Biomarkers of pre-existing risk of Torsade de Pointes under Sotalol treatment

Pablo Daniel Cruces [a,b] Drago Torkar [c] Pedro David Arini [a,b]

[a] *Instituto de Ingeniería Biomédica, UBA, Paseo Colón 850 (C1063ACV), Buenos Aires, Argentina*
[b] *Instituto Argentino de Matemática 'Alberto P. Calderón', CONICET, Saavedra 15 (C1083ACA), Buenos Aires, Argentina*
[c] *Institut 'Jožef Stefan', Department of Computer Systems, Jamova cesta 39 (SI-1000), Ljubljana, Slovenia*

Corresponding Author at: Pablo Daniel Cruces, Saavedra 15, CABA (C1083ACA), Argentina; pcruces@fi.uba.ar; 0054-11-4954-6781 int 121.

Abstract

Introduction: Antiarrhythmic drugs therapies are currently going through a turning point. The high risk that exists during the treatments has led to an ongoing search for new non-invasive toxicity risk biomarkers. **Methods:** We propose the use of spatial biomarkers obtained through the quaternion algebra, evaluating the dynamics of the cardiac electrical vector in a non-invasive way in order to detect abnormal changes in ventricular heterogeneity. In groups of patients with and without history of Torsade de Pointes undergoing a Sotalol challenge, we compute the radius and the linear and angular velocities of QRS complex and T-wave loops. From these signals we extract significant features in order to compute a risk patient classifier. **Results:** Using machine learning techniques and statistical analysis, the combinations of few indices reach a pair of sensitivity/specificity of 100%/100% when separating patients with arrhythmogenic substrate. Several biomarkers not only measure drug-induced changes significantly but also observe differences in at-risk patients outperforming current standards. **Discussion:** Alternative biomarkers were able to describe pre-existing risk of patients. Given the high levels of significance and performance, these results could contribute to a better understanding of the torsadogenic substrate and to the safe development of drug therapies.

Keywords

Cardiac electrical vector, Antiarrhytmic drugs, Proarrhythmic side effects, Quaternion theory.

1. Introduction

Cardiovascular diseases account for more than 30% of the causes of deaths in the world [1], among them the most important are the malignant ventricular arrhythmias (MVA). In recent decades, a large amount of antiarrhythmic drugs (ADs) have been developed. ADs act on permeability of the myocyte ion channels, preventing MVA and/or sudden cardiac death. However, these therapies are currently at a turning point. Recently, it has been highlighted the high risk that exists in ADs with reported adverse effects of Torsade de Pointes (TdP) [2]. Also, a compilation of hundreds of drugs that produce these effects have been constructed considering different risk categories [3].

TdP is a polymorphic ventricular tachyarrhythmia with a high risk of ventricular fibrillation. Although the mechanisms that trigger it, in non-congenital conditions, have not yet been elucidated, it has been mostly associated with the supply of different ADs such as: sotalol, amiodarone, dofetilide, quinidine sulfate, ranolazine and verapamil hydrochloride. Moreover, many commonly used non-cardiac drugs such as diuretics, antibiotics, antidepressants, among others, also have pro-arrhythmic side effects of TdP [4].

Different international agencies in favor of the development of therapies with ADs [5] propose the search for new non-invasive toxicity risk biomarkers with an efficacy that improves current biomarkers. Modern regulatory methods of cardiotoxicity assessment are based on the measurement of electrocardiogram (ECG) indices. One of the most observed characteristics of drug-induced TdP is the prolongation of the QT interval, a measure of the total time of ventricular depolarization and repolarization on the ECG. In 2005, the International Committee for Harmonization issued the S7B / E14 guidelines, requiring to date, studies of duration of the QT interval for all drugs before regulatory approval [6,7]. While this measure prevented the sale of many drugs with proarrhythmic effects, it also discouraged the development of potentially effective drugs. It is important to note that the QT interval measures have insufficient specificity. There are two key factors to explain this fact: First, this measure depends strongly on an accurate determination of the end of the T-wave, which is often not possible [8]; and second, information on morphological changes and velocities of the cardiac electrical vector is not taken into account by concentrating all valid information for the diagnosis of proarrhythmic risk only in the duration of the QT interval.

In this work, we propose as a paradigm shift the use of vectorcardiogram (VCG) biomarkers obtained from the quaternion algebra, evaluating the dynamics of the cardiac electrical vector in order to detect abnormal changes in ventricular heterogeneity. Recently, we have shown the advantages of these methods based on quaternion theory in the diagnosis of early acute coronary syndrome [9]. Biomarkers of this type are completely independent of the determination of the fiducial points of the ECG and have a significant degree of reproducibility in different contexts of the ECG recordings, such as resting, stress or Holter. These biomarkers could be very useful in the early detection of TdP risk during treatment with ADs or other drugs. We are confident that this can be a useful contribution to the optimization of therapeutic strategies and to the improvement in the design of studies for pharmacological treatments.

2. Materials and Methods

2.1. *Dataset*

The ECG recordings were obtained from "Sotalol IV and History of TdPs" database through a project with Telemetric and ECG Holter Warehouse (THEW) [10]. The population consisted of two groups: 16 subjects without history of TdP ($-TdP$) and 16 patients with documented TdP in the context of a drug with QT-prolonging potential: sotalol, sumatriptan, amiodarone, bisacodyl, cipramil, furosemide, clarithromycin, erythromycin, roxithromycin ($+TdP$). The protocol utilized in this database was detail described in [11]. In resume, dl-sotalol was supplied intravenously at a constant rate during 20 minute in two group of individuals, with ($+TdP$) and without history of drug-induced TdPs ($-TdPs$), respectively. The dose was of 2mg/kg body weight in 50ml of 0.9% saline solution at a dose of 2mg/kg body weight in 50ml of 0.9% saline solution. The protocol was carried out in the morning. Following the sotalol was injected patients were closely and continuously monitored in the intensive care unit.

