



Importance of the heart vector origin point definition for an ECG analysis: The Atherosclerosis Risk in Communities (ARIC) study



Erick Andres Perez-Alday^a, Yin Li-Pershing^a, Aron Bender^a, Christopher Hamilton^a, Jason A. Thomas^a, Kyle Johnson^a, Tiffany L. Lee^a, Ryan Gonzales^b, Aaron Li^c, Kelley Newton^a, Larisa G. Tereshchenko^{a,*}

^a Knight Cardiovascular Institute, Oregon Health & Science University, Portland, OR, USA

^b Concordia University, Portland, OR, USA

^c Carleton College, Northfield, MN, USA

ARTICLE INFO

Keywords:

Electrocardiography
Sudden cardiac death
Vectorcardiography
Electrocardiology
Vectorcardiogram
Median beat
Origin point

ABSTRACT

Aim: Our goal was to investigate the effect of a global XYZ median beat construction and the heart vector origin point definition on predictive accuracy of ECG biomarkers of sudden cardiac death (SCD).

Methods: Atherosclerosis Risk In Community study participants with analyzable digital ECGs were included ($n = 15,768$; 55% female, 73% white, mean age 54.2 ± 5.8 y). We developed an algorithm to automatically detect the heart vector origin point on a median beat. Three different approaches to construct a global XYZ beat and two methods to locate origin point were compared. Global electrical heterogeneity was measured by sum absolute QRST integral (SAI QRST), spatial QRS-T angle, and spatial ventricular gradient (SVG) magnitude, azimuth, and elevation. Adjudicated SCD served as the primary outcome.

Results: There was high intra-observer ($\kappa = 0.972$) and inter-observer ($\kappa = 0.984$) agreement in a heart vector origin definition between an automated algorithm and a human. QRS was wider in a median beat that was constructed using R-peak alignment than in time-coherent beat (88.1 ± 16.7 vs. 83.7 ± 15.9 ms; $P < 0.0001$), and on a median beat constructed using QRS-onset as a zeroed baseline, vs. isoelectric origin point (86.7 ± 15.9 vs. 83.7 ± 15.9 ms; $P < 0.0001$). ROC AUC was significantly larger for QRS, QT, peak QRS-T angle, SVG elevation, and SAI QRST if measured on a time-coherent median beat, and for SAI QRST and SVG magnitude if measured on a median beat using isoelectric origin point.

Conclusion: Time-coherent global XYZ median beat with physiologically meaningful definition of the heart vector's origin point improved predictive accuracy of SCD biomarkers.

1. Introduction

According to the American Heart Association, out-of-hospital Sudden Cardiac Death (SCD) is the most frequent and the most dramatic mode of cardiovascular death, with a survival rate of about 10% [1]. The electrocardiogram (ECG) is a ubiquitous tool in clinical practice. Every day physicians rely on (semi)automatically measured ECG parameters such as corrected QT interval (QTc), and QRS duration to guide clinical decisions. QTc [2] and other ECG metrics [3] were included in the risk models of SCD. Our group recently developed a global electrical heterogeneity (GEH) risk score of SCD [4] and showed that addition of GEH ECG metrics to traditional clinical risk factors significantly improved reclassification of SCD risk [5].

It is well established that accuracy and reproducibility of ECG

measurements are significantly improved if measurements are performed on a multi-lead, globally representative median beat [6]. However, it is unknown whether different approaches of a global multi-lead formation can affect VCG and ECG measurements. The Recommendations for the Standardization and Interpretation of the ECG [6] suggested that “a comparative study is needed of global measurements made by different methods”.

The ECG measurements are based in the lead theory which assumes that the heart is a dipole vector (heart vector) that is moving during the cardiac cycle in three-dimensional (3D) space, forming vectorcardiographic (VCG) loops [7]. Projection of the heart vector movement on any ECG lead axis produces the waveform measured in that lead. The electrical origin of the heart vector has important electrophysiological and biophysical meaning, as it is defined by the isoelectric state of the

* Corresponding author. 3181 SW Sam Jackson Park Rd, UHN62, Portland, OR, 97239, USA.

E-mail address: tereshch@ohsu.edu (L.G. Tereshchenko).

heart [8]. The electrocardiographic method, developed by Einthoven and colleagues, quantifies the magnitude and direction of the resultant electromotive force produced by the heart at a given time instance during the cardiac cycle. In 1938, Wilson and Johnston highlighted the importance of the heart vector origin point (see Fig. 3 from Ref. [8]), which corresponds to the electrocardiographic baseline. Due to technological limitations in the 1930s they were not able to define the origin point, and noted that “details are lost in the uniformly black area which surrounds the isoelectric point” [8]. Modern ECG signal resolution allows characterization of its fine features. However, contemporary ECG standards ignore VCG origin definition [6,9].

This study's goals are (1) to define the heart vector origin point, (2) to compare global measurements made by different approaches to global median beat construction and VCG origin point definition, and (3) compare differences in predictive accuracy of ECG and VCG biomarkers of SCD in the Atherosclerosis Risk in Community (ARIC) study participants.

2. Methods

The algorithms and software used to calculate median beat and define the heart vector origin point, and software code are provided at <https://github.com/Tereshchenkolab/Origin>. The data underlying our work can be obtained via ARIC Coordinating Center at the University of North Carolina—Chapel Hill, as described online [10]. Also, most ARIC data can also be obtained from the National Heart, Lung, and Blood Institute (NHLBI) – maintained BioLINCC repository [11].

2.1. Study population

The Atherosclerosis Risk in Communities (ARIC) study is an ongoing, prospective cohort study which was created to evaluate risk factors, progression, and outcomes of atherosclerosis in 15,792 participants (45% male, 74% white) enrolled in the United States in 1987–1989. The ARIC study protocol and design have been previously described [12]. The study protocol was approved by institutional review boards (IRB) at each field center, and all participants signed informed consent. In addition, this study was approved by the Oregon Health & Science University IRB.

In this study, we included all ARIC participants with available and analyzable digital ECGs ($n = 15,768$). Participants with absent or poor quality ECGs (due to noise, artifacts, missing leads; $n = 24$) were excluded.

SCD was defined as a sudden pulseless condition in a previously stable individual without evidence of a non-cardiac cause of cardiac arrest, if cardiac arrest occurred out of the hospital or in the emergency room. To identify cases of SCD in ARIC, cases of cardiovascular death that occurred by December 31, 2012, were reviewed and adjudicated by a committee of electrophysiologists, general cardiologists, and internists, as previously described [13].

2.2. Electrocardiogram analyses

All resting 12-lead ECGs from the first baseline visit of the ARIC study were analyzed [4]. The sampling rate of ECG signals was 500 Hz and amplitude resolution was $1 \mu\text{V}$.

2.3. Construction of a representative global XYZ median beat

A three-step approach was used for construction of all the comparison beats (Fig. 1):

2.3.1. Beat labeling and dominant type beat selection

First, each beat on a 10-s 12-lead ECG recording was manually labeled by the team of physicians, as recently described [14]. Non-analyzable beats due to noise and artifacts, and non-dominant ectopic beats

were excluded from analysis. Only normal sinus beats in a 10-s recording were included in the construction of a normal sinus median beat. Custom software for VCG analyses was written in MATLAB (MathWorks, Natick, MA, USA). Kors transformation matrix was used to obtain the orthogonal X, Y, and Z leads from 12-lead ECG [15].

2.3.2. Single-lead median beat construction

At the next step, all included beats on a single lead, V_x , were then aligned to construct a representative single-lead median beat, V_{MX} . The maximum absolute value of the QRS complex, maximum absolute of the first derivative (maximum $|\text{dV}/\text{dt}|$), and QRS onset were used to construct the different types of representative single-lead median beat (Fig. 1A). After alignment, the representative median beat, V_{MX} , was constructed using the *median* function of MATLAB (MathWorks, Natick, MA, USA). We also measured correlation between time-coherent global XYZ median beats constructed using single-lead X alignment by QRS peak and maximum $|\text{dV}/\text{dt}|$.

