Available online at www.sciencedirect.com
ScienceDirect
Journal of Electrocardiology 41 (2008) 588–594
JOURNAL OF Electrocardiology
www.jecgonline.com

Automatic analysis of cardiac repolarization morphology using Gaussian mesa function modeling

Fabio Badilini, PhD,[a,*] Martino Vaglio, MS,[a] Rémi Dubois, PhD,[b] Pierre Roussel, PhD,[b] Nenad Sarapa, MD,[c] Isabelle Denjoy, MD,[d] Fabrice Extramiana, MD,[d] Pierre Maison-Blanche, MD[d]

[a]*AMPS-LLC, New York, NY, USA*
[b]*ESPCI, Paris, France*
[c]*Johnson & Johnson Pharmaceutical Research and Development, Spring House, PA, USA*
[d]*Hopital Lariboisiere, Paris, France*

Received 5 June 2008; revised 14 July 2008; accepted 15 July 2008

Abstract A novel fully automated method for wave identification and extraction from electrocardiogram (ECG) waveforms is presented. This approach implements the combined use of a new machine-learning algorithm and of specified parameterized functions called Gaussian mesa functions (GMFs). Individual cardiac cycle waveforms are broken up into GMFs using a generalized orthogonal forward regression algorithm; each individual GMF is subsequently identified (wave labeling) and analyzed for feature and morphologic extraction. The GMF associated with the repolarization waveform of the main vector lead, based on principal components analysis, was analyzed, and a set of morphologic parameters were derived under 2 experimental settings: first, in 100 digital 12-lead ECG Holter recordings acquired during three 24-hour periods (baseline and after 160 and 320 mg of sotalol) from 38 healthy subjects; second, in drug-free 12-lead resting ECGs from 100 genotyped long QT syndrome (LQTS) patients (50 each with LQT1 and LQT2). QT-interval duration was measured using an on-screen method applied to the global representative beats and reviewed by a senior cardiologist. QTci (individual correction) was used for analysis. All parameters in the sotalol test showed highly significant differences between the time of peak plasma concentration (Tmax) and baseline ECGs; however, the dynamic pattern of individual parameters followed different patterns. The LQTS test confirmed the results of the sotalol test, showing that GMF-based repolarization parameters were strongly modified as compared with healthy controls. In particular, T-wave width and descending phase of repolarization were more prolonged in LQT2 compared to LQT1.

Keywords: ECG modeling; Repolarization

Introduction

Historically, several decomposition models of the electrocardiogram (ECG) signal have been proposed. Depending on the scope of analysis, algorithms based on wavelet decomposition,[1,2] radial basis function modeling,[3] neural networks,[4,5] hidden Markov Model,[6,7] and many others have been implemented.[8] Although some of these models aim to globally describe the ECG signal, others are more targeted to extract relevant information related to its components, that is, the individual waveforms that contribute to the P-QRST complex. The latter approach can be particularly suited to analyses where a morphologic aspect of an ECG portion is of particular interest, such as the repolarization segment or the T wave.

In this article, a method recently described where the continuously recorded ECG signal is automatically analyzed using a model based on parameterized functions called Gaussian mesa functions (GMFs) is reviewed.[9] Individual cardiac cycles identified with standard signal processing

* Corresponding author. AMPS-LLC, via Paolo VI 34, 25018 (BS) Montichiari, Italy. Tel.: +39 030 9650745; fax: +39 030 9650572.
E-mail address: badilini@amps-llc.com

0022-0736/$ – see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.jelectrocard.2008.07.020

techniques are first decomposed into individual GMFs (one or more for each characteristic wave/component of the ECG signal). The modeled GMFs are subsequently labeled with an assignment process based on neural network probability estimators that link each modeled GMF to characteristics ECG waves (P, Q, R, S, and T).

We extracted the GMFs associated with the T wave and derived a set of repolarization morphologic indexes from the GMF parameters from ECG signals acquired in 2 separate experimental contexts. The first reflected the dynamic patterns of repolarization after intake of sotalol, a blocker of the rapidly activating repolarizing potassium current (Ikr) known to affect repolarization and cause Torsades de paintes (TdP) in the clinic, and the second illustrated the potential of the GMF-derived parameters to discriminate between 2 genetic forms of long QT syndrome (LQT1 and LQT2).

Methods

The core of the model is to break up each individual heartbeat as the summation of M individual GMFs. A single GMF is an asymmetric function composed of 2 half-Gaussian functions connected by a horizontal line and uniquely characterized by 5 parameters (Fig. 1): the time localization of the GMF (μ), the standard deviation (SD) of the first (ascending) half-Gaussian (σ_1), the SD of the second (descending) half-Gaussian (σ_2), the length of the horizontal part (σ_L) and the amplitude (A).

