- Method 1 (M1): QT and RR are average from 3 consecutive sinus rhythm raw beats in lead II
- Method 2 (M2): Global QT/RR from one PQRST complex composed of 12 median (representative) beats superimposed from 12 leads
- Method 3 (M3): QT/RR from median (representative) beat in lead II
- Method 4 (M4): QT and RR are average from 3 consecutive beats composed of 12 superimposed raw signals from 12 leads

Method 1 is the method recommended by the CPMP guidelines. In our scenario, the CalECG internal algorithm preplaced the annotations and the reader adjusted their position as needed.

Method 2 is the typical semiautomated approach applied to representative (median) beats whereby the 12 representative waveforms are displayed superimposed and a single set of annotations automatically computed by CalECG is

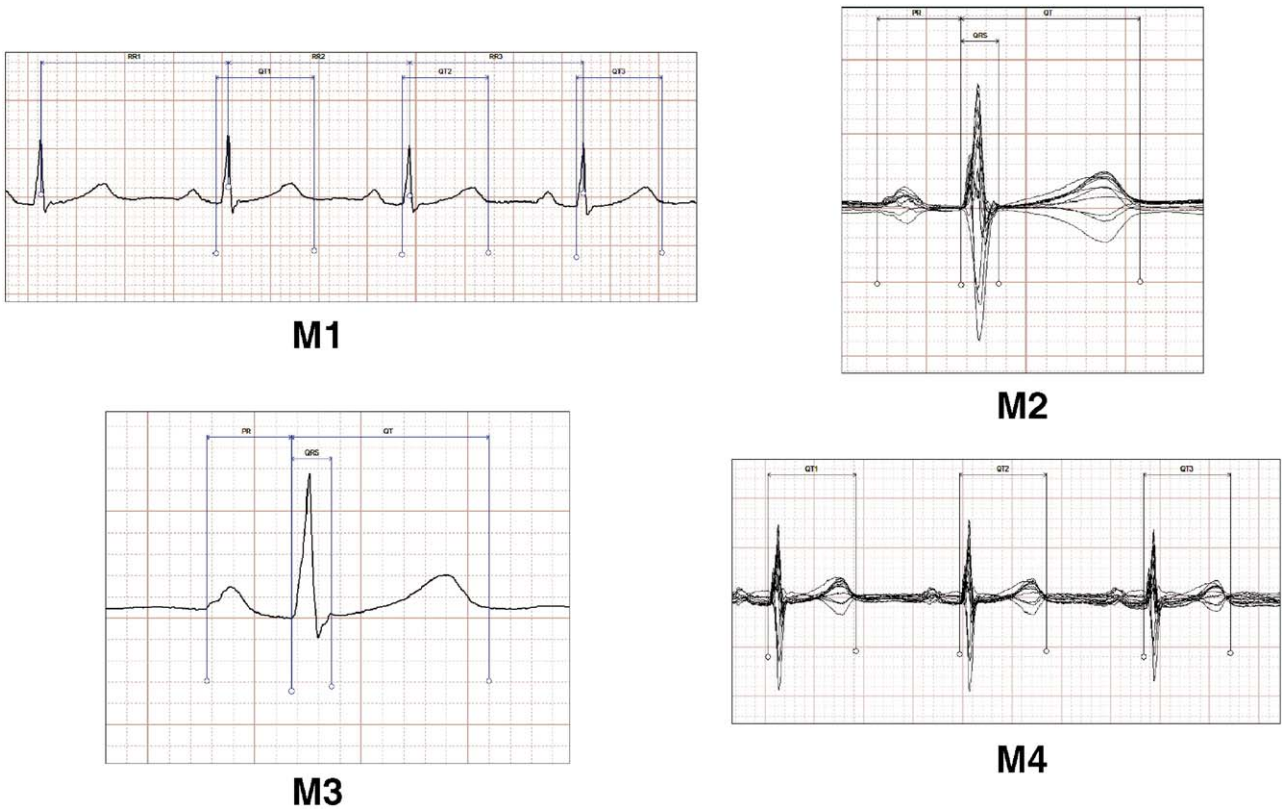

Fig. 1. Examples of methods used. M1: QT and RR are average from 3 consecutive sinus rhythm raw beats in lead II; M2: global QT/RR from one PQRST complex composed of 12 median (representative) beats superimposed from 12 leads; M3: QT/RR from median (representative) beat in lead II; M4: QT and RR are average from 3 consecutive beats composed of 12 superimposed raw signals from 12 leads.

displayed on screen. There is no unique definition for the global T-offset annotation. Some methods take the latest T-offset position of the 12 individual representative beats; other methods (like that implemented in CalECG2) derive first an integrated waveforms from the 12 leads (a variant of the vector magnitude) and then compute the different markers from the integrated waveform. The integrated waveform is generally not visible to the user who base his/her review on other criteria than those used by the algorithm. In this study, the reader was required to manually adjust the T-wave annotation whenever a mistake made by the automated algorithm was apparent.

