



# Distinct ECG Phenotypes Identified in Hypertrophic Cardiomyopathy Using Machine Learning Associate With Arrhythmic Risk Markers

Aurore Lyon<sup>1†</sup>, Rina Ariga<sup>2†</sup>, Ana Mincholé<sup>1</sup>, Masliza Mahmud<sup>2</sup>, Elizabeth Ormondroyd<sup>2</sup>, Pablo Laguna<sup>3</sup>, Nando de Freitas<sup>1</sup>, Stefan Neubauer<sup>2</sup>, Hugh Watkins<sup>2</sup> and Blanca Rodriguez<sup>1\*</sup>

## OPEN ACCESS

### Edited by:

Ahsan H. Khandoker,  
Khalifa University, United Arab  
Emirates

### Reviewed by:

Eric A. Sobie,  
Icahn School of Medicine at Mount  
Sinai, United States  
Mohammad Hasan Imam,  
American International  
University-Bangladesh, Bangladesh

### \*Correspondence:

Blanca Rodriguez  
blanca.rodriguez@cs.ox.ac.uk

<sup>†</sup>These authors have contributed  
equally to this work and first authors.

### Specialty section:

This article was submitted to  
Computational Physiology and  
Medicine,  
a section of the journal  
Frontiers in Physiology

Received: 14 November 2017

Accepted: 26 February 2018

Published: 13 March 2018

### Citation:

Lyon A, Ariga R, Mincholé A,  
Mahmod M, Ormondroyd E, Laguna P,  
de Freitas N, Neubauer S, Watkins H  
and Rodriguez B (2018) Distinct ECG  
Phenotypes Identified in Hypertrophic  
Cardiomyopathy Using Machine  
Learning Associate With Arrhythmic  
Risk Markers. *Front. Physiol.* 9:213.  
doi: 10.3389/fphys.2018.00213

<sup>1</sup> Department of Computer Science, University of Oxford, Oxford, United Kingdom, <sup>2</sup> Division of Cardiovascular Medicine, Radcliffe Department of Medicine, University of Oxford, Oxford, United Kingdom, <sup>3</sup> Biomedical Signal Interpretation & Computational Simulation Group, CIBER-BBN, University of Zaragoza, Zaragoza, Spain

**Aims:** Ventricular arrhythmia triggers sudden cardiac death (SCD) in hypertrophic cardiomyopathy (HCM), yet electrophysiological biomarkers are not used for risk stratification. Our aim was to identify distinct HCM phenotypes based on ECG computational analysis, and characterize differences in clinical risk factors and anatomical differences using cardiac magnetic resonance (CMR) imaging.

**Methods:** High-fidelity 12-lead Holter ECGs from 85 HCM patients and 38 healthy volunteers were analyzed using mathematical modeling and computational clustering to identify phenotypic subgroups. Clinical features and the extent and distribution of hypertrophy assessed by CMR were evaluated in the subgroups.

**Results:** QRS morphology alone was crucial to identify three HCM phenotypes with very distinct QRS patterns. Group 1 ( $n = 44$ ) showed normal QRS morphology, Group 2 ( $n = 19$ ) showed short R and deep S waves in V4, and Group 3 ( $n = 22$ ) exhibited short R and long S waves in V4-6, and left QRS axis deviation. However, no differences in arrhythmic risk or distribution of hypertrophy were observed between these groups. Including T wave biomarkers in the clustering, four HCM phenotypes were identified: Group 1A ( $n = 20$ ), with primary repolarization abnormalities showing normal QRS yet inverted T waves, Group 1B ( $n = 24$ ), with normal QRS morphology and upright T waves, and Group 2 and Group 3 remaining as before, with upright T waves. Group 1A patients, with normal QRS and inverted T wave, showed increased HCM Risk-SCD scores (1A: 4.0%, 1B: 1.8%, 2: 2.1%, 3: 2.5%,  $p = 0.0001$ ), and a predominance of coexisting septal and apical hypertrophy ( $p < 0.0001$ ). HCM patients in Groups 2 and 3 exhibited predominantly septal hypertrophy (85 and 90%, respectively).