To efficiently fit the parameterized GMFs to the ECG signal, a new machine-learning algorithm, called generalized orthogonal forward regression (GOFR), was developed.[10] The GOFR is based on a 4-step iterative algorithm that is applied to the signal to be modeled (in our case a single P-QRST complex). After a library of generic GMF candidates is defined, the most relevant GMF of the library (ie, the one showing maximum correlation) is selected in the first step of GOFR. In the second step, the selected GMF is tuned to best fit the original signal, using a nonlinear optimization algorithm based on minimal least squares cost functions.[11] The last 2 steps consist of orthogonalization of the GMF library and of the original signal. The GOFR algorithm implementation is represented in Fig. 2. A set of neural network classifiers are finally implemented to associate the identified GMFs with the characteristic waves of the ECG signal.[10] The set of GMF iterations is stopped once the residual signal reaches a predetermined threshold. When the goal of the model application is to extract relevant features from the ECG, the process can be stopped earlier, that is, whenever the waveform components of interest (eg, the T wave) have been identified.

Gaussian mesa function modeling method was applied under 2 experimental contexts. The first experiment (test 1) was a previously described robust ECG study in 38 healthy individuals (27 males) in whom 12-lead digital Holter ECGs (Mortara Instrument, Milwaukee, WI) were recorded for 24 hours on 3 consecutive days.[12] No drug was given on the first day. On the second and the third day, 160 and 320 mg of sotalol, respectively (Betapace 80-mg tablets, Berlex Laboratories, Montville, NY) were administered orally at 08:00 AM. Plasma concentrations of sotalol were measured at 16 time points during 24 hours after dosing so that the time of occurrence of the peak plasma concentration of sotalol (Tmax) could be closely captured. The Holter ECG data were processed using Antares (AMPS llc, New York, NY), a software application designed to automatically identify and extract stable heart rate and minimal noise ECG segments.[13] In this experiment, 10-second 12-lead ECG extractions were obtained at 73 time points during the daytime, starting at 8:00 AM and ending at 8:00 PM, with a 10-minute gap between consecutive extractions. The QT-interval duration was measured on each extracted ECG by CalECG (AMPS llc, New York, NY), an on-screen caliper system using a semiautomated approach on global (superimposed) representative (median) beats from all 12 leads.[14] QTci (individual correction) was computed from each QT interval using the "global RR interval," that is, the fully automated measurement of the RR interval from the 10-second ECG. For the QTci computation, the power law (log-log) regression formula was applied to the 73 ECG extractions from the baseline recording.[15]

Principal component analysis based on the T wave was applied to the representative beats; and the T-wave symmetry index (TWLS), the ratio between the second and the first eigenvalues, was defined.[16] Gaussian mesa function modeling was finally applied the first eigenvector (PC1), and the parameters from the function labeled as the T wave were retained for morphologic analysis. Specifically, σ_1 and σ_2 were considered as indicators of the ascending and descending times of the T wave. Because of the shape of T wave in PC1, the value of σ_L was negligible, and for simplicity it was considered null. The interval $\mu + 2\sigma_2$ was considered as a surrogate of the QT interval (based on the assumption that in a Gaussian function, 2 SDs would cover more than 95% on the total energy under the curve). Similarly, $2\sigma_1 + 2\sigma_2$ can be seen as a surrogate of T-wave width. A graphic representation of these parameters is given in Fig. 3. QTci and GMF-derived parameters from peak

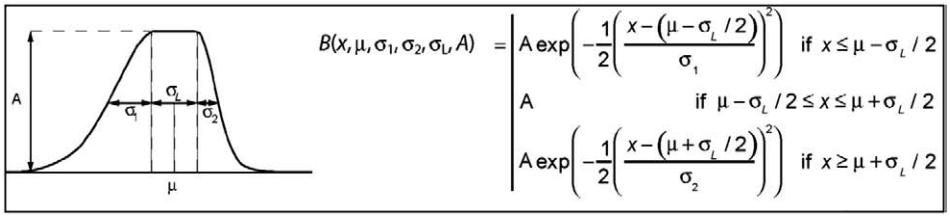

Fig. 1. Definition of the GMF (from reference [9], with permission from Elsevier).

