European Heart Journal (1996) **17**, 1035–1039

Reproducibility and automatic measurement of QT dispersion

J. M. Glancy, P. J. Weston, H. K. Bhullar*, C. J. Garratt, K. L. Woods and D. P. de Bono

*Department of Medicine and Therapeutics, University of Leicester, Leicester, U.K. and *Department of Engineering, University of Leicester, Leicester, U.K.*

This study investigated interobserver (two observers) and intrasubject (two measurements) reproducibility of QT dispersion from abnormal electrocardiograms in patients with previous myocardial infarction, and compared a user-interactive with an automatic measurement system. Standard 12-lead electrocardiograms, recorded at 25 mm . s^{-1}, were randomly chosen from 70 patients following myocardial infarction. These were scanned into a personal computer, and specially designed software skeletonized and joined each image. The images were then available for user-interactive (mouse and computer screen), or automatic measurements using a specially designed algorithm. For all methods reproducibility of the RR interval was excellent (mean absolute errors 3–4 ms, relative errors 0·3–0·5%). Reproducibility of the mean QT interval was good; intrasubject error was 6 ms (relative error 1·4%), interobserver error was 7 ms (1·8%), and observers' vs automatic measurement errors were 10 and 11 ms (2·5, 2·8%). However QTc dispersion measurements had large errors for all methods; intrasubject error was 12 ms (17·3%), interobserver error was 15 ms (22·1%), and observers' vs automatic measurement were errors 30 and 28 ms (35·4, 31·9%). QT dispersion measurements rely on the most difficult to measure QT intervals, resulting in a problem of reproducibility. Any automatic system must not only recognize common T wave morphologies, but also these more difficult T waves, if it is to be useful for measuring QT dispersion. The poor reproducibility of QT dispersion limits its role as a useful clinical tool, particularly as a predictor of events.

(Eur Heart J 1996; 17: 1035–1039)

Key Words: QT dispersion, reproducibility, automatic measurement.

Introduction

QT dispersion (maximum minus minimum QT interval) on a surface electrocardiogram may reflect the degree of dispersion of ventricular repolarization, and identify potential re-entry circuits for ventricular tachyarrythmias[1]. QT dispersion has been proposed as a useful predictor of sudden death in patients with chronic heart failure[2], and of ventricular arrhythmias in hypertrophic cardiomyopathy[3]. In other studies whilst QT dispersion measurements may have provided insight into mechanisms of arrhythmias, the overlap in QT dispersion measurements between controls and patients suggest a much more limited role as a useful predictor of events[4,5]. For any clinical measurement to be useful, particularly as a predictor, it must be reproducible. In normal electrocardiograms the reproducibility of QT dispersion has been questioned, with large relative errors for both interobserver, and intrasubject measurements[6]. This may have been partly the consequence of the absolute values for QT dispersion measurements being small in normal subjects, and therefore small absolute errors would convert into large relative errors. To address this problem we studied reproducibility of QT dispersion measurements in abnormal electrocardiograms, from patients with previous myocardial infarction. An answer to the problem of reproducibility would be a 100% accurate automated system. We also compared the two observers' measurements with an automatic algorithm, specially designed for QT interval measurement.

Methods

Patients with confirmed acute myocardial infarction were drawn from the placebo arm of the LIMIT-2 study[7], and had been admitted to Leicester Royal Infirmary between September 1987 and February 1992. Criteria for the diagnosis of acute myocardial infarction were at least two of the following: typical history; rise in

Revision submitted 15 November 1995, and accepted 21 November 1995.

This work was funded by a British Heart Foundation project grant.

Correspondence: J. M. Glancy, Department of Cardiology, The County Hospital, Union Walk, Hereford HR1 2ER.

0195-668X/96/071035+05 $18.00/0

Table 1 Intrasubject and interobserver errors of mean RR intervals, mean QT intervals, QT dispersion and QTc dispersion

	Intrasubject measurements Observation 1 minus observation 2				Interobserver measurements Observer 1 minus observer 2			
	RR interval	QT interval	QT dispersion	QTc dispersion	RR interval	QT interval	QT dispersion	QTc dispersion
Bias (ms)	0	2	4	3	0	6	7	6
Mean absolute difference (ms)	4	6	12	12	4	7	15	15
Relative difference (%)	0·5	1·4	19·8	17·3	0·5	1·8	24	22·1
Coefficient of repeatability (ms)	2	16	32	28	2	20	39	37

Table 2 Errors between observers and automatic measurement for mean RR intervals, mean QT intervals, QT dispersion and QTc dispersion