2.3.3. Global XZ multi-lead beat construction

At the third step, we constructed a global XYZ median beat. We compared three different types of global beats:

- (1) R-aligned: single-lead alignment at the maximum value of the QRS complex was performed on all 3 leads X, Y, and Z, with subsequent multi-lead alignment by QRS peak;

$$V_{MX_QRS\text{peak}} = V_{MX} \left(\text{centered at } \max_{QRS} V_{MX} \right)$$

$$V_{MY_QRS\text{peak}} = V_{MY} \left(\text{centered at } \max_{QRS} V_{MY} \right)$$

$$V_{MZ_QRS\text{peak}} = V_{MZ} \left(\text{centered at } \max_{QRS} V_{MZ} \right)$$

- (2) Q-aligned: single-lead alignment at the beginning of QRS complex was performed on all 3 leads X, Y, and Z, with subsequent multi-lead alignment by QRS onset;

$$V_{MX_QRS\text{onset}} = V_{MX} (\text{centered at } QRS_{\text{onset}} (V_{MX}))$$

$$V_{MY_QRS\text{onset}} = V_{MY} (\text{centered at } QRS_{\text{onset}} (V_{MY}))$$

$$V_{MZ_QRS\text{onset}} = V_{MZ} (\text{centered at } QRS_{\text{onset}} (V_{MZ}))$$

- (3) time-coherent: single-lead alignment was performed only on a single lead (X). Corresponding time points on leads Y and Z were used for construction of time-coherent global XYZ median beat (Fig. 1B).

$$V_{MX_Time} = V_{MX} \left(\text{centered at } \max_{QRS} \left| \frac{\partial V_{MX}}{\partial t} \right| \right)$$

$$V_{MY_Time} = V_{MY} \left(\text{centered at } \max_{QRS} \left| \frac{\partial V_{MY}}{\partial t} \right| \right)$$

$$V_{MZ_Time} = V_{MZ} \left(\text{centered at } \max_{QRS} \left| \frac{\partial V_{MZ}}{\partial t} \right| \right)$$

2.4. Vectorcardiogram origin point definition

We defined VCG origin point based on its electrophysiological and biophysical meaning: the electrically quiet, or isoelectric state of the heart when the heart vector does not move in 3D space. The electrically silent period was defined as the segment between the end of repolarization and the onset of P-wave (if present), or the beginning of QRS

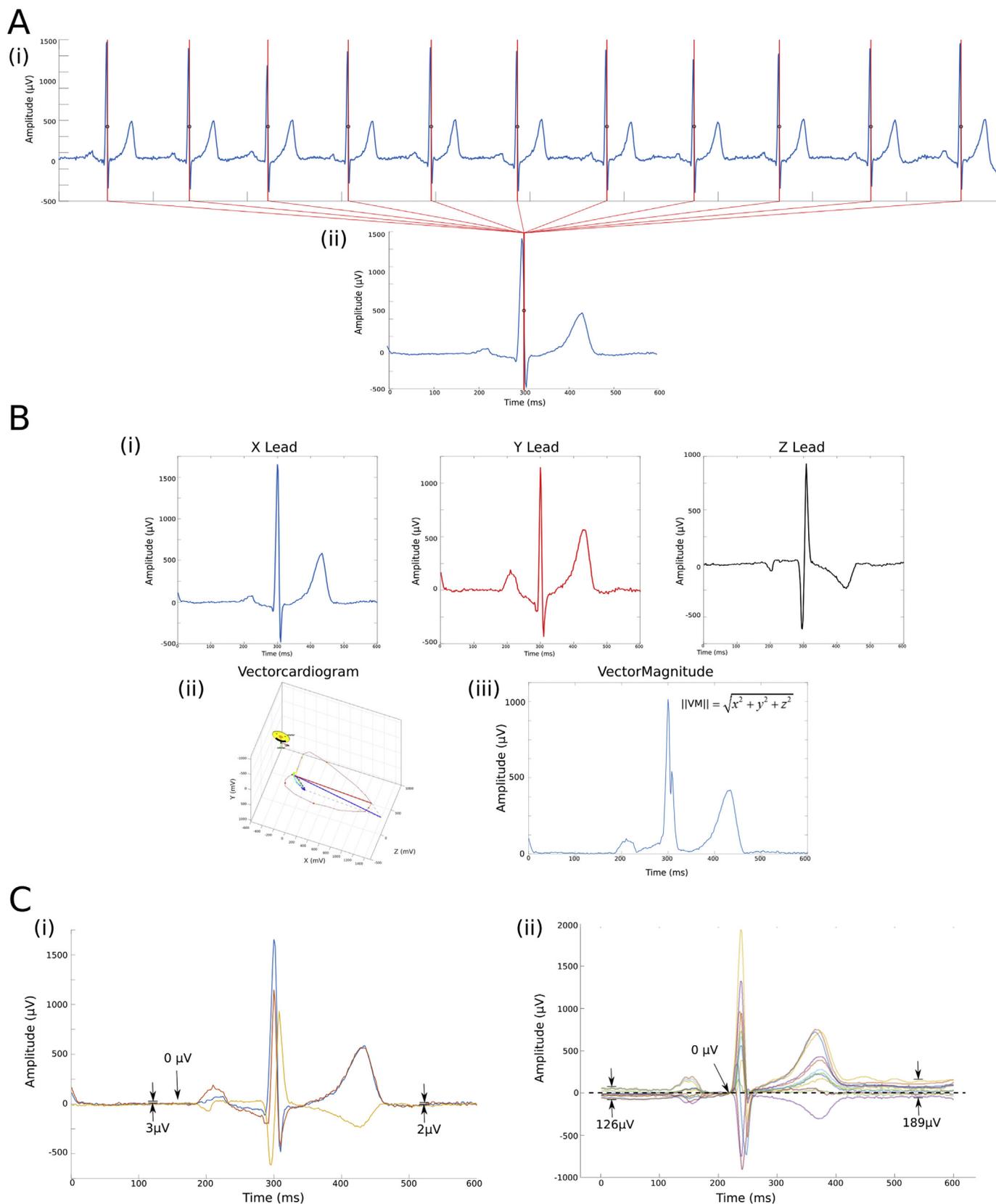


Fig. 1. Construction of the Vector-Magnitude. (A) (i) Normal sinus beats with the absolute maximum $|dV/dt|$ marked in each beat are used to (ii) construct the median single-lead (X-lead) beat. Beats are aligned by maximum $|dV/dt|$. (B) Time-coherent global XYZ (i) multi-lead beat construction. Resultant vectorcardiogram (ii) and vector magnitude (iii) are shown. (C) Vectorcardiogram origin point definition, using (i) the physiological approach with the isoelectric heart vector origin definition, (ii) current industry standard, defined baseline at the beginning of QRS complex.

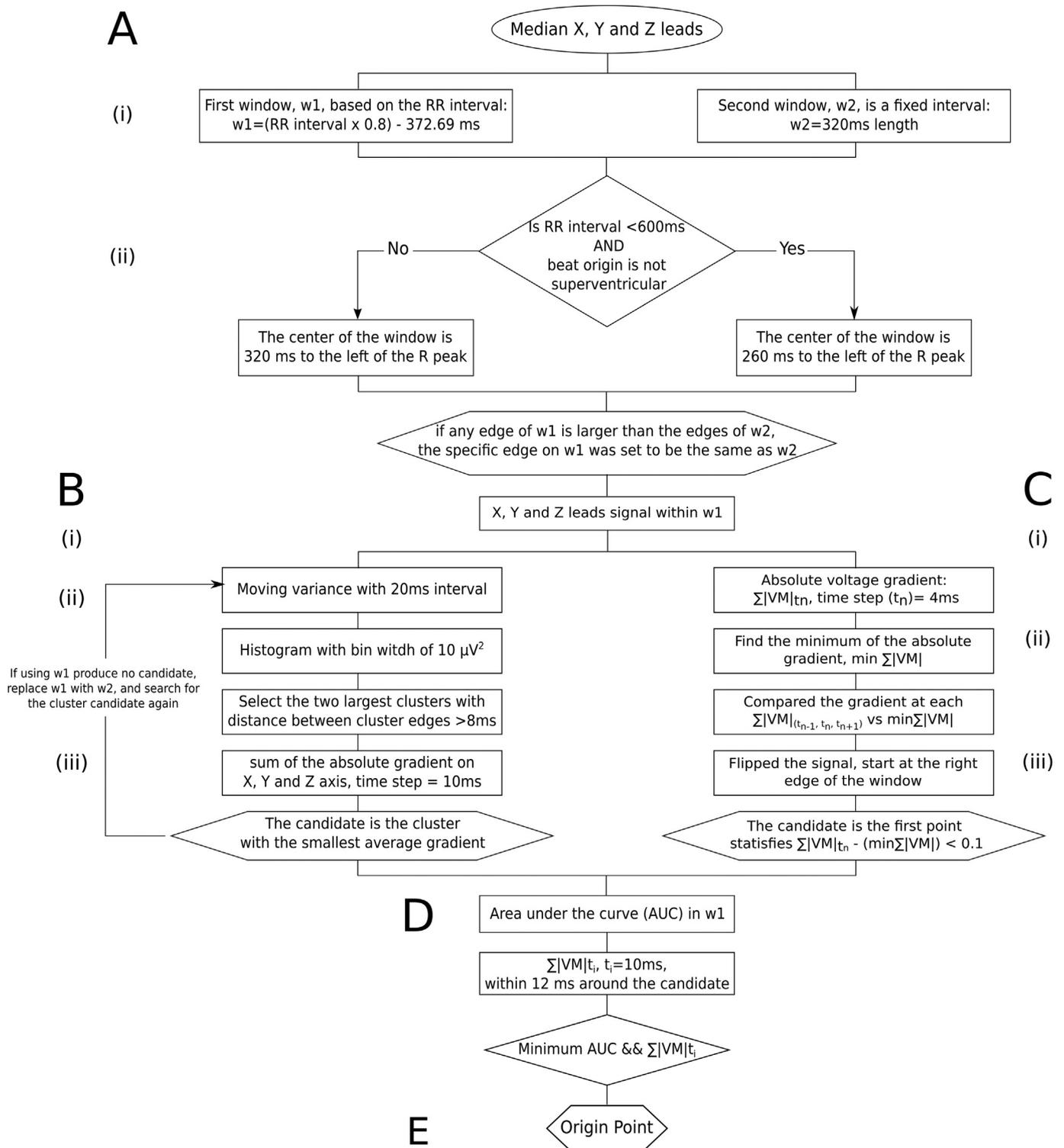


Fig. 2. Schematic of the algorithm to automatically detect the electrical origin point.

complex. The Vector-Magnitude (VM) was obtained by computing the Euclidean norm, $\|VM\| = \sqrt{V_{MX}^2 + V_{MY}^2 + V_{MZ}^2}$, for each global XYZ media beat constructed ($V_{M_QRSpeak}$, $V_{M_QRSonset}$, and V_{M_time}). The definition of the VCG origin point determines the definition of the VM ECG baseline. We compared our definition of the VCG origin point with the current industry standard [16], which defines the zero value of an ECG baseline in a representative median beat at the beginning of a QRS (Fig. 1C).