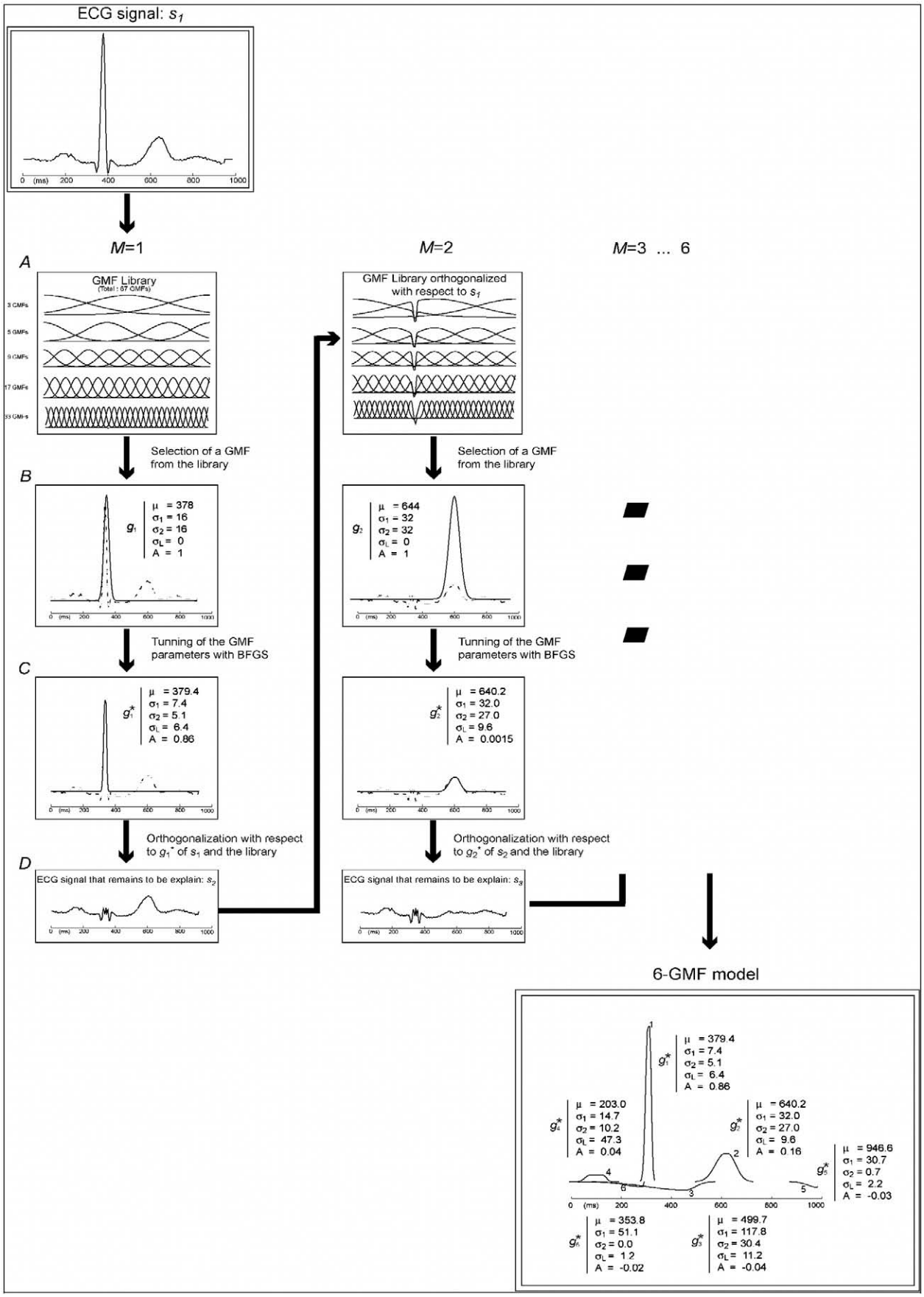

Fig. 2. Generalized orthogonal forward regression algorithm for selection, tuning, and orthogonalization of GMF modeling an ECG signal (from reference [9], with permission from Elsevier).