	Observer 1 minus automatic measurement				Observer 2 minus automatic measurement			
	RR interval	QT interval	QT dispersion	QTc dispersion	RR interval	QT interval	QT dispersion	QTc dispersion
Bias (ms)	0	2	22	22	0	8	14	17
Mean absolute difference (%)	3	10	28	30	3	11	24	28
Relative difference (%)	0·3	2·5	35·7	35·4	0·3	2·8	30	31·9
Limits of agreement (ms)	−16, 16	−15, 19			−10,10	−11,27		
[% for log transformed data]			−70%, 65%	−64%, 63%			−82%, 59%	−86%, 63%

serum creatine kinase to at least twice the upper limit of normal; evolving electrocardiogram changes consistent with acute infarction. A total of 135 patients had electrocardiograms available from at least 28 days post infarct. Seventy electrocardiograms were randomly chosen for analysis. Sixty of these electrocardiograms had inverted or biphasic T waves. Of the remaining 10 electrocardiograms, eight had Q waves, and eight had prominent U waves. Two electrocardiograms looked morphologically normal.

QT interval analysis was performed by two observers on 12-lead standard electrocardiograms recorded at 25 mm . s^{-1} (Marquette inkwriter, FCP-4101U, Cambridge 3038, or Hewlett Packard pagewriter). Observer 1 re-analysed the electrocardiograms at least 6 months after the first measurement. Electrocardiograms were scanned by a flatbed scanner (Hewlett Packard Scanjet Plus) interfaced with a personal computer. Each image was then cut and copied into 12 files (HP Paintbrush, Hewlett Packard) in the .PCX format, corresponding to the 12 leads of the electrocardiogram. Specially designed software skeletonized and joined each image. Each image can be magnified on the computer monitor for optimal QT interval measurement using a 'mouse'. At a recording speed of 25 mm . s^{-1}, and a scan density of 300 dots per inch, one pixel on the screen corresponds to 4 ms. Repeated estimates of one normal electrocardiogram waveform have shown errors less than 4 ms for QT interval measurements. The automatic algorithm relies on the creation of area maps for each electrocardiogram waveform, and the application of a series of rules to determine the QRS onset, and T wave offset. The rules were specifically designed to identify most common T wave morphologies. Visual monitoring of the automatic algorithm very occasionally revealed gross misinterpretation of a QT interval, such a lead was excluded from analysis. The 20 electrocardiograms with greatest error for QTc dispersion measurements between the automatic system and observers were re-analysed automatically. Observer 1 monitored the re-analysis and substituted his measurement of a QT interval if he felt the automatic system to be markedly wrong. Further details and validation of both the user-interactive and automatic systems have been described elsewhere[8]. Mean RR interval, uncorrected mean QT interval, QT dispersion, and QTc dispersion for each electrocardiogram analysed were calculated. The QT dispersion is the difference between the maximum and minimum QT

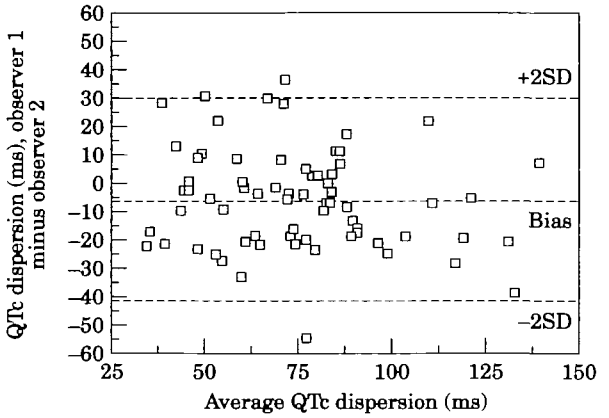

Figure 1 **Interobserver error of QTc dispersion.**

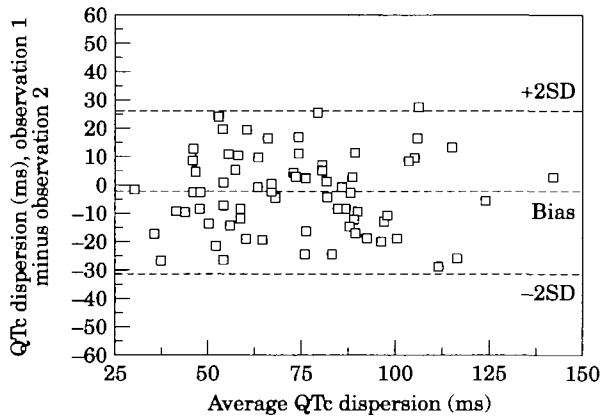

Figure 2 **Intrasubject error of QTc dispersion.**

across the 12-lead electrocardiogram. The QTc dispersion is the QT dispersion adjusted for heart rate. QT measurements were by the definitions of Lepeschkin and Surawicz[9], and rate adjustment was by Bazett's formula ($QTc = QT/RR^{1/2}$).