2.5. Algorithm to detect the VCG electrically quiet origin point

We developed an algorithm to detect the VCG origin point on a median beat and provided the software code at <https://github.com/Tereshchenkolab/Origin>. A schematic flowchart of the algorithm can be seen in Fig. 2.

Two time interval windows were identified to select an isoelectric segment in the TP segment (Fig. 3A). The first window (w_1) was dynamically calculated from the RR' interval using the following equation

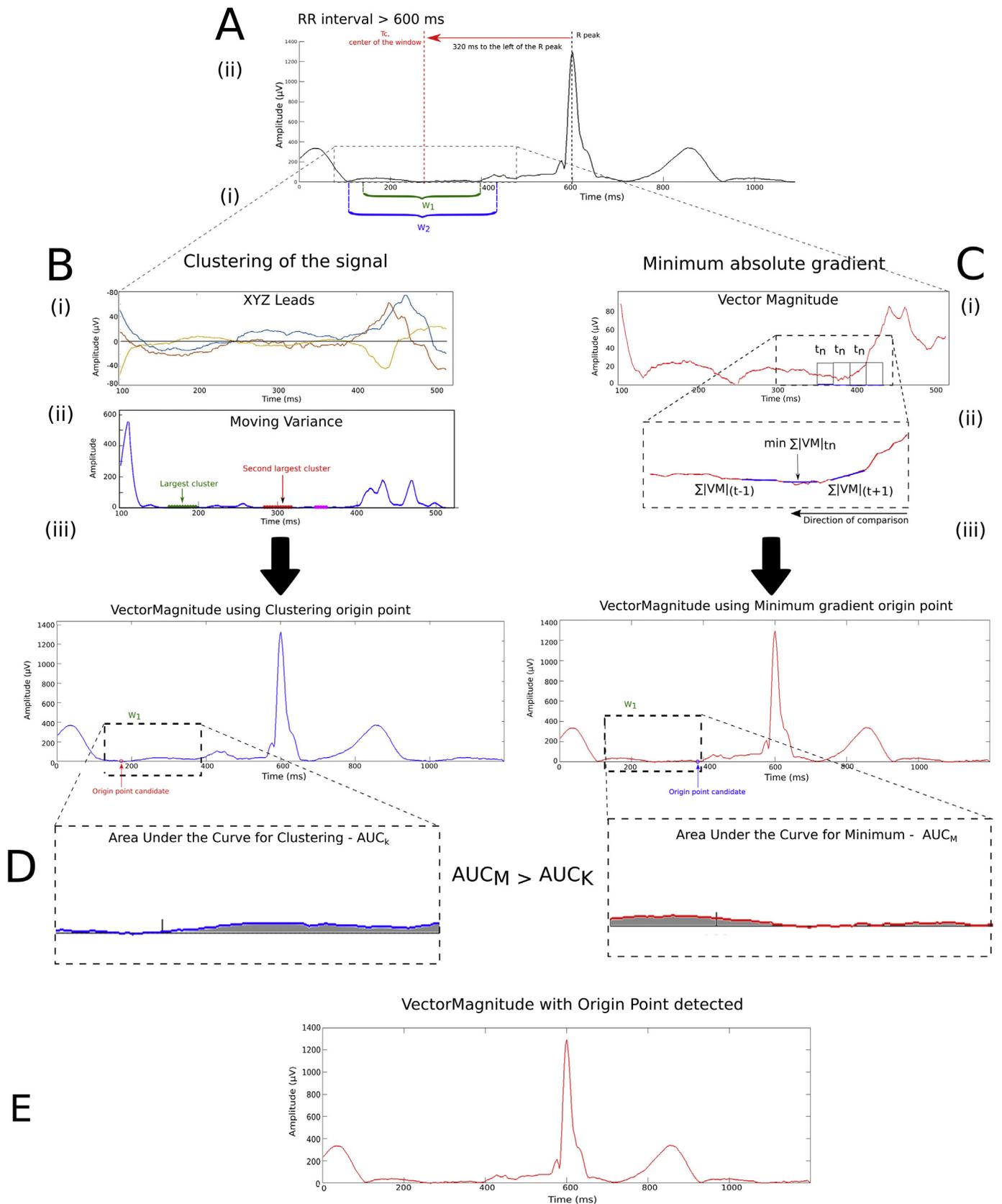


Fig. 3. The process to detect the electrical origin point. (A) Selection of the time interval window. Methods to detect origin point: (B) Clustering the signal with moving variance, and (C) Minimum absolute gradient. The AUC obtained from the minimum absolute gradient, AUC_M , has a lower value compared to the AUC obtained from the clustering method, AUC_C . (D) Therefore, the origin point calculated with the minimum method (red signal) is selected.

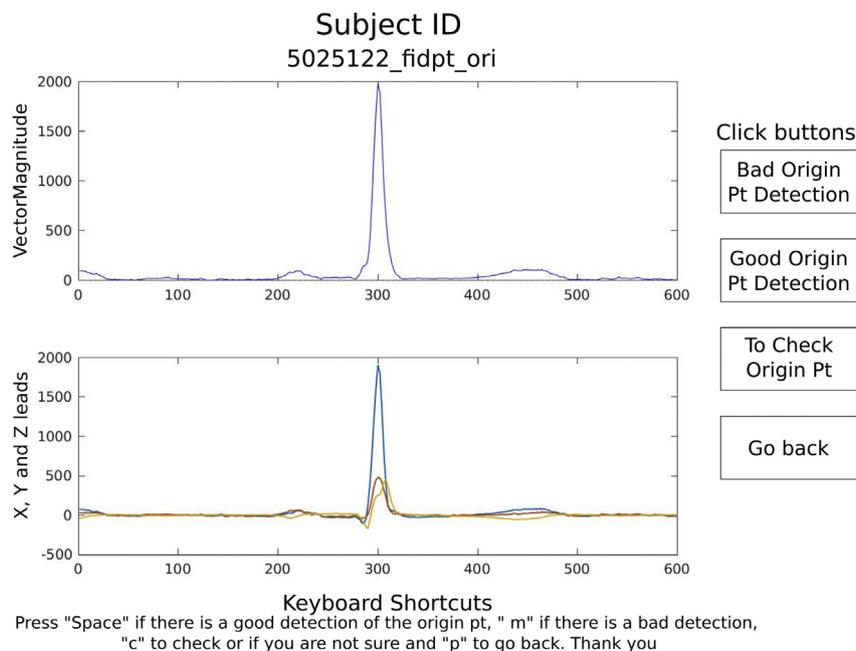


Fig. 4. Viewer to assess origin point detection accuracy. The click buttons (i) and keyboard shortcuts (ii) are included.

(eq. (1)):

$$W_1 = \text{floor}(RR'_{\text{interval}} \times 0.8 - 372.69 * 0.5). \quad (1)$$

The center of the window, T_C , also depended on the RR interval and the type of beat under investigation. Two cases were identified: (1) $T_C = 260$ ms, if RR interval (RR_{int}) < 600 ms and median beat \neq Supraventricular (S). $T_C = 320$ ms, for any other case: $RR_{\text{int}} \geq 600$ ms for any type of beat or $RR_{\text{int}} < 600$ but median beat = S.

$$T_C = \begin{cases} 260 \text{ ms to the left of } R_{\text{peak}}, & \text{if } RR_{\text{int}} < 600 \text{ ms and beat } \neq S \\ 320 \text{ ms to the left of } R_{\text{peak}}, & \text{Any other case} \end{cases} \quad (2)$$

The second window (w_2) was a fixed window of 320 ms with same T_C center (Fig. 3A-ii). If the edges of w_1 were larger than the edges of w_2 , the edges of w_1 were set to be the same as w_2 .

Following selection of the isoelectric baseline, the algorithm used two approaches to select the flattest line on VM and XYZ leads signal within w_1 : clustering of the signal with the least variance in a given window and minimum absolute change of magnitude closest to the end of the window.