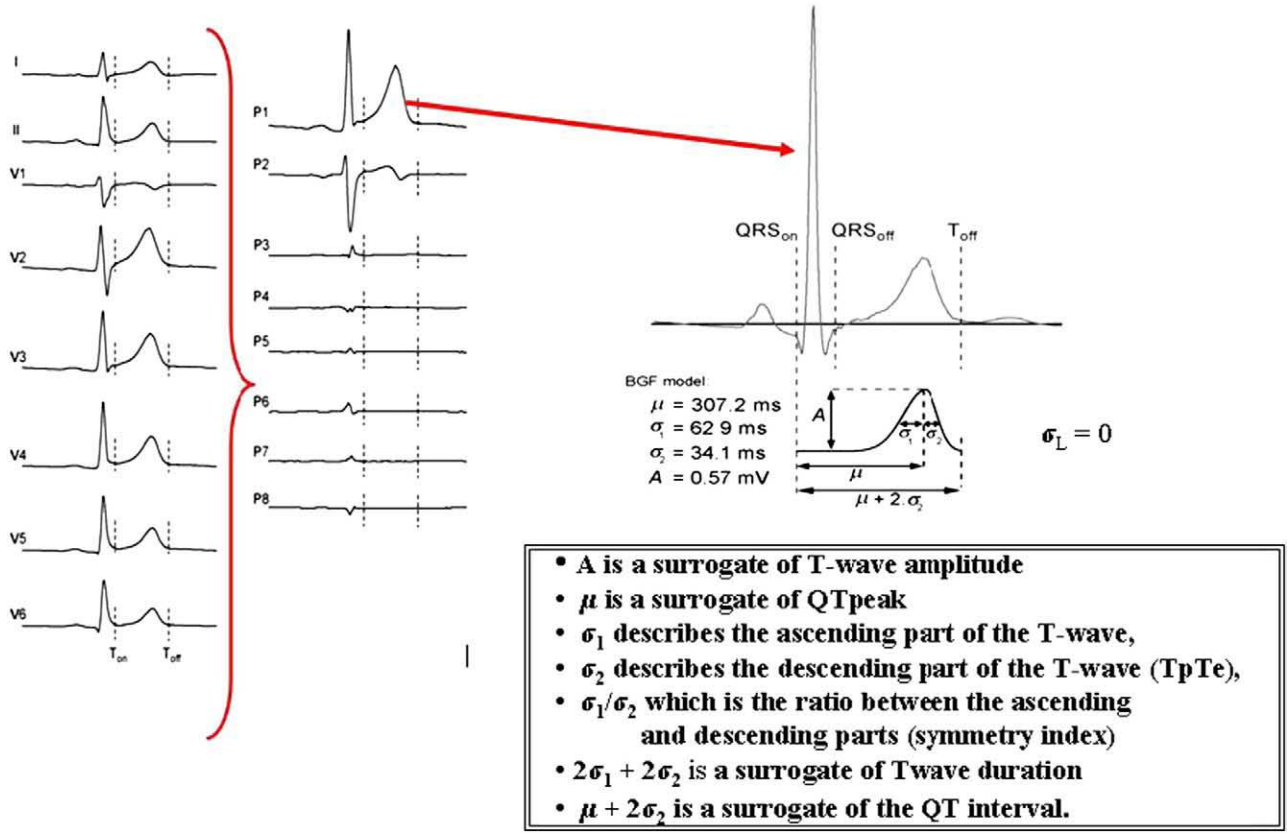

Fig. 3. Gaussian mesa function–based repolarization parameters.

concentration vs time-matched ECGs (separately for the 2 dosing levels inspected) were compared using Kruskal-Wallis nonparametric test for medians.

In the second experiment (test 2), 100 resting 12-lead ECGs (MacVu and Mac5000 electrocardiograph, GE HealthCare, Milwaukee, WI) were selected from patients with the confirmed genotype of LQTS from the French LQTS registry.[17] This set consisted of 50 ECGs from LQT1 patients (mean age, 29 ± 18; 26 males) and 50 ECGs from LQT2 patients (mean age, 32 ± 19; 22 males). Baseline (drug-free) ECGs from the healthy subjects in the sotalol study were used as the control set. As with first experiment, GMF modeling was applied to the first eigenvector of the principal component, and the same indexes were derived form the GMF labeled as T wave. QT, Bazett-corrected QTcB, and all morphologic parameters were compared between groups using Kruskal-Wallis nonparametric test for medians. Patients in each LQTS group were finally categorized as symptomatic (patients with known occurrence of syncope) vs asymptomatic, and the behavior of each parameter to characterize a symptomatic phenotype was investigated.

Results

Test 1 (the sotalol experiment)

Results of baseline vs sotalol comparisons at Tmax (time when peak plasma concentration is reached) are reported in Table 1. As plasma concentration dynamic is different in each patient, the Tmax time point varies across subjects: Tmax was observed 2 hours 49 minutes (±45 minutes) and 2 hours 39 minutes (±47 minutes) after dosing, respectively, for 160 and 320 mg sotalol administration. Upon exposure to peak plasma levels of sotalol, all GMF parameters showed significant changes indicative of a prolonged repolarization and reduced T-wave energy. Both effects were significantly more pronounced after 320 vs 160 mg dose of sotalol. T-wave symmetry index exhibited an increased pattern, although not reaching statistical significance.

The QT interval and its GMF-based automated surrogate ($\mu + 2\sigma_2$) were highly correlated ($r = 0.96$, $P < .001$); to assess their degree of agreement, Bland-Altman analysis from the pooled sample of 6236 ECGs from all time points was performed (Fig. 4). The difference

Table 1
Comparisons of baseline vs peak sotalol plasma concentrations (160 and 320 mg of sotalol in healthy subjects, test 1)

	Baseline	160mg (2 h 49 min [±45 min])	320mg (2 h 39 min [±47 min])
RR (ms)	892 ± 143	1103 ± 140†	1149 ± 117*,†
QTci (ms)	387 ± 17	428 ± 26†	438 ± 28*,†
TWLS (%)	0.19 ± 0.09	0.22 ± 0.13	0.22 ± 0.07
$\mu + 2\sigma_2$ (ms)	345 ± 26	423 ± 33†	441 ± 37*,†
σ_1 (ms)	66 ± 17	99 ± 19†	116 ± 24*,†
σ_2 (ms)	30 ± 5	41 ± 13†	45 ± 11*,†
A (μV)	701 ± 71	640 ± 80†	611 ± 79*,†

* $P < .05$ vs single dose.
† $P < .05$ vs baseline.