Group	Age	Gender	AF_h	CAD	MI_h	HT_h	EF	H_k
$+TdP$	58 ± 12 y.o.	F 70% M 30%	62.5%	18.7%	18.7%	43.7%	25%	31.2%
$-TdP$	61 ± 12 y.o.	F 56% M 44%	93.7%	18.7%	6.2%	56.2%	0%	0%

Table 1
Clinical information for each group: History of Atrial Fibrillation (AF_h), Coronary Artery Disease (CAD), History of Myocardial Infarction (MI_h), History of Hypertension (HT_h), Ejection Fraction below normal $< 55\%$ (EF), Hypokalemia (HK).

The age and gender of the subjects are reported in Table 1 together with the relevant information concerning the clinical characteristics. All the patients have normal heart rate. Presence of history of hypertension, coronary artery disease and myocardial infarction were similar between the two groups. There were several patients with history of atrial fibrillation in both of them. None of the subjects were carrying mutation linked to the major congenital forms of the LQTS. Baseline ECG measurements shows one control and five patients slightly above the gender-specific thresholds for LQTS. T-wave and QRS complex widths present normal values. Four patients in $+TdP$ group have reduced values of left ventricular ejection fraction. Each individual record includes two dataset of continuous 3 to 4-minute 12-lead ECG at 1kHz of sampling frequency (*Fs*) and $5\mu V$ resolution. The former is acquired at rest in supine position before (baseline) the drug supply and the later at 20-min steady state phase (peak concentration) after Sotalol injection. None of the subjects experienced episodes of TdP during the challenge.

2.1.1. *Drug selection criterion*

As mentioned above, the mechanisms behind drug-induced TdP have not been elucidated. There is some consensus that it is mainly originated in repolarization impairments, such as: changes in ventricular heterogeneity, reduced repolarization reserve, and cardiac electrical instability. Sotalol is a well-known medication with torsadogenic side effects. In consequence, a protocol which exposed patients to this drug, is very useful to the assessment of spatial biomarkers during repolarization changes.

Almost every drug that have been linked with TdPs modify the same ion current, i.e. the rapid components of the delayed rectifier potassium current (I_{Kr}) of the cardiomyocites [12]. In this sense, some studies have experimentally shown repolarization impairments provoked by Class III antiarrhythmic agents just like by other non-antiarrhythmic medications [13]. We have also observed similar impairments in spatial electrocardiographic signals [14], which were obtained from an artificial thorax from In-Vitro rabbit hearts during d-sotalol supply.

Lastly, the database used in this work [10] was proper to obtain provocative drug testing to unmask latent abnormalities in cardiac repolarization, as was the case of the Sotalol described in [11].

2.2. *Preprocessing of spatial signals*

The study of the dynamics of cardiac vector needs to be approached from a spatial perspective. For this purpose, ECG signals are transformed to XYZ using the inverse Kors matrix. All the fiducial points of each ECG and XYZ signals have been obtained through a Wavelet-transform based method. A 5^{TH} order Butterworth high-pass filter (0.5Hz, bidirectional) has been applied for baseline wander correction. Likewise, high frequency noise has been removed using a 5^{TH} order Butterworth low-pass filter (40Hz, bidirectional). Cut-off frequencies were selected according current standards in this kind of algorithm [15,16].

2.3. *Dynamic biomarkers*

The ECG is a measure of the projection on the chest of all the action potentials of the heart. The potential of cardiac cells is determined by the concentrations of ions (Na^+, K^+, Ca^{2+}, Cl^-) on either side of the membrane. Ions require aqueous channels to move across cell membranes at specific times during cardiac cycle. These movements produce currents that define the basis of the cardiac action potential.

Antiarrhythmic drugs act on these channels in order to prolong refractory period or reduce reentry circuits. It has been shown that the measure of JTp interval (time between J point and the peak of the T-wave)

can distinguish QT prolonging drugs that mainly block I_{Kr} currents from those affecting multiple channels [17]. In addition, the relationship between Tpe (time between peak and end of the T-wave) and ventricular heterogeneity has been extensively studied [18,8,19].

In this sense, we observe that all these biomarkers evaluate only durations, overlooking all the spatial information contained in the dynamics of the cardiac electrical vector. Drug supply involves ion channels blockage causing conductivity alterations that consequently alter the velocity and magnitude of the vector during both the QRS complex and the T-wave. Also, considering current biomarkers JTp, Tpe and QT, it is clear that the halves of the T-wave reflect differences between class III drugs (blockage of potassium currents) and miscellaneous drugs. For this reason, we propose new biomarkers based on the dynamics of the cardiac electrical vector during the QRS complex, the T-wave and during the first half (early repolarization, ER) and the second half (late repolarization, LR) of the latter. These biomarkers are studied in the stage prior to drug supply and in the peak concentration of sotalol.

2.3.1. *Velocities and magnitude vector*

Using the Kors transformation, XYZ signals of vectorcardiogram have been obtained. The three dimensional movement of the tip of cardiac vector, computed from these signals, is described through the radius and the linear and angular velocities. The radius is merely the 2-norm of XYZ vectors, ie a n^{TH} point $\mathbf{P_n}(x, y, z)$ of a beat, considering a given sampling frequency *Fs*, has a radius defined by $\|\mathbf{P_n}\|_2$. On the other hand, while linear velocity can be computed by direct differenciation of the prior vectors (Eq. 1), the angular velocity is quite hard to be computed.

$$\boldsymbol{v_n} = (\mathbf{P_{n+1}}(x, y, z) - \mathbf{P_n}(x, y, z)) \cdot Fs \tag{1}$$

We have previously shown that by means of a quaternion transformation of 3D-space it is possible to obtain the expression of the instantaneous angular velocity of the cardiac electrical vector [9]. Quaternions are hypercomplex numbers that constitute a non-commutative field. They are very useful in the study of rotations and it has been shown that they are very efficient in terms of uncertainty propagation and computing time by comparison with traditional methods such as Euler matrices [20]. Every quaternion (Eq. 2) has a real part (associated with an amount of rotation) and three imaginary parts (associated with the rotation axis) that satisfy the Hamilton rule ($\boldsymbol{i}^2 = \boldsymbol{j}^2 = \boldsymbol{k}^2 = \boldsymbol{ijk} = -1$).