Statistical analysis were by methods described by Bland and Altman[10], with log transformation of data if necessary. The repeatability coefficient is the expected value below which 95% of the differences will fall for intrasubject and interobserver reproducibility. For comparison between the two methods of measurement 95% of expected differences will fall within the limits of agreement. If A and B are the repeat measurements A minus B is the absolute error. Mean absolute error is calculated by ignoring the direction of the error, for this calculation the absolute value of each error is used: $(A - B)/(A + B/2)$ is the relative error[6]. The bias is the difference between the mean measurements of sample A and sample B[10].

Results

Mean absolute errors, relative errors, and repeatability coefficients or limits of agreement for RR intervals, uncorrected QT intervals, QT dispersion, and QTc dispersion are shown for intrasubject and interobserver variability in Table 1; and for comparison between the user-interactive and automatic systems in Table 2.

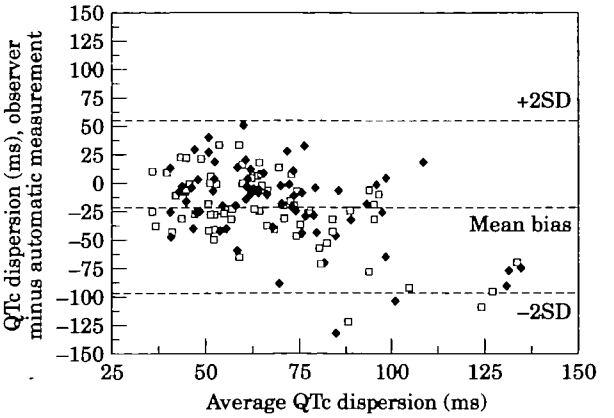

Figure 3 **Errors between observers and automatic algorithm for QTc dispersion measurements. □, observer 1 vs automatic; ◆, observer 2 vs automatic.**

Figures 1 and 2 are Bland–Altman plots for interobserver and intrasubject errors of QTc dispersion. Figure 3 is a similar plot combining the data from both observers vs the automatic system for QTc dispersion measurements.

There was a small bias of 6 ms for observer 2 to measure shorter mean QT intervals, compared to observer 1. There was bias for the automatic system to measure both longer mean QT intervals and higher QT and QTc dispersion than either observer. Greater errors of QT and QTc dispersion measurement between the automatic system and observers existed at higher average QT or QTc dispersion values.

For the 20 electrocardiograms re-analysed automatically observer 1 detected and corrected marked errors of QT interval measurements in one lead for 18 electrocardiograms, two leads for two electrocardiograms, and three leads for one electrocardiogram. Repeatability of the automatic system was confirmed as 100%. In nine leads a part of the T wave was misinterpreted as a U wave, for four leads the converse error was seen. In six leads the automatic system attempted to measure T waves too flat for accurate measurement, and in two leads misinterpreted baseline wander as a T wave. The remaining four leads had errors in the measurement of the QRS onset. Correction of these leads by observer 1 reduced the absolute error (relative error) in QTc dispersion measurements between observer 1 and the automatic system for these 20 electrocardiograms from 53 ms (53%) to 14 ms (14%). Figure 4 shows examples of leads causing greatest error for (a) the automatic system, and (b) between observers.

Discussion

Reproducibility of QT dispersion is a major problem. In normal subjects Kautzner *et al.*[6] found relative errors