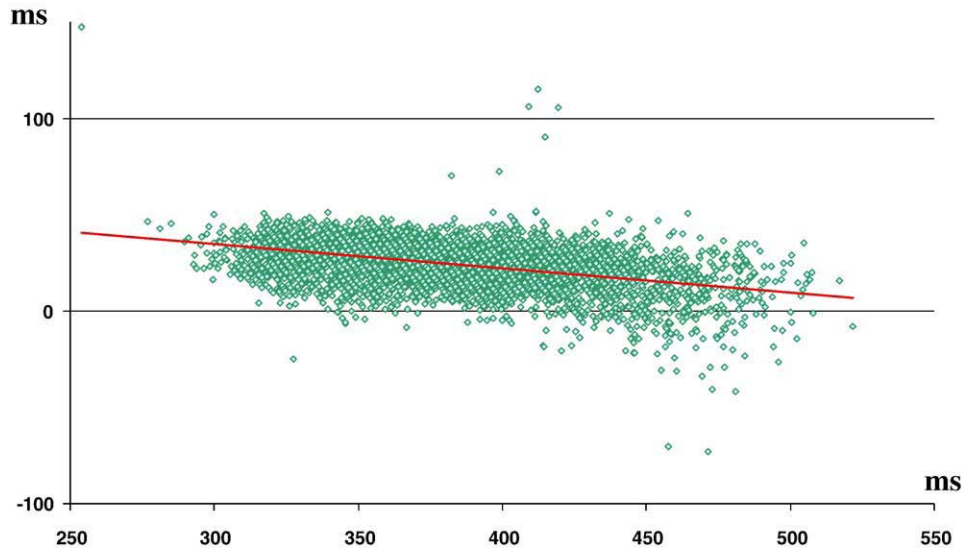

Fig. 4. Bland-Altman plot of QT interval vs GMF-related repolarization duration from the overall data population (6236 ECGs from test 1).

between QT and $\mu + 2\sigma_2$ exhibited a significant trend (intercept, 73 milliseconds; slope, −0.1266; $P < .001$), thus indicating a large positive bias ($QT > \mu + 2\sigma_2$) at fast heart rates and a closer agreement at slower heart rates.

Fig. 5 presents an example of time-course patterns in a typical subject receiving a 320-mg dose of sotalol: at 1 hour after dosing, the T wave is still very close to the baseline shape, whereas at 1.5 hours after dosing, the repolarization prolongation becomes apparent, and it is mainly related to a change of the ascending phase (σ_1). At maximum concentration (plasma Tmax of sotalol, between 2 and 2.5 hours after dosing), the overall duration of repolarization has approximately reached its maximum value, but the morphology changes are still progressing. The maximum morphologic changes in the example depicted in Fig. 5 were seen 7 hours after dosing, and even at 12 hours, the morphology of the T wave has not yet returned to what it was at peak concentration. In Fig. 6, the trends over time of the differences of both QT and $\mu + 2\sigma_2$, together with the averaged distribution of sotalol plasma concentration, are shown.

Test 2 (the LQTS experiment)

Comparisons between LQT1 and LQT2 with respect to controls are reported in Table 2. All parameters reflecting