$$\mathbf{q_n} = a_1 + a_2\boldsymbol{i} + a_3\boldsymbol{j} + a_4\boldsymbol{k} \quad , \quad a_{1\ldots4} \in \mathbb{R} \tag{2}$$

Similarly, $\mathbf{q_n}$ can be written in terms of a rotation angle α from a XYZ point $\mathbf{P_n}$ to its consecutive point $\mathbf{P_{n+1}}$:

$$\mathbf{q_n} = cos\left(\frac{\alpha}{2}\right) + \mathbf{u}.sin\left(\frac{\alpha}{2}\right) \tag{3}$$

where $\mathbf{u}$ represents the rotation axis. Both trigonometric functions are easily obtained from dot and cross products between $\mathbf{P_n}$ and $\mathbf{P_{n+1}}$.

Then, if we take every point of XYZ signals, we have a sequence of quaternions and thereby we can compute the instantaneous angular velocity by solving the Poisson (Eq. 4) equation [21]. Quaternion inverse can be expressed in terms of its conjugate and its norm: $\mathbf{q}^{-1} = \overline{\mathbf{q}}/\|\mathbf{q}\|^2$.

$$\dot{\mathbf{q}_n} = \frac{1}{2}.\boldsymbol{\omega_n}.\mathbf{q_n} \;\rightarrow\; \boldsymbol{w_n} = (\mathbf{q_{n+1}} - \mathbf{q_n}) \cdot Fs \cdot \frac{\bar{\mathbf{q}_n}}{\|\mathbf{q_n}\|^2} \tag{4}$$

2.3.2. *Spatial features*

For the two velocities and the magnitude vector (ω, v and E) the maxima and areas are computed. As explained previously, this procedure is performed for the QRS complex, for the T-wave, and for the two halves of the latter: early and late repolarization. An explanatory graph is shown in Figure 1. Herein, 3-D velocity signals are obtained from a T-wave loop and from these, the maxima and area indices. Superscript indicates the corresponding wave (QRS, T, ER, LR) and the subscript indicates the signal (E, ω and v). The maxima are computed as,

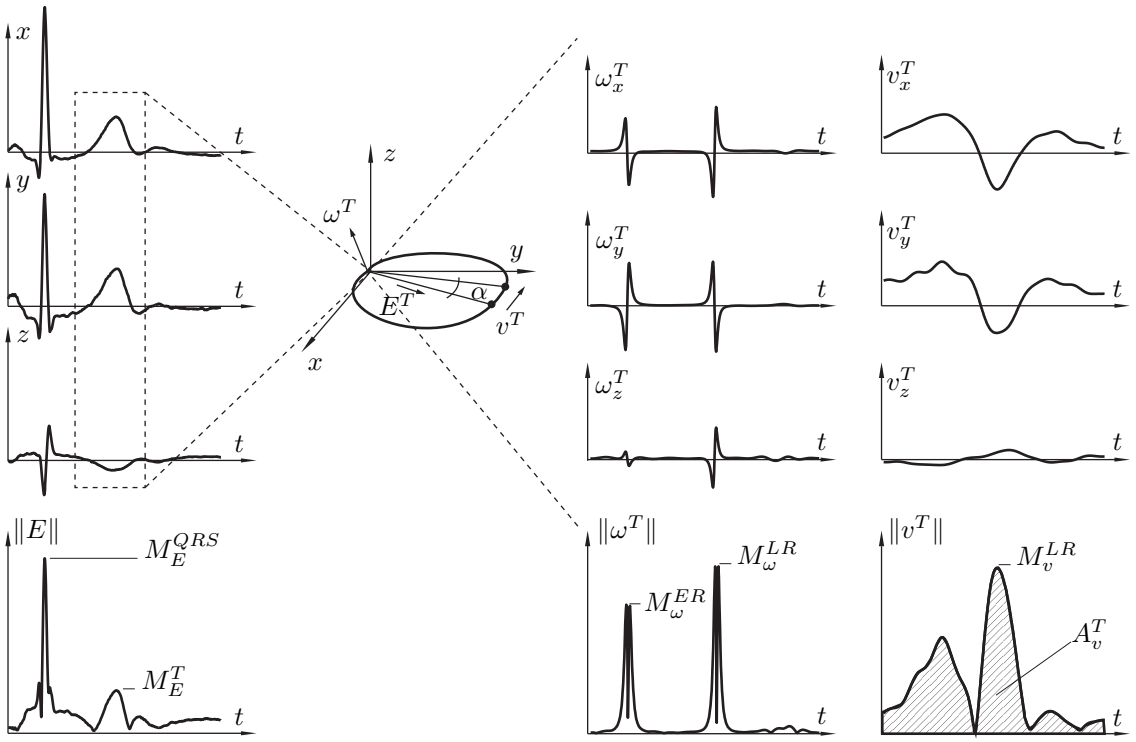

Figure 1. Description of spatial biomarkers. Magnitude vector (E), angular (ω) and linear (v) velocities are obtained from XYZ signals using a quaternion method. From these, some of the area and maxima indices are shown. Superscript indicates the corresponding wave (QRS, T, ER, LR) and the subscript indicates the signal.

$$M_{\mathsf{S}}^{\mathsf{b}} = max(\|\mathsf{S}^{\mathsf{b}}\|_2) \tag{5}$$

where S is the signal E, ω or v, and b is the wave QRS, T, ER or LR. On the other hand, the areas are obtained as,

$$A_{\mathsf{S}_i}^{\mathsf{b}} = \frac{\|\mathsf{S}_i{}^{\mathsf{b}}\|_1}{\|(\|\mathsf{S}^{\mathsf{b}}\|_2)\|_1} \quad , \quad i = x, y, z \tag{6}$$

It is important to mention that in order to make valid comparisons, the areas have been normalized. The energy of the projection of a vector wave on each axis is obtained as a relationship with the total energy of the vector. Also, for the early and late repolarization waves, their energy is normalized by the total energy of the T-wave. Finally, the area of the QRS complex is normalized by the maximum amplitude.