Method 3 was based on the representative beat in limb lead II. Method 4 applied the concept of global annotation to the 3 consecutive sinus beats using raw signal from all 12 leads superimposed. Both M3 and M4 are very rarely used in the analysis of ECG interval duration in clinical trials.

Each of the 4 sets of interval measurements on 104 ECGs was performed by the same reader but at least 3 weeks apart and the on screen display environment was kept constant throughout the analysis. For each method, a semiautomated approach was used: CalECG2 prepositioned the annotations and the reader adjusted the caliper positions whenever he/she noticed a discrepancy between his/her visual assessment of the optimal placement and the annotations placed by the algorithm. For the 2 methods using 3 consecutive sinus beats within the 10-second raw data (M1 and M4), the RR and QT intervals were the averages of the 3 individual RR/QT pairs. For the 2 methods using the representative beats (M2 and M3), a single RR interval based on the whole 10 seconds ("global RR") and a single QT interval were measured. Fridericia's correction formula was applied to each QT/RR pair to derive the corrected QT interval (QTc).

Fig. 2 shows the between-method differences in absolute QTcF duration. As expected, different methods resulted in different QTcF duration but the between-method differences were small. The QTcF measured by the global method applied on 12 median beats (M2) was consistently longer

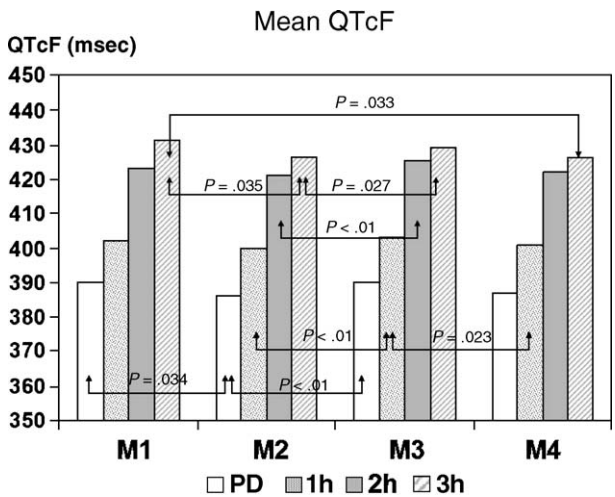

Fig. 2. Means of QTc Fridercia plots obtained with the 4 methods at baseline, P1, P2, and P4. Significant *P* values (by paired Student *t* test) are also reported.

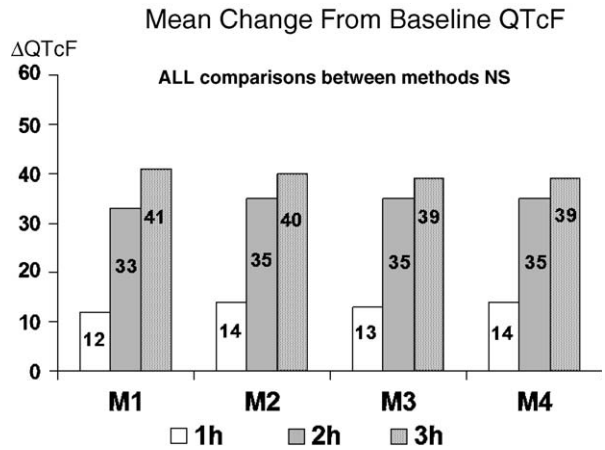

Fig. 3. Mean changes from baseline of QTc Fridericia with the 4 methods.

than the QTcf by the most commonly used method on 3 consecutive raw beats in lead II (M1). This difference was statistically significant, but it would not have clinical implications because of the comparable statistics for QTcF change from baseline.

Fig. 3 presents between-method comparisons of the change from baseline in QTcF for each of the postdose time points. QTcF changes detected by all methods were comparable with no differences reaching statistical significance.

Despite small but statistically significant between-method differences in absolute QTcF, detection of a prominent drug effect on cardiac repolarization (QTcF change from baseline induced by sotalol) was independent of the ECG interval measurement methodology. If confirmed for other types of drug-induced changes, our results would diminish the concerns around the choice of optimal method of QT interval measurement in clinical trials.