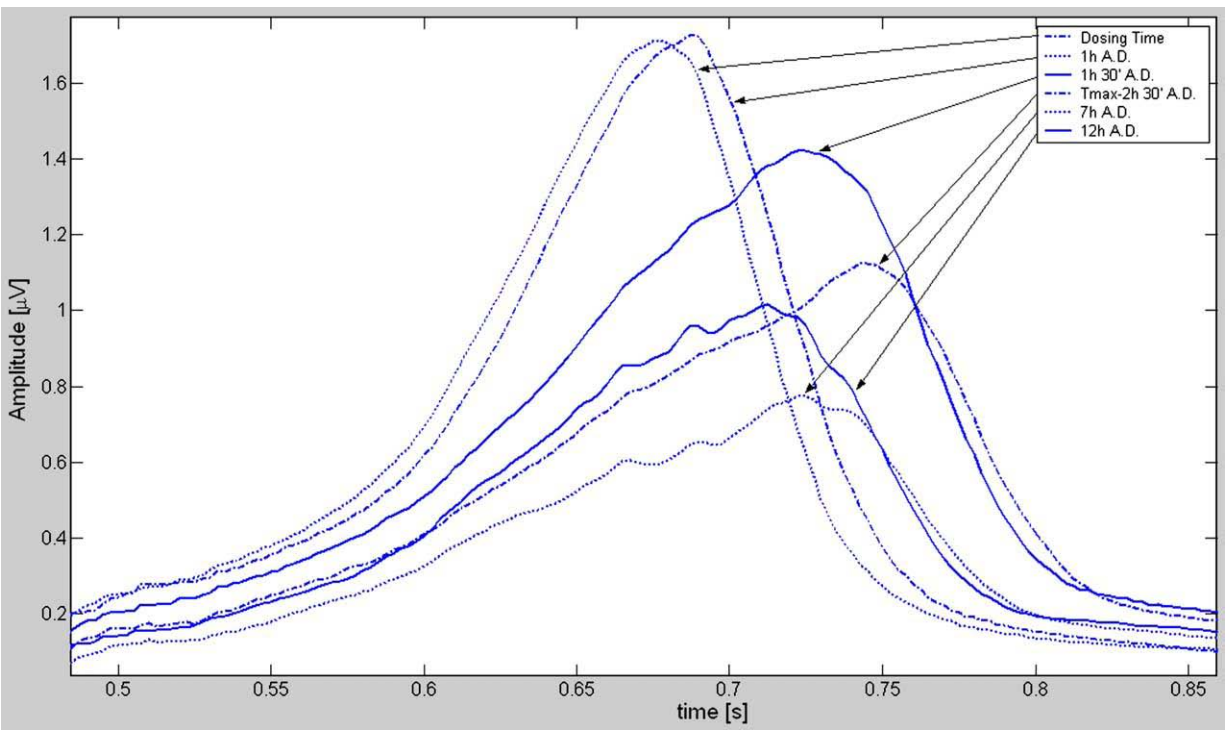

Fig. 5. Overlap of T wave after 320 mg sotalol administration in a representative healthy subject.

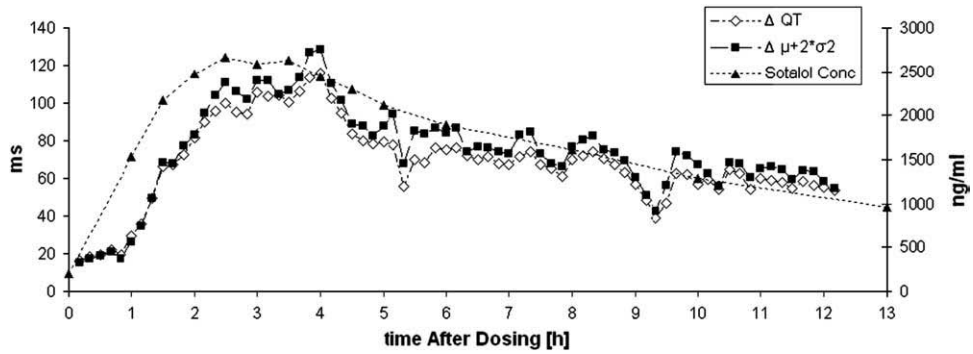

Fig. 6. Trend display of QT interval vs GMF-related repolarization duration differences with respect to time-matched baseline ECGs.

repolarization duration (QT, μ, $\mu + 2\sigma_2$) were significantly prolonged in both LQT1 and LQT2 patients. However, the morphologic parameters showed less consistency between LQT types; for example, the descending phase of the T wave (σ_2), the T wave amplitude, and the GMF-related T-wave duration surrogate $2\sigma_1 + 2\sigma_2$ were clearly altered only in the LQT2 group. The TWLS in the LQT2 group was significantly increased with respect to both the control and the LQT1 groups. Although a tendency toward discrimination between the symptomatic and asymptomatic patients was observed for most parameters in both LQT types, statistical significance was only reached for the difference in the GMF-based T-wave duration $2\sigma_1 + 2\sigma_2$ in the LQT2 group (Fig. 7).

Discussion

Gaussian mesa function modeling is a novel method with a promising potential for the analysis of abnormal cardiac repolarization. Being fully automated, GMF modeling is a powerful tool for efficient beat detection and classification of ECG waveforms from both resting ECGs and continuous ECG recordings.[9] This study shows the applications of this approach in the assessment of drug-induced morphologic changes of cardiac repolarization (test 1) and for the identification of new discriminating indexes between different forms of the inherited LQTS (test 2).