Additionally, the QT interval is obtained with the aim of comparing with the current standard. The QRS complex onset to T-wave end (QT) interval, quantifies the full time of depolarization and repolarization of ventricles and it is currently the major ECG marker used for decision on drug cardiotoxicity [5]. It is computed as Eq. 7, where the interval is corrected with RR interval (the cardiac cycle).

$$QTc = (T_{\text{END}} - QRS_{\text{ON}})/\sqrt{RR} \tag{7}$$

On the other hand, using the maxima obtained in the angular velocity of both T and QRS vectors, we determine an alternative computation of this time. We have called it "QT_ω" and it is computed as Eq. 8, also corrected with RR interval. This cardiomarker improves the sensitivity of QT interval measurement, since it does not depend on the T-wave end point.

$$QT_\omega \;=\; \frac{(\text{LAST}_{max}(\|\omega_n\|_2)^T \;-\; \text{FIRST}_{max}(\|\omega_n\|_2)^Q)}{\sqrt{RR}} \tag{8}$$

2.3.3. *PCA features*

Principal Component Analysis (PCA) is a widely used technique in ECG and VCG studies [22]. This technique usually seeks to obtain a triad of orthogonal axes that lie along the direction of maximum variation in the data. This is useful for assessing spatial dispersion alterations and in some cases reducing dimensions with redundant information. In this case, the blocking of the ion channels produces morphological alterations in the dynamics of the currents, which we seek to capture with the linear and angular velocities, and the magnitude of the cardiac electrical vector. Since these signals are three-dimensional, the aim of applying PCA is based on the possibility of studying dispersion that arises in the signals when supplying the drug.

In this work, we apply PCA through a singular value decomposition [22]. We hypothesize that the increase in velocity dispersion should involve an increase in the energy of the second eigenvalue as a balancing response to the decrease in the first. We denote the eigenvalues as $\lambda_1 \geq \lambda_2 \geq \lambda_3$ and the sum of these three quantify the total energy of the corresponding signal (ω, v, E). Then, we define the fraction of energy that each principal component represents as:

$$\lambda_{iS} = \frac{\lambda_i}{\sum_{j=1}^{3} \lambda_j} \cdot 100 \quad , \quad i = 1, 2, 3 \tag{9}$$

where S is the corresponding signal (ω, v, E).

2.4. *Population analysis*

We seek to differentiate two populations: subjects without history of TdP ($-TdP$) and patients with documented TdP in the context of a drug with QT-prolonging potential ($+TdP$). In consequence, we have analyzed three instances using the biomarkers described above. The first represents the baseline-value of the features before the drug supply (**baseline dataset**); the second represents the value of the features during the peak concentration of blood sotalol (**sotalol-peak dataset**); and the third represents the induced changes in the features computed as the differences between the previous two instances (**drug-effect dataset**). The latter represents the most important part of this study because it provides information about changes in ventricular depolarization and repolarization caused by sotalol supply.

Since we have a large number of features which may be potential candidates for becoming part of a classifier, first we need to select the more representative ones. Then, we can try to use a linear classifier combining some of them. However, if it is not possible to reach the classification objective with a straight line, more complex methods should be used. We applied the following procedure.

(1) A statistical significance study is carried out using a two-sided Wilcoxon signed rank test for the three datasets. If a certain parameter reaches a value $p < 0.05$ there are significant differences between $-TdP$ and $+TdP$ populations.

(2) Additionally, for **drug-effect dataset**, a single-column sign test is computed in order to evaluate significant drug effects in each group.

(3) By means of a combination of the significant features, a linear classifier is sought that allows the separation of $+TdP$ and $-TdP$ populations with the best possible accuracy.

(4) If the accuracy is less than 90% we apply various methods of attribute selection for looking candidates for non-linear classifiers. Attribute or feature selection is an effective way to reduce high dimensionality of problems in machine learning which can further reduce computation time, improve learning accuracy, and facilitate a better understanding of learning [23]. We utilized 13 methods implemented in Weka [24], some of which can be evaluated in two ways: as a full training set and as a 10 fold cross-validation and the others evaluated only one way. Then, 19 attribute rankings were obtained. From each ranking we selected 10 best ranked attributes and we counted how many times each attribute was ranked as first, as second, and so on. We assigned 10 points to each attribute for each ranking as first, 9 points for each ranking as second and so on to 1 point for each 10 placement. If an attribute within a certain ranking was not among 10 best no points were assigned. Finally, we counted all points for each attribute and determine final ranks for **baseline**, **sotalol-peak** and **drug-effect** datasets.

(5) Finally, we performed non-linear classifications using the best classifier considering separately every attribute among the 10 best from step 1 and also all combinations of the highest three ranked attributes. The best classifier is selected applying the Auto-Weka mechanism which treats many of the of Weka classification algorithms as a single, highly parametric machine learning framework, and using Bayesian optimization to find a strong instantiation for a given dataset [25].

3. Results

The dynamic study of the cardiac electrical vector has been carried out by analyzing its radius and the linear and angular velocities. Additionally, the relative areas and maximums of each signal has been studied along with the energies of the loops obtained by PCA. Finally, this task has been done for the T-wave loop as well as for each of its halves, which were identified as early and late repolarization. The loop of the QRS complex has also been studied from its areas and absolute maxima. This resulted in a total of 43 dynamic features which can be candidates for use in a classifier of $-TdP$ and $+TdP$ populations. The QTc standard has been computed too and an alternative index QT_ω has been proposed for computation from the maxima angular velocities of the T-wave and QRS complex loops. Both were corrected by the Bazzet formula as well as the classic Tpe index.