2.5.1. Searching for a candidate segment with the least variance

We calculated the variance on the X, Y, and Z lead signals with a moving window of 20 ms ($movvar_{XYZ}$). Using the MATLAB function *movvar*, we identified the segment with the least amount of variance i.e. the flattest portion of the signal (Fig. 3B). Candidate points were identified by looking for the segments containing the largest number of counts from the histogram of $movvar_{XYZ}$ with a fixed bin width of 10 μV^2 . Once candidate points were identified, they were separated into clusters. The clusters were then defined such that the minimum distance between the edges of neighboring clusters was larger or equal to 8 ms in order to reduce the impact of noise within the signal. The algorithm selected two candidate clusters comprised of the largest number of elements with which to perform origin point calculation.

The absolute voltage gradient across X, Y, and Z was calculated with a time step of 10 ms. Then, the average of the sum of the absolute voltage gradient across the X, Y, and Z leads of a given cluster was obtained. The cluster with the smallest average gradient was chosen to best approximate the isoelectric segment, and the origin point was chosen at the median time point of that candidate cluster. If the last

criteria was not met and the algorithm could not identify an origin point candidate within w_1 , the search for an origin point was carried on to w_2 with the algorithm described above.

2.5.2. Searching for a candidate segment with the least change

As an alternative approach, we attempt to identify the isoelectric segment by computing the sum of the absolute voltage gradient on the X, Y, and Z leads (Fig. 3C):

$$\sum_{x,y,z} |\nabla VM| = \left| \frac{dV_x}{dt_n} \right| + \left| \frac{dV_y}{dt_n} \right| + \left| \frac{dV_z}{dt_n} \right| \quad (3)$$

Then, we compared the gradient at any time step Δt_n and their neighboring time points (Δt_{n-1} and Δt_{n+1}) against the minimum of the sum of the voltage absolute gradient (eq. (4)) in the window w_1

$$\min \sum_{x,y,z} |\nabla VM| = \operatorname{argmin}_{t_n: t_n \in w_1} \left(\left| \frac{dV_x}{dt_n} \right| + \left| \frac{dV_y}{dt_n} \right| + \left| \frac{dV_z}{dt_n} \right| \right) \quad (4)$$

The candidate origin point was the middle point of the segment where the comparison was below a threshold of 0.1 mV, to look for a stable isoelectric line. The comparison was performed starting at the right edge of the window w_1 (closest to the R peak), then working backward toward the preceding T wave with a time step of $\Delta t_n = 4$ ms.

2.5.3. Final selection of the origin point

At the final step of the algorithm, the two previously identified candidate origin points were compared using the following criteria: area under the curve of reconstructed VM and sum of average absolute gradient in 10 ms time steps. Individual VMs were reconstructed for each candidate by defining the candidate as the zero-value baseline. The area under the curve for each VM was computed only in window w_1 (Fig. 3D). The sum of the average absolute gradient of each VM, with a time step of 10 ms, was calculated around each candidate. The final heart vector origin point with the smallest area under the windowed VM curve and the minimum average voltage gradient was selected.

2.6. Accuracy of the VCG origin point detection algorithm

Appropriate selection of the VCG origin point was confirmed manually by investigators (CH, RG, AL, KN, and KJ) with the aid of a

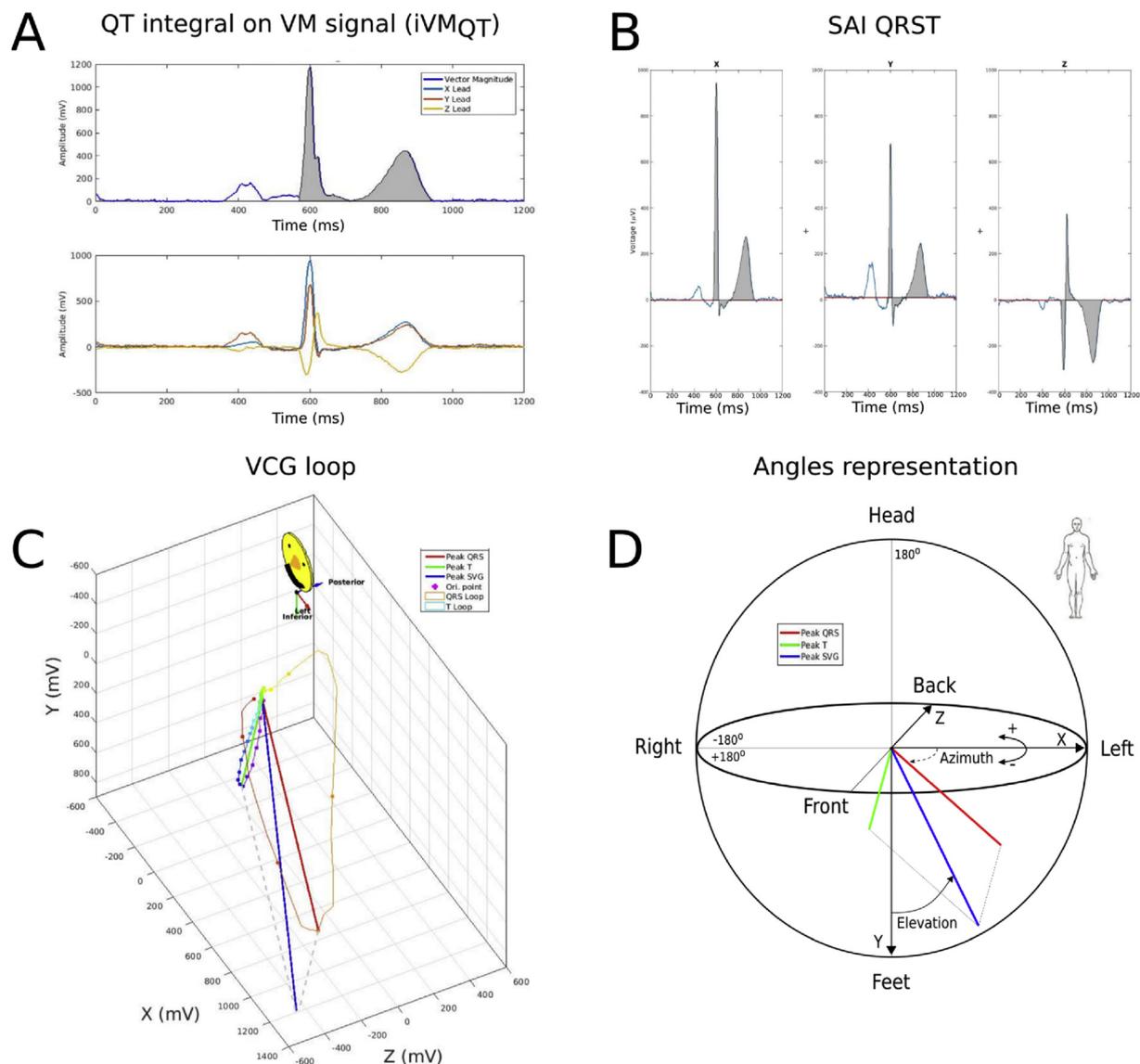


Fig. 5. GEH measurement example of peak and area vectors. (A) Vector Magnitude, and XYZ leads. (B) Sum Absolute QRST integral (SAI QRST) measurement. (C) Vectorcardiographic loops. (D) Peak QRS, T, and spatial ventricular gradient (SVG) vectors, and measurement of SVG direction (azimuth and elevation) is shown.

graphical display (Fig. 4) for all ECGs, separately for different beat construction and origin point methods. Agreement between an automatic detection and investigator's opinion was assessed. Cases of disagreement were reviewed by the 2nd investigator (EAPA), who made a final determination and corrected origin point location as needed. Inter-observer and intra-observer agreement were evaluated.

2.7. Detection of ECG fiducial points on global XYZ Vector-Magnitude

All ECG fiducial points were automatically detected on global XYZ median beat VM, using previously validated algorithms [17]. QRS onset and offset were detected using the Pan and Tompkins approach [17]. T-offset was detected using the curve-length transformation by Zong et al. [18].

2.8. Measurement of global electrical heterogeneity

GEH was measured as previously described [4]. We have provided the software code for GEH analysis at Physionet (<https://physionet.org/physiotools/geh/>). The spatial peak area QRS, T, and SVG vectors were defined (Fig. 5). The scalar value of SVG was measured by the sum

absolute QRST integral (SAI QRST) as previously described [19,20]. In addition, scalar value of SVG was also assessed as a QT integral on Vector-Magnitude signal (iVM_{QT}), or an area-under-the-Vector-Magnitude signal curve from the QRS-onset to T-offset [21]. For comparison of the effect of beats alignment and origin point definition on ECG and VCG measurements, we did not validate the accuracy of fiducial points (QRS onset and offset, T offset) measurements on VM.

2.9. Statistical analysis

Statistical analysis was performed using STATA MP 15.1 (StataCorp LP, College Station, TX, USA). The κ -statistic was used to evaluate inter-observer and intra-observer agreement between an automated algorithm and investigators, and between different investigators. Paired multivariate test of means (Hotelling's t-squared statistic) was used to compare ECG and VCG measurements performed on a median beat constructed using 3 different types of beats alignment: by QRS onset, QRS peak, and time-coherent. Paired *t*-test was used to compare ECG and VCG measurements performed on a median beat constructed using two different definitions of origin point: with origin points defined as the isoelectric flattest segment of T_{end}-P_{onset} or located at the QRS onset.