We have observed a strong correlation ($r = 0.96$) between the semiautomated measurement of the QT interval duration and the fully automated determination of the GMF-based repolarization duration index $\mu + 2\sigma_2$. Despite this correlation, Bland-Altman analysis evidenced a nonstable bias, whereby $\mu + 2\sigma_2$ is shorter than QT, particularly at faster heart rates (Fig. 4). Gaussian mesa function–based repolarization duration may address different features of cardiac repolarization, and further investigation will tell whether this biomarker also carry better diagnostic and prognostic values. Nevertheless, it is important that our novel fully automated approach produced an index highly correlated with the QT interval duration measured by the standard manually supervised method routinely used by central ECG laboratories.

Results from the sotalol experiment evidenced a strong dose-related prolongation of repolarization (Table 1); this finding confirms what was previously reported using standard QT and novel T-wave area-based parameters.[18,19] More interestingly, different dynamics were observed with the sotalol-induced GMF-based morphologic indexes. Specifically, the ascending phase duration σ_1 was affected very early after dosing, whereas the changes in the descending phase duration σ_2 were more closely associated with the morphologic changes and correlated with the amplitude decrease observed later in time (Fig. 5). These

Table 2
Comparisons between standard and GMF-related repolarization parameters in LQT1 and LQT2 patients vs healthy controls (test 2)

	Control (n = 38)	LQT1 (n = 50)	LQT2 (n = 50)
RR (ms)	892 ± 143	941 ± 222	985 ± 192*
QT (ms)	372 ± 25	433 ± 73*	445 ± 72*
QTcB (ms)	395 ± 18	449 ± 42*	450 ± 50*
TWLS (%)	0.19 ± 0.09	0.22 ± 0.11	0.32 ± 0.20*,†
$\mu + 2\sigma_2$ (ms)	345 ± 26	401 ± 67*	412 ± 64*
μ (ms)	286 ± 23	343 ± 60*	343 ± 65*
σ_1 (ms)	66 ± 17	61 ± 12	69 ± 21
σ_2 (ms)	30 ± 5	29 ± 7	36 ± 14*,†
$2\sigma_1 + 2\sigma_2$ (ms)	182 ± 29	180 ± 32	210 ± 55*,†
A (μV)	1043 ± 451	974 ± 501	749 ± 384*,†

The ECGs from drug-free baseline in healthy subjects in test 1 were used as controls.

* $P < .05$ vs control.

† $P < .05$ vs LQT1.

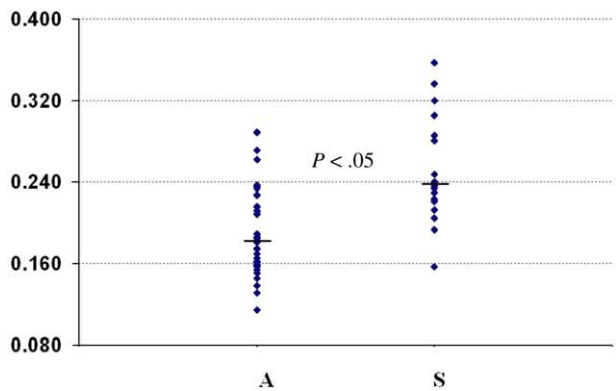

Fig. 7. GMF-based T-wave duration in symptomatic and asymptomatic LQT2 patients.

findings indicate that the prolongation of the QT interval occurs early in time (even before peak concentration of the drug is reached), whereas more critical sotalol-induced changes in ECG waveform morphology only occur at later stages.

Our findings on the LQTS experiment showed a marked prolongation of all repolarization duration indexes in both the genotyped LQT1 and LQT2 patient groups (Table 2). This result confirms previous findings[20]; however, morphologic parameters seem to indicate a stronger differentiation between the 2 genetic variants of LQTS than the standard QT duration, thus offering enhanced diagnostic potential of GMF modeling. For example, none of the GMF parameters were affected in LQT1 patients, implying that in this population, the T-wave duration is shifted temporally without any critical change in its width, amplitude, slope, or other morphologic aspects. Conversely, in LQT2 patients, all parameters except σ_1 were altered compared to controls and (for TWLS, σ_2, $2\sigma_1 + 2\sigma_2$, and A) vs the LQT1 group. This implies that in LQT2 the T wave is shifted temporally, it is wider, and has a reduced amplitude mainly because of changes in the descending phase of the T wave (σ_2). These findings are in agreement with previously published data[16,20] and particularly with the pioneer work of Moss et al who first described patterns of different forms of LQTS in 1995.[21]

Comparisons between symptomatic and asymptomatic patients with LQTS enhanced a tendency for more pronounced alterations in the symptomatic group, although statistical significance was only reached for the GMF surrogate of T-wave duration $2\sigma_1 + 2\sigma_2$ in LQT2 patients (Fig. 7). The interpretation of this result is limited by the relatively small numbers of symptomatic patients with LQT2 (19 vs 31 asymptomatic). However, the trend toward a separation in this morphologic T-wave parameter between the 2 clinical phenotypes of LQT2 points toward the possible superior discrimination power of GMF-based parameters over QT interval duration and warrants further study in larger cohorts of patients with LQTS.