We have done the assessment of the indices in the ECG data coming from the instance prior to the drug supply (**Baseline** dataset) and during the peak concentration of blood sotalol (**Sotalol** dataset). Also, we have studied the sotalol-induced changes in the indices (**Drug-effects** dataset). The results observed in the statistical analysis (see Sec. 2.4) are shown in Table 2. Herein, we show two tests. The first is the cross Wilcoxon test between $+TdP$ and $-TdP$. A $p < 0.05$ indicates significant differences and the mean and standard deviation of the index in each population appear bolded. The second test evaluates if the data in the index come from a distribution whose median is zero or not. A significant level $p < 0.05$ implies that the induced effect of the drug increases or decreases the index values.

We have observed that in baseline dataset only three indices shown significant differences. The QT_ω index includes both depolarization and repolarization times; however, parameters $A^T_{\omega x}$ and A^{ER}_{Ex} suggest that differences occur in the T-wave, particularly in early repolarization. Since these alterations are found prior to conducting the study, they may be useful for finding signs of latent repolarization abnormalities.

Regarding the indices during the peak concentration of blood sotalol, QT_ω and QTc shown significant differences between $+TdP$ and $-TdP$. Additionally, significant alterations appeared in the angular velocity indices during the early repolarization phase. The changes of the main components ${\lambda_1}^{ER}_\omega$ and ${\lambda_3}^{ER}_\omega$ suggest an increase in the roundness of the angular velocity loop in at-risk patients, which is accompanied by a reduction in the maximum of the early repolarization vector M^{ER}_ω. This may be linked to drug-induced alterations in the conductivity of myocardial tissue by blocking K^+ ion channels.

Significant sotalol-induced changes are shown in the Drug-effect column of the Table 2. Here, changes has been detected in both QT_ω and QTc. However, only QT_ω shows significance in the discrimination between $+TdP$ and $-TdP$ groups. Moreover, several dynamic features shown differential effects both in the T-wave and in the QRS complex. These indices are promising features for classification.

In the next step, considering the attribute candidates from Table 3, we have studied the possible linear classifications in the three datasets in order to differentiate the at-risk patients. High accuracy has been achieved only for Drug-effect data. The result is shown in Fig. 2. As can be seen, a pair sensitivity/specificity of 100%/100% is achieved with a simple linear division.

Continuing with the procedure, we have then address a different approach from the statistical study by assessing several attribute selection methods as it was explained above. This gives possibility of evaluate from a different perspective the significance of the features and so to extend the results in order to obtain highly sensitive biomarkers in the datasets where no linear classifier could be found. The results of the best ten candidates in each instance is shown in the Table 3a. Also, the indices obtained from early repolarization signal (first half of the T-wave) were best qualified than those from late repolarization (second half) excepting the $A^{LR}_{\omega y}$ index.

Finally, with these results we then applied the mechanisms described in Section 2.4 in order to find an optimal classification. Considering the best candidates from Table 3a we have tested classification perfor-

Index	Baseline		Sotalol-peak		Drug-effect	
	$+TdP$	$-TdP$	$+TdP$	$-TdP$	$+TdP$	$-TdP$
QT_ω	$\mathbf{432.8 \pm 27.4}$	$\mathbf{410.8 \pm 18.8}$	$\mathbf{498.1 \pm 31.5}$	$\mathbf{453.9 \pm 44.2}$	‡ $\mathbf{65.4 \pm 20.6}$	‡ $\mathbf{43.0 \pm 31.5}$
$A^T_{\omega x}$	$\mathbf{53.7 \pm 17.3}$	$\mathbf{42.2 \pm 15.6}$	-	-	-	-
A^{ER}_{Ex}	$\mathbf{56.9 \pm 15.2}$	$\mathbf{68.1 \pm 13.3}$	-	-	-	-
${\lambda_2}^{ER}_{\omega}$	-	-	$\mathbf{10.8 \pm 12.4}$	$\mathbf{2.5 \pm 4.0}$	$*$ $\mathbf{4.4 \pm 10.4}$	$\mathbf{-1.6 \pm 8.3}$
M^{ER}_{ω}	-	-	$\mathbf{3.9 \pm 10.1}$	$\mathbf{16.5 \pm 26.6}$	-	-
A^{QRS}_{E}	-	-	$\mathbf{46.1 \pm 4.5}$	$\mathbf{44.9 \pm 11.2}$	$*$ 2.2 ± 2.8	$*$ 1.3 ± 2.2
${\lambda_1}^{ER}_{\omega}$	-	-	$\mathbf{87.2 \pm 13.3}$	$\mathbf{96.8 \pm 4.7}$	$*$ -5.3 ± 11.7	1.5 ± 9.2
${\lambda_3}^{ER}_{\omega}$	-	-	$\mathbf{1.9 \pm 1.7}$	$\mathbf{0.7 \pm 0.9}$	-	-
QTc	-	-	$\mathbf{536.6 \pm 29.0}$	$\mathbf{501.2 \pm 48.2}$	‡ 55.7 ± 16.7	$*$ 37.4 ± 45.6
$A^T_{\omega y}$	-	-	-	-	$*$ $\mathbf{-15.6 \pm 28.6}$	$\mathbf{6.5 \pm 12.3}$
$A^T_{\omega z}$	-	-	-	-	† $\mathbf{16.6 \pm 19.1}$	$\mathbf{-5.6 \pm 21.8}$
$A^{ER}_{\omega y}$	-	-	-	-	$\mathbf{-3.6 \pm 20.9}$	$\mathbf{9.1 \pm 21.7}$
$A^{LR}_{\omega y}$	-	-	-	-	$\mathbf{-15.7 \pm 28.7}$	$\mathbf{2.3 \pm 8.3}$
M^{QRS}_{E}	-	-	-	-	† $\mathbf{-59.2 \pm 64.7}$	$*$ $\mathbf{24.2 \pm 34.1}$
M^{QRS}_{v}	-	-	-	-	$*$ $\mathbf{-2.07 \pm 25.5}$	$\mathbf{0.13 \pm 30.2}$
Tpe	-	-	-	-	$*$ 15.5 ± 35.1	$*$ 13.5 ± 32.6
M^T_v	-	-	-	-	† -3.11 ± 6.59	-0.96 ± 3.91
M^T_ω	-	-	-	-	-9.0 ± 30.4	$*$ 18.2 ± 35.6
${\lambda_1}^T_\omega$	-	-	-	-	1.7 ± 17.1	$*$ 4.2 ± 8.3
${\lambda_2}^T_\omega$	-	-	-	-	-1.9 ± 16.2	$*$ -4.2 ± 6.7
A^{ER}_E	-	-	-	-	-0.6 ± 8.9	$*$ -2.3 ± 5.3
A^{ER}_{Ez}	-	-	-	-	-0.5 ± 15.6	$*$ 5.8 ± 9.1
A^{LR}_E	-	-	-	-	-0.5 ± 9.0	$*$ 2.3 ± 5.4