Table 1
Clinical characteristics of study population.

Characteristic	n = 15,768
Age \pm SD, years	54.2 \pm 5.8
Female, n(%)	8696(55.2)
White, n(%)	11,471(72.8)
Diabetes, n(%)	1876(12.0)
Hypertension, n(%)	5498(34.9)
Anti-hypertensive drugs, n(%)	4767(35.0)
Coronary Heart Disease, n(%)	769(4.9)
Heart failure, n(%)	739(4.8)
Stroke, n(%)	278(1.8)
Peripheral artery disease, n(%)	635(4.2)
Atrial fibrillation, n(%)	36(0.2)
Current smoking, n(%)	4120(26.2)
Body-mass-index \pm SD, kg/m ²	27.7 \pm 5.4
Total cholesterol \pm SD, mmol/L	5.6 \pm 1.1
Triglycerides \pm SD, mmol/L	1.5 \pm 1.0
Alcohol consumption \pm SD, g/wk	42.4 \pm 97.0
Heart rate \pm SD, bpm	66.3 \pm 10.3
Corrected QT \pm SD, ms	416.4 \pm 19.7
QRS duration \pm SD, ms	92.3 \pm 12.7
Bundle branch block, n(%)	677(4.3)

Harrell's C, and Gönen and Heller's K concordance coefficients with 95% confidence intervals (CI) were calculated. Equality of the areas under the (receiver operating characteristic, ROC) curve (AUC) for studied ECG and VCG predictors of SCD was compared using a Wald test of the null hypothesis that all classifier AUC values are equal [22]. In addition, as we recently showed that the best predictive accuracy of ECG biomarkers is within 2 years after ECG recording [14], time-dependent (2-year) ROC AUC analysis was performed, using survival analysis framework [23,24].

3. Results

3.1. Study population

Clinical characteristics of the study population are shown in Table 1. Approximately half of the study participants were female, and 73% were white. Approximately one-third of the study population had a common cardiovascular risk factor (e.g. hypertension, smoking). Average traditional ECG parameters were normal. Over a median follow-up of 24.4 years, 581 SCDs occurred (incidence 1.77 (95%CI 1.63–1.92) per 1000 person-years).

3.2. Agreement in VCG origin point definition

Human observers agreed with automated detection of isoelectric origin point in 97.97% of time-coherent global beats, 97.65% of R-peak-aligned, and 96.84% of QRS-onset-aligned global XYZ beats. The intra-observer agreement was high (kappa 0.972), as well as inter-observer agreement (kappa 0.984).

3.3. Effect of beats alignment on ECG and VCG measurements

Marked changes were observed in the ECG and VCG morphology in different types of global beats (Fig. 6). The type of beat significantly affected all ECG and VCG measurements (Table 2). QRS duration and QT interval were larger in an R-aligned global median beat compared to a time-coherent beat. Direction of peak QRS, T, and SVG vectors was affected by different alignment to the largest degree, whereas differences in direction of area vectors were less prominent, albeit statistically significant. The narrowest peak QRS-T angle was observed on the time-coherent global median beat. Alignment by QRS onset produced peak QRS-T angle that was wider by 5°. Alignment by R-peaks leads to further widening of peak QRS-T angle by additional 5° (Table 2).

Different types of beats were characterized by significant differences in SCD predictive accuracy (Table 3 and Fig. 7). Predictive accuracy of spatial peak QRS-T angle measured on a time-coherent global median beat was significantly better than predictive accuracy of peak QRS-T angle measured on a global median beat that was constructed using QRS onset alignment. Overall, across all studied ECG and VCG predictors of SCD, time-coherent global XYZ median beat facilitated significantly higher predictive accuracy of SCD (Table 3 and Fig. 7).

3.4. Effect of VCG origin point location on ECG and VCG measurements

Noticeable differences were observed in the ECG and VCG morphology when comparing the two definitions of origin point (Fig. 6). The VCG origin point significantly affected all ECG and VCG measurements (Table 2). QRS duration, QTc, SAI QRST, and SVG were significantly larger if measured on a median beat constructed using QRS-onset definition of the heart vector origin. However, spatial peak QRS-T angle was significantly wider if measured on a median beat constructed using isoelectric origin point (Table 2).

The definition of the ECG origin point as an electrically-silent isoelectric segment of the cardiac cycle resulted in the most physiologically accurate measurement of GEH and traditional ECG parameters on a time-coherent global XYZ beat, which was highlighted by significantly higher ROC AUCs (Table 3 and Fig. 7).

4. Discussion

In this study, we defined the origin point of the heart vector as the electrically silent portion of the cardiac cycle (isoelectric period). We showed that the physiologically meaningful definition of the heart vector's origin point and time-coherent global XYZ median beat construction affected ECG measurements, and improved predictive accuracy of ECG and VCG biomarkers of SCD.

4.1. The origin of the heart vector

Our results highlight the importance of the electrophysiological definition of the electrical origin point of the heart vector (Fig. 6). We defined the heart vector origin as a point in 3D space which corresponds to the interval of time within the cardiac cycle where the heart vector does not move in 3D space. An appropriate selection of the heart vector origin point defines the ECG's baseline, which then affects all subsequent ECG and VCG measurements. Current ECG guidelines do not define the heart vector origin point [6,9]. Traditionally, the beginning of the QRS complex is defined as a “zero” value on a 12-lead ECG (Fig. 1Cii) [16]. However, the beginning of the QRS complex is *not* electrically silent: it is a time of atrial repolarization. Moreover, detection of QRS onset is challenging, resulting in disagreement between different algorithms [25]. A physiologically accurate definition of the heart vector origin point is crucial for accurate VCG and ECG measurements. Our algorithm searches for the flattest line within the cardiac cycle to provide a robust and reproducible solution, which is important for accurate detection of QRS, T, and SVG peak vectors. To facilitate its implementation, we provided our open-source Matlab software on GitHub.

A heart dipole model is a mathematical approximation with limitations. While a single dipole model is a good approximation of the healthy heart's electrical activity in healthy individuals, a single dipole model may be less accurate in diseased hearts [26]. Multipolar modeling can better characterize electrical activity of the heart, but it requires body surface mapping. Widely used 12-lead ECG technology is based on a single dipole model. In our work, we demonstrated that it is possible to localize the heart vector's origin point, and to construct a time-coherent, global median XYZ beat. Vector-Magnitude that is zeroed at the origin of the heart vector provides the most physiologically accurate cardiac beat. This is a fundamental step for any further