Conclusions

We described a new method for the automatic analysis of ECG waveforms based on GMF modeling suited for both classification and parametric feature extraction. The results presented suggest that this method could be useful to characterize repolarization-related abnormalities in both the clinical health care and drug development environments.

Limitations

In this preliminary study, analysis is based on the first principal component of the repolarization phase. Although most of the energy is confined in the first component, it has been shown that morphologic changes can also be visible in the second and third components and further analysis will be needed to determine the effectiveness of GMF modelization extended to other components.

The discriminative power of the new biomarkers presented in this article, particularly with respect to different variants of LQTS, will need to be further explored with more powerful statistical tools once the sample size of the data population will have reached suitable size.

References

1. Li C, Zheng C, Tai C. Detection of ECG characteristic points using wavelet transforms. IEEE Trans Biomed Eng 1995;42:21.
2. Martinez J, Almeida R, Olmos S, Rocha A, Laguna P. A wavelet-based ECG delineator: evaluation on standard databases. IEEE Trans Biomed Eng 2004;51:570.
3. Suppappola S, Sun Y, Chiaramida SA. Gaussian pulse decomposition: an intuitive model of electrocardiogram waveforms. Ann Biomed Eng 1997;25:252.
4. Simon BP, Eswaran C. An ECG classifier designed using modified decision based neural networks. Comput Biomed Res 1997;30:257.
5. Osowski S, Linh T. ECG beat recognition using fuzzy hybrid neural network. IEEE Trans Biomed Eng 2001;48:1265.
6. Koski A. Modelling ECG signals with hidden Markov models. Artif Intell Med 1996;8:453.
7. Thoraval L, Carrault G, Bellanger JJ. Heart signal recognition by hidden Markov models: the ECG case. Methods Inform Med 1994;33:10.
8. Kanters JK, Fanoe S, Larsen LA, Bloch Thomsen PE, Toft E, Christiansen M. T wave morphology analysis distinguishes between KvLQT1 and HERG mutations in long QT syndrome. Heart Rhythm 2004 September;1:285.
9. Dubois R, Maison-Blanche P, Quenet B, Dreyfus G. Automatic ECG wave extraction in long-term recordings using Gaussian Mesa functions models and nonlinear probability estimators. Comput Methods Programs Biomed 2007;88:217.
10. Dubois R, Quenet B, Faisandier Y, Dreyfus G. Building meaningful representations for nonlinear modeling of 1d and 2d-signals: applications to biomedical s ignals. Neurocomputing 2006;69:2180.
11. Broyden CG. The convergence double rank minimization algorithms 2. The New Algorithm. J Inst Maths Appl 1970;6:222.
12. 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.
13. Badilini F, Vaglio M, Sarapa N. Automatic extraction of ECG strips from continuous 12-lead Holter recordings for QT analysis at prescheduled versus optimized time points. Ann Noninvasive Electrocardiol 2008 [in press].
14. Badilini F, Sarapa N. Implication of methodological differences in digital electrocardiogram interval measurement. J Electrocardiol 2006; 39:S152.
15. Malik M, Färbom P, Batchvarov V, Hntakova K, Camm AJ. Relation between QT and RR intervals is highly individual among healthy subjects: implications for heart rate correction of the QT Interval. Heart 2000;87:220.
16. Badilini F, Fayn J, Maison-Blanche P, et al. Quantitative aspects of ventricular repolarization: relationship between three dimensional T-wave loop morphology and scalar QT dispersion. Ann Noninvasive Electrocardiol 1997;2:146.
17. Imboden M, Swan H, Denjoy I, et al. Female predominance and transmission distortion in the long-QT syndrome. N Engl J Med 2006; 355:2744.
18. Extramiana F, Badilini F, Sarapa N, Leenhardt A, Maison-Blanche P. Contrasting time and rate based approaches for the assessment of drug-induced QT changes. JCP 2007;47:1129.
19. Couderc JP, Vaglio M, Xia X, et al. Impaired T-amplitude Adaptation to Heart Rate Characterizes I(Kr) Inhibition in the Congenital and Acquired Forms of the Long QT Syndrome. J Cardiovasc Electrophysiol 2007;18:1299.
20. Vaglio M, Couderc JP, McNitt S, Xia X, Moss AJ, Zareba W. A quantitative assessment of T-wave morphology in LQT1, LQT2, and healthy individuals based on Holter recording technology. Heart Rhythm 2008;5:11.
21. Moss AJ, Zareba W, Benhorin J, et al. ECG T-wave patterns in genetically distinct forms of the hereditary long QT syndrome. Circulation 1995;92:2929.