Table 2
Features with statistical significance in either of the two tests. (1) Cross test between $+TdP$ and $-TdP$ (bolded). (2) Single test for drug-effect dataset: ‡ for $p < 0.0005$, † for $p < 0.005$ and $*$ for $p < 0.05$.

mance of each one and different combinations as well as different classifiers. As expected, a variety of combinations were obtained for each instance. In Table 3b, we show the results of the best positioned classifier, AdaBoost M1, which reached 100% accuracy in Baseline data (combining M^{ER}_ω, $A^{ER}_{\omega y}$ and $A^{ER}_{\omega x}$) and 91% accuracy in Sotalol-peak data (combining M^{ER}_ω and A^{ER}_{Ex}). AdaBoost M1 is a boosting algorithm used to improve the performance of decision trees on binary classification problems [25]. Here it can be seen that although QT_ω is the best positioned in the attribute selection step, the combination of the second and the third achieves better accuracy.

4. Discussion

It becomes clear that ion channels blocking in the membrane of myocytes modifies the global conduction of the heart and thus the cardiac electrical gradient. In hearts with a certain risk substrate it is possible that the alterations of this gradient are modified differently to a healthy heart. In this sense, we proposed a study of the dynamics of the cardiac electrical vector in order to detect the latent depolarization and repolarization abnormalities that differentiate at-risk patients from subjects without a cardiac history. These abnormalities

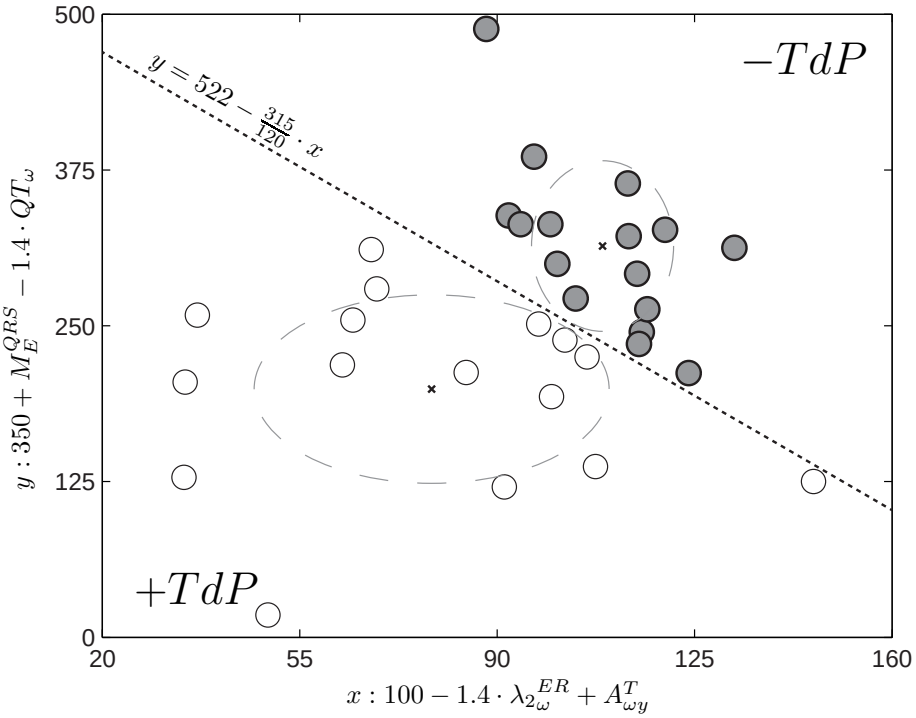

Figure 2. Territorial map showing both groups in drug-effect dataset bounded by dotted line. Centroids are marked along with a standard deviation ellipse. Unfilled circles indicate at-risk patients ($+TdP$) and filled ones indicate subjects without history of Torsade de Pointes ($-TdP$). Herein, classification using four indices achieves an accuracy of 100%

could be associated to arrhythmogenic factors, particularly, risk of TdP.

Because of his high sensitivity, the QTc interval is currently the major index used in decision of drug-induced TdP risk. However, its low specificity is still inadequate. Alternatively, we proposed to use a quaternion method [9] to obtain a different biomarker associated with the angular velocity maxima, so-called QT_ω, which measures the same time interval than QTc. Since this measurement is independent of the waves onset and offset, the accuracy would increase. Its effectiveness can be seen in Table 2, where QT_ω shows significant differences between $+TdP$ and $-TdP$ groups in the three instances while QTc only in **Sotalol-peak** dataset. On the other hand, although QTc and QT_ω could be of great help in the task of making a decision on drug-induced TdP risk, both of them show drug-induced changes at $-TdP$ and $+TdP$, as it can be noted in Table 2. This suggests the need to use alternative parameters that evaluate dispersion changes beyond temporal phenomena as it was highlighted previously [2].