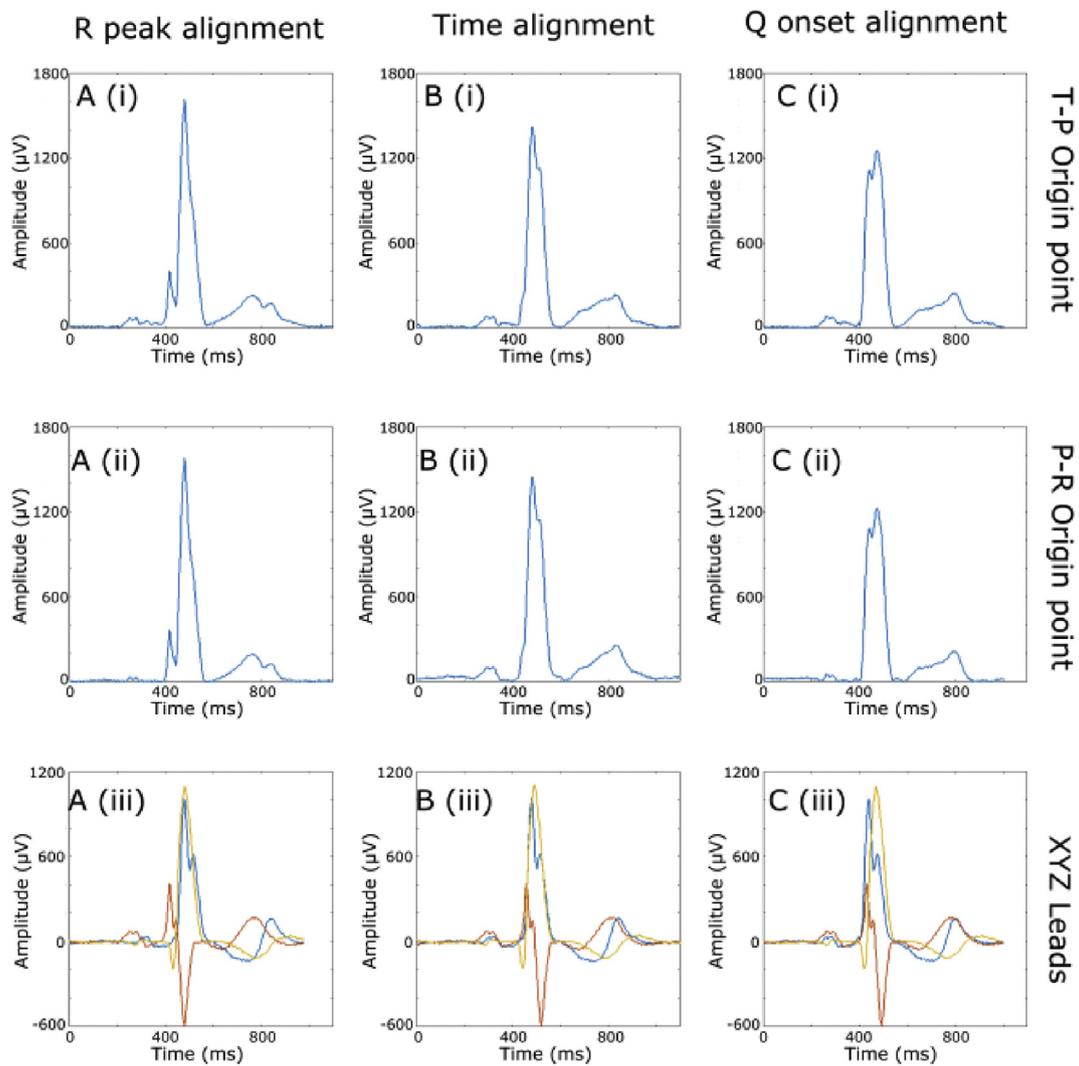


Fig. 6. XYZ leads and Vector-Magnitude obtained using three different alignment methods and two different electrical origin point location. (A) Signal alignment using the R-peak; (B) time-coherent global beat; (C) signal alignment using the beginning of the QRS complex. (i) Vector-Magnitude with origin point calculated in the TP interval; (ii) Vector-Magnitude with the origin point calculated in the PR interval; (iii) overlapped X, Y and Z leads.

global ECG and VCG measurements.

4.2. Appropriate construction of the global XYZ median beat

Modern ECG industry standards endorse the use of a representative median beat [6] for accurate ECG measurements. For the development

of a median beat, every beat included in the template beat must be properly aligned (Fig. 6). Several approaches to beat alignment are currently employed although some algorithm developers do not disclose details of their beat alignment process. A recent comparison of leading ECG algorithms revealed significant differences in ECG measurements by ECG manufacturers [25]. Different approaches to beat

Table 2
Effect of the global beat construction and origin point location on ECG and VCG measurements.

Measurement	Cross-corr (12SL)	OP(+) T-coherent	T-coherent	OP(+) R-aligned	R-aligned	OR(+) Q-aligned	Q-aligned	P-value
QRS duration(SD),ms	92.3(12.7)	83.7(15.9)	86.7(15.9)	88.1(16.7)	89.5(16.3)	83.9(17.0)	87.2(16.8)	< 0.0001
QT interval(SD), ms	400.1(31.3)	378.3(29.3)	379.5(29.8)	385.0(30.7)	384.3(31.6)	377.7(30.0)	379.1(30.0)	< 0.0001
Peak QRS-T angle(SD),	45.9(34.0)	42.6(31.7)	42.6(31.7)	55.7(32.1)	49.6(30.0)	50.6(34.8)	47.0(32.8)	< 0.0001
Area QRS-T angle(SD),	68.4(34.1)	57.2(30.5)	57.2(30.5)	69.5(33.8)	53.7(31.2)	68.1(33.4)	57.0(30.3)	< 0.0001
Peak SVG azimuth(SD),	4.3(25.3)	3.2(24.2)	3.2(24.2)	17.1(20.1)	17.6(20.1)	7.2(30.3)	5.8(28.7)	< 0.0001
Area SVG azimuth(SD),*	-2.6(24.9)	-6.4(21.5)	-6.4(21.5)	-3.2(23.3)	2.1(24.9)	-3.6(23.1)	-6.2(21.0)	< 0.0001
Peak SVG elevation(SD),	63.7(15.6)	62.6(15.3)	64.7(16.0)	64.7(16.0)	64.0(16.1)	66.9(19.2)	65.7(19.0)	< 0.0001
Area SVG elevation(SD),	64.3(16.3)	59.3(15.7)	64.1(15.8)	64.1(15.8)	60.3(17.0)	63.9(15.6)	60.0(16.8)	< 0.0001
Peak SVG magnitude(SD), mV	1.60(0.43)	1.67(0.45)	1.67(0.45)	1.69(0.43)	1.77(0.44)	1.54(0.44)	1.61(0.45)	< 0.0001
iVM _{QT} (SD), mV*ms	80.0(24.2)	86.5(24.9)	86.5(24.9)	79.0(24.2)	85.0(25.5)	80.8(24.2)	87.5(25.0)	< 0.0001
Wilson's SVG(SD), mV*ms	48.2(17.8)	61.3(21.5)	61.3(21.5)	47.9(18.1)	61.1(21.4)	48.4(18.1)	61.5(21.3)	< 0.0001
SAI QRST(SD), mV*ms	121.6(37.0)	132.7(38.8)	132.7(38.8)	121.4(36.6)	131.4(38.5)	121.2(36.8)	133.1(38.8)	< 0.0001

OP(+) = isoelectric origin point identified. T = time. SVG = Spatial Ventricular Gradient. SAI QRST = Sum Absolute QRST integral. iVMQT = QT integral on Vector Magnitude signal.

Table 3

Effect of the global beat construction and origin point location on predictive accuracy, measured by cumulative ROC AUC.

Measurement	Global beat type	Harrell's C	Gönen and Heller's K (95%CI)	P-value
QRS duration	OP(+) Time-Coherent	0.572	0.542 (0.531–0.552)	< 0.05
	Time-Coherent	0.573	0.537 (0.528–0.546)	
	OP(+) R-aligned	0.548	0.538 (0.523–0.551)	
	R-aligned	0.557	0.541 (0.529–0.553)	
	OP(+) Q-aligned	0.588	0.536 (0.528–0.543)	
QTc interval	Q-aligned	0.578	0.538 (0.528–0.547)	< 0.05
	OP(+) Time-Coherent	0.516	0.540 (0.518–0.561)	
	Time-Coherent	0.515	0.527 (0.504–0.549)	
	OP(+) R-aligned	NS	NS	
	R-aligned	NS	NS	
Peak QRS-T angle	OP(+) Q-aligned	0.520	0.535 (0.516–0.554)	< 0.05
	Q-aligned	0.509	0.525 (0.503–0.546)	
	OP(+) Time-Coherent	0.704	0.628 (0.617–0.639)	
	Time-Coherent	0.699	0.622 (0.611–0.632)	
	OP(+) R-aligned	0.700	0.656 (0.643–0.669)	
Area QRS-T angle	R-aligned	0.697	0.649 (0.637–0.661)	< 0.05
	OP(+) Q-aligned	0.678	0.627 (0.613–0.640)	
	Q-aligned	0.679	0.625 (0.612–0.637)	
	OP(+) Time-Coherent	0.697	0.660 (0.646–0.675)	
	Time-Coherent	0.690	0.651 (0.637–0.664)	
Peak SVG azimuth	OP(+) R-aligned	0.698	0.662 (0.647–0.676)	< 0.05
	R-aligned	0.690	0.650 (0.637–0.664)	
	OP(+) Q-aligned	0.701	0.662 (0.648–0.677)	
	Q-aligned	0.695	0.653 (0.640–0.666)	
	OP(+) Time-Coherent	0.584	0.567 (0.552–0.582)	
Area SVG azimuth	Time-Coherent	0.586	0.589 (0.554–0.584)	< 0.05
	OP(+) R-aligned	0.609	0.580 (0.561–0.598)	
	R-aligned	0.605	0.577 (0.557–0.596)	
	OP(+) Q-aligned	0.577	0.552 (0.535–0.569)	
	Q-aligned	0.577	0.551 (0.534–0.568)	
Peak SVG elevation(SD),°	OP(+) Time-Coherent	0.562	0.567 (0.550–0.584)	< 0.05
	Time-Coherent	0.564	0.567 (0.550–0.585)	
	OP(+) R-aligned	0.563	0.566 (0.547–0.585)	
	R-aligned	0.547	0.545 (0.522–0.567)	
	OP(+) Q-aligned	0.568	0.578 (0.560–0.595)	
Area SVG elevation(SD),°	Q-aligned	0.566	0.564 (0.546–0.581)	< 0.05
	OP(+) Time-Coherent	0.620	0.599 (0.582–0.616)	
	Time-Coherent	0.619	0.602 (0.585–0.618)	
	OP(+) R-aligned	0.620	0.596 (0.579–0.613)	
	R-aligned	0.621	0.596 (0.579–0.612)	
iVM _{QT}	OP(+) Q-aligned	0.604	0.598 (0.577–0.620)	< 0.05
	Q-aligned	0.611	0.603 (0.582–0.623)	
	OP(+) Time-Coherent	0.614	0.596 (0.579–0.612)	
	Time-Coherent	0.612	0.597 (0.580–0.614)	
	OP(+) R-aligned	0.600	0.590 (0.572–0.608)	
SAI QRST	R-aligned	0.604	0.584 (0.567–0.600)	< 0.05
	OP(+) Q-aligned	0.600	0.591 (0.573–0.609)	
	Q-aligned	0.601	0.587 (0.570–0.604)	
	OP(+) Time-Coherent	0.573	0.553 (0.540–0.566)	
	Time-Coherent	0.548	0.547 (0.531–0.563)	
SAI QRST	OP(+) R-aligned	0.563	0.549 (0.535–0.563)	< 0.05
	R-aligned	0.549	0.544 (0.528–0.560)	
	OP(+) Q-aligned	0.569	0.552 (0.538–0.565)	
	Q-aligned	0.550	0.546 (0.529–0.562)	
	OP(+) Time-Coherent	0.551	0.547 (0.531–0.562)	
SAI QRST	Time-Coherent	0.529	0.536 (0.517–0.555)	< 0.05
	OP(+) R-aligned	0.547	0.544 (0.527–0.560)	
	R-aligned	0.531	0.536 (0.517–0.554)	
	OP(+) Q-aligned	0.548	0.544 (0.529–0.560)	
	Q-aligned	0.530	0.536 (0.518–0.555)	