In a separate study, we will investigate the intrinsic variability of QT interval measurement by each method. If methodological consistency is preserved, any method will be capable to adequately detect prominent QTc prolongation induced by drugs in development. By modifying the morphology of repolarization, sotalol can affect the precision of QT interval measurement by methods such as the tangent approach, which is very sensitive to changes in T-wave morphology. However, the 4 methods compared in our study were well balanced with respect to the potential influence of abnormal T-wave morphology.

Given the observed equivalence in detecting sotalol-induced QTc prolongation, M2 based on representative waveforms might be preferred because of its intrinsic better quality (less noise than raw signal). Furthermore, semi-automated interval measurement on representative beats requires only the review (and if necessary, adjustment) of annotations on a single PQRST complex (instead of 3 beats in limb lead II) and therefore has potentially lower intrareader and interreader variability and improved speed of analysis. One potential negative aspect of methods based on representative beats is that the RR interval used is generally automatically computed from the 10-second raw data by the ECG machine and cannot be visually reviewed. Because of this, it is very important to attempt to take

10-second ECGs at stable heart rate so that the global RR measurement is more credible.

Conclusions

Our study is the first to compare digital on-screen measurements of QT and RR intervals by 4 different semiautomated methods. Although small but statistically significant differences between individual methods were observed for the absolute QTcF duration, the QTcF change from baseline induced by sotalol at the time of peak concentration in plasma 1 to 4 hours after a single 160-mg dose was equivalent for all 4 methods. Intrinsic variability of each method (to be evaluated in a subsequent analysis) will define the utility of each individual method for precise detection of drug-induced QTc prolongation in clinical trials.

With the introduction of digital ECG signal processing, the CPMP guideline has become obsolete. New recommendations are warranted to define a working environment for digital on-screen ECG interval measurement so that the higher precision offered by electronic tools can be coupled with optimal consistency of measurement of drug-induced QTc prolongation in clinical trials.

Such guidelines will help the industry to better understand the risk of methodological biases and be more successful in the quest for reliable markers of drug-induced arrhythmias.

References

1. Center for Drug Evaluation and Research (CDER). Digital FDA initiative. http://www.fda.gov/cder/regulatory/ersr/default.htm#ECG.
2. Stockbridge N, Throckmorton DC. Regulatory advice on evaluation of the proarrhythmic potential of drugs. J Electrocardiol 2004; 37(Suppl):40.
3. Stockbridge N, Brown BD. Annotated ECG waveform data at FDA. J Electrocardiol 2004;37(Suppl):83.
4. Points to consider: the assessment of the potential for QT interval prolongation by non-cardiovascular medicinal products. London: Committee for Proprietary Medicinal Products (CPMP); 1997.
5. Murray A, McLaughlin NB, Bourke JP, Doig JC, Furniss SS, Campbell RWP. Errors in manual measurement of QT intervals. Br Heart J 1994;71:386.
6. Morganroth J, Silber SS. How to obtain and analyze electrocardiograms in clinical trial. ANE 1999;4:425.
7. Pratt CM, Ruberg S, Morganroth J, et al. Dose-response relation between terfenadine (Seldane) and the QTc interval on the scalar electrocardiogram: distinguishing a drug effect from spontaneous variability. Am Heart J 1996;131:472.
8. Malik M, Bradford A. Human precision of operating a digitizing board: implications for electrocardiogram measurements. PACE 1998; 21:1656.
9. McLaughlin NB, Campbell RWF, Murray A. Comparison of automatic QT measurement techniques in the normal 12-lead electrocardiogram. Br Heart J 1995;74:84.
10. Savelieva I, Yi G, Guo X, Hnatkova K, Malik M. Agreement and reproducibility of automatic versus manual measurement of QT interval and QT dispersion. Am J Cardiol 1998;81:471.
11. Darpo B, Agin M, Kazierad DJ, et al. Man versus machine: Is there an optimal method for QT measurement in thorough QT studies? J Clin Pharmacol 2006;46:598.
12. The clinical evaluation of QT/QTc interval prolongation and proarrhythmic potential for non-antiarrhythmic drugs. International Conference on Harmonization (ICH), London, May 25, 2005.
13. Azie NE, Adams G, Darpo B, et al. Comparing methods of measurement for detecting drug-induced changes in the QT interval: Implications for the thoroughly conducted ECG studies. Ann Noninvasive Electrocardiol 2004;9:166.
14. Sarapa N, Morganroth J, Couderc JP, et al. Electrocardiographic identification of drug-induced QT prolongation: assessment by different recording and measurement methods. Ann Noninvasive Electrocardiol 2004;9:48.