In this sense, we shown a set of indices extracted from the dynamics of the cardiac electrical vector, particularly associated with the radius and the linear and angular velocities. Table 2 provides a variety of features with significant drug-induced changes and high statistical significance between $-TdP$ and $+TdP$. The **Baseline** features of $+TdP$ patients have shown more extended QT_ω. The three features which shown significant differences between the two groups suggest that there are latent abnormalities hidden in the early repolarization phase. Also, in **Sotalol-peak** dataset, $+TdP$ subjects show even more prolonged QT_ω as it was expected considering the Class III effects of Sotalol. Furthermore, most of the features are linked to first half of the T-wave and no significance is found in late repolarization phase. These results are consistent with [4]. Furthermore, the features associated with the linear velocity were only found in **Drug-effect** dataset. These results agree with a recent study that suggests that linear T-wave velocity may be useful in discrimination of cardiac drug effects [26]. However, we found greater sensitivity in the features obtained from the angular velocity. In this latter dataset, it was possible to separate $+TdP$ from $-TdP$ population with a linear classifier achieving 100% accuracy (see Fig. 2).

Regarding classification in **Baseline** and **Sotalol-peak** datasets, more complex algorithms were needed to explore some solutions to at-risk patients separation task. Table 3a shows that many features are better

(a)

Rank	Baseline		Sotalol-peak	
	Feature	Points	Feature	Points
1	M_{ω}^{ER}	93	QT_{ω}	136
2	$A_{\omega y}^{ER}$	72	M_{ω}^{ER}	108
3	$A_{\omega x}^{ER}$	64	A_{Ex}^{ER}	79
4	$A_{\omega z}^{ER}$	58	${\lambda_1}_{\omega}^{ER}$	71
5	$A_{\omega x}^{T}$	54	A_{E}^{QRS}	66
6	QT_{ω}	53	QTc	58
7	QTc	44	$A_{\omega x}^{ER}$	46
8	${\lambda_1}_{\omega}^{ER}$	36	${\lambda_3}_{\omega}^{ER}$	43
9	$A_{\omega z}^{T}$	32	${\lambda_3}_{\omega}^{LR}$	40
10	A_{ω}^{ER}	24	$A_{\omega y}^{ER}$	36

(b)

Instance	Feature	Acc	RMSE
Baseline	M_{ω}^{ER}		
	$A_{\omega y}^{ER}$	100%	0.0684
	$A_{\omega x}^{ER}$		
Sotalol-peak	M_{ω}^{ER}	92%	0.2804
	A_{Ex}^{ER}		

Table 3
(a) Feature ranking obtained from the algorithm that considers the score of 19 attribute selection methods. (b) Best combination obtained for the best classifier selected: AdaboostM1.

positioned than QTc. Herein, PCA biomarkers shown high accuracy and also for both instances the early repolarization biomarkers seem to be more efficient. In fact, only with the combination of features of the first half of the T-wave, very high accuracy values were achieved. It is important to highlight that these models trained to classify the combinations of the features are not linear, so we can not define a simple linear split between the $+TdP$ and $-TdP$ populations unlike the case of **Drug-effect** dataset.

It should be noted that we have exploited the potential of a three-dimensional study of the cardiac electrical vector. Consequently, this work requires the use of VCG signals. Despite all the advantages of VCG that are nowadays known [27], it is not so widely used in clinical practice unlike ECG which is the most common method for initial cardiac diagnosis. On the other hand, there are several possible ways to reconstruct this orthogonal system from ECG leads. One of these is the Kors regression transformation, which is used herein and that it is being extensively used in many others ongoing lines of research, including myocardial infarction diagnosis [28], hypertrophy [29] and ischaemic heart diseases [9,30]. Furthermore, the Kors matrix is the one with the best accuracy compared to alternative methods [27]. Holter studies are also turning to three-dimensional perspectives through the use of the EASI lead system [31]. This 5-lead configuration provides a three-dimensional portrayal of the electrical activity of heart, rather than independent channels of unipolar and bipolar energy. Transferring these ideas to the everyday clinical practice could offer substantial improvements in the detection of patients with a predisposition to TdP. Likewise, for the pharmaceutical industry, it would represent a promising proposal to replace current regulatory indices, which have insufficient specificity.

Finally, this study provides relevant information regarding the features that are affected by the drug supply and which explain the latent abnormalities in patients with torsadogenic risk. Although these results

are highly promising, these outcomes should be evaluated in the future with more patients and with drugs that block different ion channels.

5. Conclusion

Through an in-depth study of the dynamics of the cardiac electrical vector during a sotalol challenge, alternative biomarkers were obtained that were able to describe pre-existing risk of patients with a history of Torsade de Pointes. By means of a suitable combination of these ones a sensitivity/specificity pair of 100%/100% was reached in **Drug-effect** dataset using a simple linear classifier. These results could contribute to a better understanding of the torsadogenic substrate and to the safe development of drug therapies.

Acknowledgment

This work was supported by CONICET, under project PIP #112-20130100552CO and Agencia MINCYT, under project PICT 2145-2016, Argentina. Also, the authors acknowledge the financial support from the Slovenian Research Agency (research core funding No. P2-0098).