OP(+) = isoelectric origin point identified on TP segment. Other abbreviations as defined in the Table 2 legend.

alignment may be one of the reasons of discrepant results of ECG measurements. Results of our study showed that time-coherent beat construction is the best approach, which preserves the physiological sequence of cardiac activation and recovery. The time-coherent approach takes an easily detectable fiducial point as detected only on one ECG lead (e.g., X). Corresponding time-points on two other leads (e.g., Y and Z) are taken regardless of their morphological features. Such an approach significantly reduces the chance of error and preserves cardiac beat morphology. Our study shows that the time-coherent beat construction and a physiological definition of the heart vector origin

improved the dynamic predictive value of ECG and VCG measurements.

4.3. Novelty of the work

In this work we, for the first time, defined the origin point of the heart vector, and developed a novel algorithm to identify XYZ coordinates of the origin point in VCG. In addition, we developed a Matlab software application, and provided the open-source software at: <https://github.com/Tereshchenkolab/Origin>. For the first time, we compared different approaches of global beat construction. We

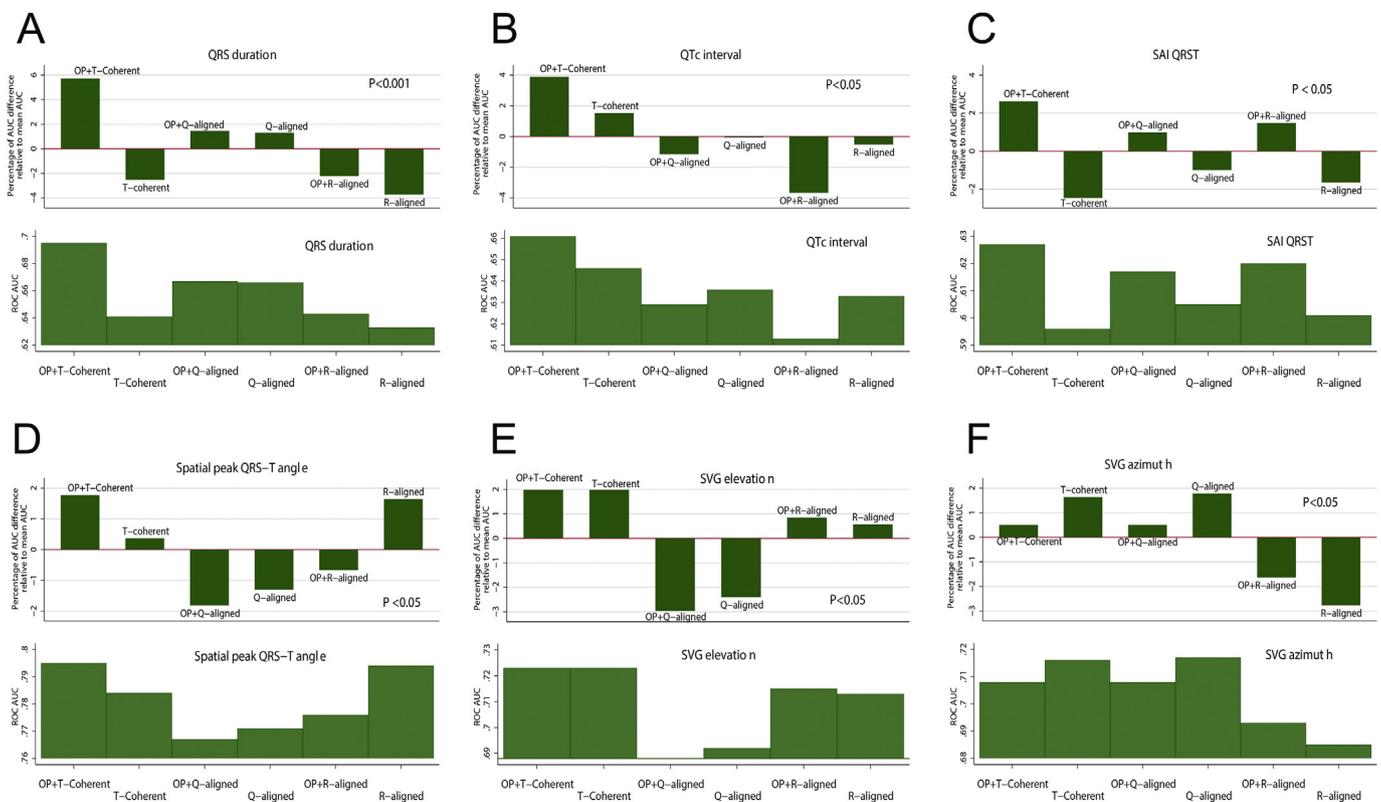


Fig. 7. Comparison of time-dependent AUCs for window of SCD prediction within 2 years, for (A) QRS duration, (B) QTc, (C) SAI QRST, (D) Spatial peak QRS-T angle, (E) peak SVG elevation, (F) peak SVG azimuth. An upper panel shows an AUC difference relative to mean ROC for each ECG/VCG parameter. A lower panel shows specific AUC.

proposed a time-coherent approach of global median XYZ (VCG) beat construction, zeroed in the origin point. “Recommendations for the Standardization and Interpretation of the Electrocardiogram” [6] suggested that “a comparative study is needed of global measurements made by different methods.” We used data of the a large prospective cohort study of more than 15,000 participants, and showed significant differences in global ECG measurements (QT interval, QRS duration) as a result of different approaches to global median beat construction.

4.4. Strengths and limitations

The strength of our study derives from the large prospective cohort design, and use of a clinically significant, well-adjudicated SCD outcome. However, as we analyzed resting 12-lead ECGs in a community, we did not encounter cardiac arrhythmias with heart rate above 139 bpm. Our developed algorithm is not designed to define an origin point if the heart rate exceeds 139 bpm. Additional studies are needed to define origin point in tachycardia.

As we focused on a global XYZ beat construction, we did not compare all possible single-lead beats alignment approaches. Specifically, we did not compare template matching or cross-correlation, which is an important step which can potentially affect the dominant beat selection [27]. Theoretically, if performed separately on each lead, cross-correlation can shift beats in time to a different degree in different leads, therefore distorting global temporal relationships and precluding formation of time-coherent global XYZ beat. The effect of cross-correlation on the formation of a time-coherent global XYZ beat deserves further study.

We used single-lead alignment by QRS peak. While we found a high degree of correlation between global time-coherent XYZ beats constructed using single X-lead alignment by QRS peak and maximum $|dV/dt|$ using 10-sec stationary resting ECGs (see Supplementary Material), use of QRS peak in the alignment might produce an error as the QRS

peak can be affected by respiration and noise [28]. Maximum $|dV/dt|$ is more stable fiducial point, and, therefore, single-lead alignment by maximum $|dV/dt|$ was implemented in the final version of our algorithm.

5. Conclusion

A physiologically meaningful definition of the heart vector's origin point and time-coherent global XYZ median beat preserved cardiac beat morphology and improved predictive accuracy of SCD biomarkers.

Funding sources

The Atherosclerosis Risk in Communities study has been funded in whole or in part with Federal funds from the National Heart, Lung, and Blood Institute, National Institutes of Health, Department of Health and Human Services, under Contract nos. (HHSN268201700001I, HHSN268201700002I, HHSN268201700003I, HHSN268201700004I, HHSN268201700005I). This work was supported by 1R01HL118277 and 2R56HL118277 (Principal Investigator Larisa Tereshchenko).

Disclosures

None.

Acknowledgements

The authors thank the staff and participants of the ARIC study for their important contributions. We would like to acknowledge the SCD mortality classification committee members: Nona Sotoodehnia (lead), Selcuk Adabag, Sunil Agarwal, Lin Chen, Rajat Deo, Leonard Ilkhanoff, Liviu Klein, Saman Nazarian, Ashleigh Owen, Kris Patton, and Larisa Tereshchenko.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.compbio.2018.11.013>.