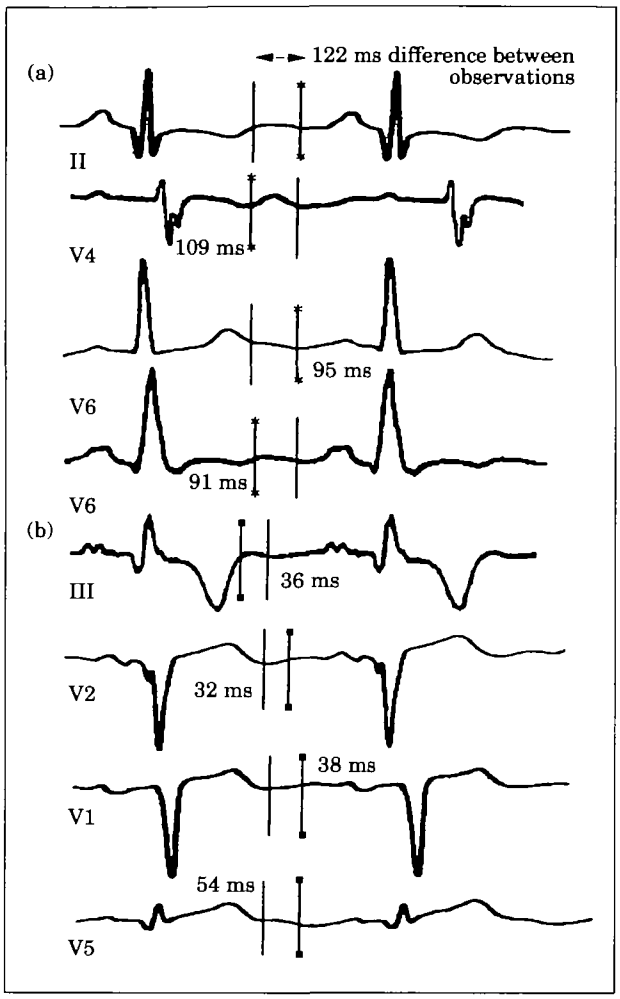

Figure 4 **Examples of electrocardiogram leads causing greatest error (a) for the automatic system, and (b) between observers. Vertical lines are T wave ends as defined by: |, observer 1; ⁞, observer 2; ‡, automatic system.**

between observers of between 26·8 and 33·2% for measures of QT dispersion. Others have shown absolute interobserver errors of up to 20 ms for QT dispersion in normal subjects, and suggested poor reproducibility[1]. The data presented concurs with those studies. In contrast van de Loo *et al.*[12] studied 77 electrocardiograms from 120 patients during acute myocardial infarction (43 patients were excluded from the study), in combination with 50 electrocardiograms from age- and sex-matched normal controls. They reported a mean interobserver error of only 7 ms for QTc dispersion, and a mean intrasubject error of 6 ms for QT dispersion in the total group studied. Unfortunately errors within each group were compared by correlation coefficients, making comparison of their population of abnormal electrocardiograms with our data difficult. They concluded their reason for good reproducibility was that electrocardiograms were recorded at 50 mm . s^{-1} and not 25 mm . s^{-1}. They compared a small subgroup of electrocardiograms at different recording speeds, again by correlation coefficients to try to confirm this hypothesis.

If the problem of reproducibility was due to paper speed recording, and therefore problems of resolution, then one would expect poor reproducibility for all QT intervals. These data and others do not show this[6]. Improved reproducibility of mean QT interval measurement has been shown with faster paper speeds in small numbers of electrocardiograms. However, the interobserver error of mean QT intervals reported here of 7 ms is less than that study's reported error in electrocardiograms recorded at 100 mm . s^{-1}[13].

Large errors in the determination of QT intervals during acute myocardial infraction have been reported between observers not specifically experienced in QT interval measurement[14]. Accurate QT interval measurement requires some experience, the end of a T wave can be difficult to determine. The definitive paper by Lepeschkin and Surawicz in 1952[9] remains the only benchmark for accuracy and repeatability of QT interval measurement, and its criteria for determining a T wave end must be adopted strictly. Even here there are problems. Some of their methods used to define T wave ends rely on the synchronicity of all T waves across a surface electrocardiogram, an idea contrary to the concept of a QT dispersion. The greatest potential source for error is the analysis of a T–U interval. If a notch or kink is present between a T and U wave then the measurement is easier. The often misquoted phrase 'if a U wave was present we defined the end of the T wave as the nadir between the T and U wave' in fact refers to the nadir of a notch if it is present. This misinterpretation from Lepeschkin's paper may be responsible for considerable error. In the abnormal electrocardiograms we studied an inverted T wave often flowed smoothly into a U wave, without any obvious notch, resulting in difficulty in determining the T wave end. When a U wave is superimposed on a T wave, e.g. the final waveform in Fig. 4, phonocardiography was used by Lepeschkin and Surawicz to define the T wave end[9]. The method of extrapolating the downslope of the T wave to the baseline would seem to be the best method of analysing such leads[15]. Where problems with T and U waves are encountered, interlead variability of QT intervals is significantly decreased by excluding such leads[16]. However, if these leads were ignored in analysis, a falsely low value of QT dispersion would be obtained. From our data it is likely that these difficult to measure QT intervals are those most likely to contribute to QT dispersion.