References

[1] W. H. Organization, Noncommunicable diseases country profiles, www.who.int (2014).
[2] A. Camm, Hopes and disappointments with antiarrhythmic drugs, Int J Cardiol 237 (2017) 71–74.
[3] R. Woosley, K. Romero, C. Heise, T. Gallo, J. Tate, R. Woosley, S. Ward, Adverse drug event causality analysis (ADECA): A process for evaluating evidence and assigning drugs to risk categories for sudden death, Drug Saf 40 (6) (2017) 465–474.
[4] J. Couderc, S. Kaab, M. Hinterseer, S. McNitt, X. Xia, A. Fossa, B. Beckmann, S. Polonsky, W. Zareba, Baseline values and sotalol-induced changes of ventricular repolarization duration, heterogeneity, and instability in patients with a history of drug-induced torsades de pointes, J Clin Pharmacol 49 (2009) 6–16.
[5] D. G. Strauss, G. Gintant, Z. Li, W. Wu, K. Blinova, J. Vicente, T. J.R., S. P.T., Comprehensive in vitro proarrhythmia assay (CiPA) update from a cardiac safety research consortium / health and environmental sciences institute / FDA meeting, Ther Innov Regul 53 (2018) 519–525.
[6] International Council on Harmonization, S7B the non-clinical evaluation of the potential for delayed ventricular repolarization (QT interval prolongation) by human pharmaceuticals, https://database.ich.org/sites/default/files/S7B_Guideline.pdf (2005).
[7] International Council on Harmonization, E14 the clinical evaluation of QT/QTc interval prolongation and proarrhythmic potential for non-antiarrhythmic drugs, https://database.ich.org/sites/default/files/E14_Guideline.pdf (2005).
[8] M. Malik, H. Huikuri, F. Lombardi, G. Schmidt, R. Verrier, M. Zabel, Is the Tpeak-Tend interval as a measure of repolarization heterogeneity dead or just seriously wounded?, Heart Rhythm 16 (6) (2019) 952–953.
[9] P. Cruces, P. Arini, Quaternion-based study of angular velocity of the cardiac vector during myocardial ischaemia, Int J Cardiol 248 (2017) 57–63.
[10] Telemetric, Holter ECG Warehouse, Cardiac patients with and without a history of drug-induced torsades de pointes, http://thew-project.org/Database/E-OTH-12-0068-010.html.
[11] S. Kaab, M. Hinterseer, M. Nabauer, G. Steinbeck, Sotalol testing unmasks altered repolarization in patients with suspected acquired long-QT-syndrome–a case-control pilot study using i.v. sotalol, Eur Heart J 24 (7) (2003) 649–657.
[12] A. Brown, Drugs, hERG and sudden death, Cell Calcium 35 (6) (2004) 543–547.
[13] B. Broux, D. De Clercq, A. Decloedt, L. Vera, M. Devreese, R. Gehring, C. S., G. van Loon, Pharmacokinetics and electrophysiological effects of sotalol hydrochloride in horses, Equine Vet J 50 (3) (2018) 377–383.
[14] P. Cruces, P. Arini, Cardiomarkers of ventricular repolarization based on theory of quaternions, XVIII Reunión de Trabajo en Procesamiento de la Información y Control, http://sgk-ar.com/rpic-2019/index.php/papers/ 18 (2019) 435–440.
[15] J. Proakis, D. Manolakis, Digital signal processing. Principles, algorithms, and applications, Prentice-Hall International, Inc., 1996.
[16] P. Cruces, P. Arini, A novel method for cardiac vector velocity measurement: Evaluation in myocardial infarction, Biomed Signal Proc and Control 28 (2016) 58–62.
[17] J. Couderc, S. Ma, A. Page, C. Besaw, J. Xia, W. Chiu, J. de Bie, J. Vicente, M. Vaglio, F. Badilini, S. Babaeizadeh, C. S. Chien, M. Baumert, An evaluation of multiple algorithms for the measurement of the heart rate corrected JTpeak interval, J Electrocardiol 50 (6) (2017) 769–775.
[18] N. Srinivasan, M. Orini, R. Providencia, R. Simon, M. Lowe, O. Segal, A. Chow, R. Schilling, R. Hunter, P. Taggart, P. Lambiase, Differences in the upslope of the precordial body surface ECG T wave reflect right to left dispersion of repolarization in the intact human heart, Heart Rhythm 16 (6) (2019) 943–951.
[19] C. Antzelevitch, J. Di Diego, Counterpoint tpeak-tend interval as a marker of arrhythmic risk, Heart Rhythm 16 (6) (2019) 954–955.

[20] B. Barsky (Ed.), Rethinking Quaternions. Theory and Computation, Morgan & Claypool, California, 2010.
[21] A. Poznyak, Modelado Matemático de los Sistemas Mecánicos, Eléctricos y Electromecánicos, Pearson, 2005, Ch. 2, pp. 73–83.
[22] F. Castells, P. Laguna, L. Sörnmo, A. Bollman, M. Roig, Principal component analysis in ECG signal processing, EURASIP Journal on Advances in Signal Processing.
[23] C. Jie, L. Jiawei, W. Shulin, Y. Sheng, Feature selection in machine learning: A new perspective, Neurocomp 300 (2018) 70–79.
[24] E. Frank, M. Hall, I. Witten, The WEKA Workbench. Online Appendix for Data Mining: Practical Machine Learning Tools and Techniques, Morgan Kaufmann, 2016.
[25] L. Kotthoff, C. Thornton, H. Hoos, F. Hutter, K. Leyton-Brown, Auto-weka 2.0: automatic model selection and hyperparameter optimization in weka, J Machine Learning Res 18 (1) (2017) 826–830.
[26] W. Bystricky, C. Maier, G. Gintant, D. Bergau, K. Kamradt, P. Welsh, D. Carter, T vector velocity: A new ECG biomarker for identifying drug effects on cardiacventricular repolarization, PLoSONE 14 (7) (2019) 1–22.
[27] R. Jaros, R. Martinek, L. Danys, Comparison of different electrocardiography with vectorcardiography transformations, Sensors 19 (3072) (2019) 1–19.
[28] S. Choudhuri, T. Ghosal, D. Goswami, A. Sengupta, Planarity of the spatial qrs loop of vectorcardiogram is a crucial diagnostic and prognostic parameter in acute myocardial infarction, Med Hypotheses 130 (2019) 109251.
[29] D. Cortez, T. Schlegel, M. Ackerman, J. Bos, ECG-derived spatial QRS-T angle is strongly associated with hypertrophic cardiomyopathy, J Electrocardiol 50 (2017) 195–202.
[30] D. Cortez, J. Bos, M. Ackerman, Vectorcardiography identifies patients with electrocardiographically concealed long qt syndrome, Heart Rhythm 14 (2017) 894–899.
[31] M. Jahrsdoerfer, K. Giuliano, Clinical usefulness of the easi 12-lead continuous electrocardiographic monitoring system, Clin Care Nurse 25 (5) (2005) 28–38.