References

- [1] E.J. Benjamin, S.S. Virani, C.W. Callaway, A.M. Chamberlain, A.R. Chang, S. Cheng, S.E. Chiuve, M. Cushman, F.N. Delling, R. Deo, Heart disease and stroke statistics—2018 update: a report from the American Heart Association, *Circulation* 137 (2018) e67–e492.
- [2] R. Deo, F.L. Norby, R. Katz, N. Sotoodehnia, S. Adabag, C.R. DeFilippi, B. Kestenbaum, L.Y. Chen, S.R. Heckbert, A.R. Folsom, R.A. Kronmal, S. Konety, K.K. Patton, D. Siscovick, M.G. Shlipak, A. Alonso, Development and validation of a sudden cardiac death prediction model for the general population, *Circulation* 134 (2016) 806–816 <https://doi.org/10.1161/CIRCULATIONAHA.116.023042>.
- [3] A.L. Aro, K. Reinier, C. Rusinaru, A. Uy-Evanado, N. Darouian, D. Phan, W.J. Mack, J. Jui, E.Z. Soliman, L.G. Tereshchenko, S.S. Chugh, Electrical risk score beyond the left ventricular ejection fraction: prediction of sudden cardiac death in the Oregon Sudden Unexpected Death Study and the Atherosclerosis Risk in Communities Study, *Eur. Heart J.* 38 (2017) 3017–3025 <https://doi.org/10.1093/eurheartj/ehx331>.
- [4] J.W. Waks, C.M. Sitlani, E.Z. Soliman, M. Kabir, E. Ghafoori, M.L. Biggs, C.A. Henrikson, N. Sotoodehnia, T. Biering-Sorensen, S.K. Agarwal, D.S. Siscovick, W.S. Post, S.D. Solomon, A.E. Buxton, M.E. Josephson, L.G. Tereshchenko, Global electric heterogeneity risk score for prediction of sudden cardiac death in the general population: the atherosclerosis risk in Communities (ARIC) and cardiovascular health (CHS) studies, *Circulation* 133 (2016) 2222–2234 <https://doi.org/10.1161/CIRCULATIONAHA.116.021306>.
- [5] J.W. Waks, L.G. Tereshchenko, Global electrical heterogeneity: a review of the spatial ventricular gradient, *J. Electrocardiol.* 49 (2016) 824–830 <https://doi.org/10.1016/j.jelectrocard.2016.07.025>.
- [6] P. Kligfield, L.S. Gettes, J.J. Bailey, R. Childers, B.J. Deal, E.W. Hancock, H.G. van, J.A. Kors, P. Macfarlane, D.M. Mirvis, O. Pahlm, P. Rautaharju, G.S. Wagner, M. Josephson, J.W. Mason, P. Okin, B. Surawicz, H. Wellens, Recommendations for the standardization and interpretation of the electrocardiogram: part I: the electrocardiogram and its technology a scientific statement from the American heart Association electrocardiography and arrhythmias committee, Council on clinical Cardiology; the American College of Cardiology Foundation; and the heart Rhythm Society endorsed by the International Society for Computerized Electrocardiology, *J. Am. Coll. Cardiol.* 49 (2007) 1109–1127 [pii];10.1016/j.jacc.2007.01.024 [doi] [https://doi.org/S0735-1097\(07\)00232-X](https://doi.org/S0735-1097(07)00232-X).
- [7] E. Frank, An accurate, clinically practical system for spatial vectorcardiography, *Circulation* 13 (1956) 737–749.
- [8] F.N. Wilson, F.D. Johnston, The vectorcardiogram, *Am. Heart J.* 16 (1938) 14–28 [https://doi.org/10.1016/S0002-8703\(38\)90899-3](https://doi.org/10.1016/S0002-8703(38)90899-3).
- [9] J.J. Bailey, A.S. Berson, A. Garson Jr., L.G. Horan, P.W. MacFarlane, D.W. Mortara, C. Zywiets, Recommendations for standardization and specifications in automated electrocardiography: bandwidth and digital signal processing. A report for health professionals by an ad hoc writing group of the Committee on Electrocardiography and Cardiac Electrophysiology of the Council on Clinical Cardiology, American Heart Association, *Circulation* 81 (1990) 730–739.
- [10] ARIC, Study Description, Available at <http://www2.csc.unc.edu/aric/distribution-agreements>, (1987) <http://www2.csc.unc.edu/aric/distribution-agreements> June 22nd, 2018.
- [11] NIH/NHLBI, Biologic Specimen and data repository Information coordinating center, Available at: <https://biolincc.nhlbi.nih.gov/>, (1987) , Accessed date: 22 June 2018.
- [12] The ARIC Investigators, The atherosclerosis risk in community (ARIC) study: design and Objectives, *Am. J. Epidemiol.* 129 (1989) 687–702.
- [13] W.T. O'Neal, M.J. Singleton, J.D. Roberts, L.G. Tereshchenko, N. Sotoodehnia, L.Y. Chen, G.M. Marcus, E.Z. Soliman, Association between QT-interval Components and sudden cardiac death: the ARIC study (atherosclerosis risk in Communities), *Circ Arrhythm Electrophysiol* 10 (2017), <https://doi.org/10.1161/CIRCEP.117.005485>.
- [14] A. Perez Alday Erick, D. German, A. Bender, A. Li-Pershing, C. Hamilton, A. Thomas Jason, R. Gonzales, A. Li, K. Newton, L. Tereshchenko, Abstract 12033: dynamic predictive accuracy of electrocardiographic biomarkers of sudden cardiac death within a survival framework: the atherosclerosis risk in Communities (ARIC) study, *Circulation* 138 (2018) A12033-A12033 https://doi.org/10.1161/circ.138.suppl_1.12033.
- [15] J.A. Kors, H.G. van, A.C. Sittig, J.H. van Bommel, Reconstruction of the Frank vectorcardiogram from standard electrocardiographic leads: diagnostic comparison of different methods, *Eur. Heart J.* 11 (1990) 1083–1092.
- [16] G. Healthcare, Marquette 12SL ECG Analysis Program, Physician's Guide, Revision B, 2015.
- [17] J. Pan, W. Tompkins, A real-time QRS detection algorithm, *IEEE Trans. Biomed. Eng.* 32 (1985) 230–236.
- [18] W. Zong, M. Saeed, T. Heldt, A QT interval detection algorithm based on ECG curve length transform, *Comput. Cardiol.* 33 (2006) 377–380.
- [19] S. Sur, L. Han, L.G. Tereshchenko, Comparison of sum absolute QRST integral, and temporal variability in depolarization and repolarization, measured by dynamic vectorcardiography approach, in healthy men and women, *PLoS One* 8 (2013) e57175 <https://doi.org/10.1371/journal.pone.0057175>.
- [20] L.G. Tereshchenko, A. Cheng, B.J. Fetis, B. Butcher, J.E. Marine, D.D. Spragg, S. Sinha, D. Dalal, H. Calkins, G.F. Tomaselli, R.D. Berger, A new electrocardiogram marker to identify patients at low risk for ventricular tachyarrhythmias: sum magnitude of the absolute QRST integral, *J. Electrocardiol.* 44 (2011) 208–216 [https://doi.org/S0022-0736\(10\)00381-X](https://doi.org/S0022-0736(10)00381-X) [pii];10.1016/j.jelectrocard.2010.08.012 [doi].
- [21] J.A. Thomas, A.P. A.E., A. Junell, K. Newton, C. Hamilton, Y. Li-Pershing, D. German, A. Bender, L.G. Tereshchenko, Vectorcardiogram in athletes: the Sun valley Ski study, *Ann. Noninvasive Electrocardiol.* (2018) e12614 <https://doi.org/10.1111/anec.12614>.
- [22] E.R. DeLong, D.M. DeLong, D.L. Clarke-Pearson, Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach, *Biometrics* 44 (1988) 837–845.
- [23] P.J. Heagerty, T. Lumley, M.S. Pepe, Time-dependent ROC curves for censored survival data and a diagnostic marker, *Biometrics* 56 (2000) 337–344.
- [24] M. Cattaneo, P. Malighetti, D. Spinelli, Estimating receiver operative characteristic curves for time-dependent outcomes: the stroccurve package, *STATA J.* 17 (2017) 1015–1023.
- [25] P. Kligfield, F. Badilini, I. Denjoy, S. Babaeizadeh, E. Clark, J. De Bie, B. Devine, F. Extramiana, G. Generali, R. Gregg, E. Helfenbein, J. Kors, R. Leber, P. Macfarlane, P. Maison-Blanche, I. Rowlandson, R. Schmid, M. Vaglio, G. van Herpen, J. Xue, B. Young, C.L. Green, Comparison of automated interval measurements by widely used algorithms in digital electrocardiographs, *Am. Heart J.* 200 (2018) 1–10 <https://doi.org/https://doi.org/10.1016/j.ahj.2018.02.014>.
- [26] Y. Okamoto, Y. Teramachi, T. Musha, H. Tsunakawa, K. Harumi, Moving multiple dipole model for cardiac activity, *Jpn. Heart J.* 23 (1982) 293–304.
- [27] J.H. van Bommel, J.A. Kors, G. van Herpen, Methodology of the modular ECG analysis system MEANS, *Methods Inf. Med.* 29 (1990) 346–353.
- [28] J.A. Kors, J.L. Talmon, J.H. van Bommel, Multilead ECG analysis, *Computers and biomedical research, an international journal* 19 (1986) 28–46.