The advantage of a fully automated system of QT dispersion measurement is 100% reproducibility. This automatic algorithm was designed to recognize most common T wave morphologies. To this extent it appears quite successful, repeatability for mean QT intervals between this system and the two observers appear acceptable. However, recognition of rarer T wave morphologies appears poor, resulting in inaccurate measurements of QT dispersion. Methods of designing most automatic computer algorithms for QT interval

measurement rely on recognizing common T wave morphologies, and have been validated using normal electrocardiograms[17,18]. The specific problem we have encountered with our automated measurement system is probably a general one. Increasing the repertory of algorithms available to automatic systems is feasible, but some form of 'operator monitoring' remains desirable. No system, whether automated or manual, will cope with the completely isoelectric T wave.

The poor reproducibility of QT dispersion measurement limits its usefulness. With an intrasubject coefficient of repeatability of 32 ms for QT dispersion and 28 ms for QTc dispersion, reported differences of this order for measurements in the same study for one observer should be treated with caution. Comparisons of QT dispersion measurements between either different observers or across studies call for even greater care.

References

[1] Day CP, McComb JM, Campbell RWF. QT dispersion: an indication of arrhythmia risk in patients with long QT intervals. Br Heart J 1990; 63: 342–4

[2] Barr CS, Naas A, Freeman M, Lang CC, Struthers AD. QT dispersion and sudden unexpected death in chronic heart failure. Lancet 1994, 343. 327–9.

[3] Buja G, Miorelli M, Turrini P, Melacini P, Nava A. Comparison of QT dispersion in hypertrophic cardiomyopathy between patients with and without ventricular arrhythmias and sudden death. Am J Cardiol 1993; 72. 973–6.

[4] Linker NJ, Colonna P, Kekwick CA, Till J, Camm AJ, Ward DE Assessment of QT dispersion in symptomatic patients with congenital long QT syndromes Am J Cardiol 1992; 69· 634–8.

[5] Pye M, Quinn AC, Cobbe SM. QT interval dispersion: a non-invasive marker of susceptibility to arrhythmia in patients with sustained ventricular arrhythmias? Br Heart J 1993; 71: 511–4.

[6] Kautzner J, Gang Yi, Camm AJ, Malik M. Short- and Long-term reproducibility of QT, QTc and QT dispersion measurement in healthy subjects. PACE 1994; 17: 928–37.

[7] Woods KL, Fletcher S, Roffe C, Haider Y. Intravenous magnesium sulphate in suspected acute myocardial infarction: results of the second Leicester Intravenous Magnesium Intervention Trial (LIMIT-2). Lancet 1992; 339: 1553–8.

[8] Bhullar HK, Fothergill JC, Goddard WP, de Bono DP. Automated measurement of QT interval dispersion from hard-copy electrocardiograms. J Electrocardiol 1993; 26 321–31.

[9] Lepeschkin E, Surawicz B. The measurement of the QT interval of the electrocardiogram. Circulation 1952; 6: 378–88.

[10] Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986; 1(8476): 307–10.

[11] Fei L, Statters DJ. QT-interval dispersion on 12-lead electrocardiogram in normal subjects: its reproducibility and relation to the T wave Am Heart J 1994; 127: 1654–5.

[12] Van de Loo A, Arendts W, Hohnloser SH. Variability of QT dispersion measurements in the surface electrocardiogram in patients with acute myocardial infarction and in normal subjects. Am J Cardiol 1994; 74: 1113–8.

[13] Murray A, McLaughlin NB, Bourke JP, Doig CJ, Furniss SS, Campbell RWF. Errors in manual measurement of QT intervals. Br Heart J 1994; 71: 386–90.

[14] Ahnve S Errors in the visual determination of corrected QT (QTc) interval during acute myocardial infarction. JACC 1985; 5: 699–702.

[15] Cui G, Sen L, Sager P, Uppal P, Singh BN. Effects of amiodarone, sematilide, and sotalol on QT dispersion. Am J Cardiol 1994; 74: 898–900

[16] Sylven JC, Horacek BM, Spencer CA, Klassen GA, Montague TJ. QT interval variability on the body surface. J Electrocardiol 1984; 17: 179–88.

[17] O'Donnell J, Knoebel SB, Lovelace DE, McHenry PJ. Computer quantitation of Q-T and terminal T wave (aT-eT) intervals during exercise: methodology and results in normal men. Am J Cardiol 1981; 47: 1168–72.

[18] Puddu PE, Bernard PM, Chaitman BR Bourassa MG. QT interval measurement by a computer assisted program: a potentially useful clinical parameter. J Electrocardiol 1992; 15. 15–22.