**Conclusion:** HCM patients were classified in four subgroups with distinct ECG features. Patients with primary T wave inversion not secondary to QRS abnormalities had increased HCM Risk-SCD scores and coexisting septal and apical hypertrophy,

suggesting that primary T wave inversion may increase SCD risk in HCM, rather than T wave inversion secondary to depolarization abnormalities. Computational ECG phenotyping provides insight into the underlying processes captured by the ECG and has the potential to be a novel and independent factor for risk stratification.

**Keywords:** hypertrophic cardiomyopathy, electrocardiography, e-cardiology, phenotyping, computational clustering

## INTRODUCTION

Hypertrophic cardiomyopathy (HCM) remains a common yet challenging genetic heart muscle disease due to its heterogeneous clinical course. Ventricular arrhythmias are a major cause of sudden cardiac death (SCD) in young people (Maron et al., 2000; Maron, 2002). Accurate identification of high risk patients is a clinical priority since implantable cardioverter-defibrillators (ICD) can successfully treat ventricular arrhythmias triggering SCD.

In HCM, both ionic remodeling (Passini et al., 2016) and structural abnormalities [hypertrophy (Spirito et al., 2000), myocyte disarray (Varnava et al., 2001b) and fibrosis (Adabag et al., 2008)] create a pro-arrhythmic substrate to different extents in specific patients. Previous studies have attempted to assess the electrophysiological signature of HCM by visually inspecting the standard 12-lead paper electrocardiogram (ECG). Abnormalities such as abnormal Q waves, wide and high amplitude QRS complexes, ST segment displacement as well as giant inverted T waves have been reported in HCM (Savage et al., 1978; Lakdawala et al., 2011). However, no single abnormality was shown to be characteristic of HCM patients (Maron et al., 1983) and it is unclear whether T wave inversion is secondary to abnormal depolarization or a consequence of abnormal repolarization dynamics and heterogeneity. Furthermore, previous studies including cohorts of high-risk HCM patients also failed to produce reliable stratification, finding no differences in ECG between patients with and without appropriate ICD shocks (Maron et al., 1982; Sherrid et al., 2009). Some studies for example reported TWI to be related to increase in SCD risk (Ostman-Smith et al., 2010) but others did not (Maron et al., 1982; Sherrid et al., 2009). This may be due to the limitations in the methodology and therefore, more sophisticated approaches and new knowledge are required to improve the information extracted from the ECG for HCM phenotyping.

In the absence of reliable ECG biomarkers, five conventional risk factors [non-sustained ventricular tachycardia (NSVT), unexplained syncope, family history of SCD, massive left ventricular hypertrophy (LVH) and abnormal exercise blood

pressure response] provide clinical utility for predicting SCD in HCM (Maron et al., 2007; Christiaans et al., 2010). More recently, the prospectively validated HCM Risk-SCD prediction model (O'Mahony et al., 2014) recommended by the 2014 ESC guidelines (Elliott et al., 2014), has performed better than conventional risk factors albeit with limitations (Maron et al., 2015). However, neither method captures the degree of the underlying myocardial abnormalities which lead to arrhythmic risk.

Our aim was to identify discrete subgroups of HCM patients with differences in electrophysiological and structural phenotype using novel computational analysis of high fidelity 12-lead Holter ECGs, through combined machine learning and mathematical modeling. To this end, we first extracted morphology-based biomarkers from the QRS using a mathematical model based on Hermite functions. We then applied an unsupervised clustering approach to ECG-based biomarkers, to automatically identify the presence of different phenotypic subgroups in the HCM population. Cardiac magnetic resonance (CMR) imaging and arrhythmic risk markers were evaluated to further characterize the patient subgroups. The low incidence of SCD in HCM of <1% per year (O'Mahony et al., 2013), and our reliance on patients without comorbidities, precluded its use as an endpoint in this study. Instead, we provide a deeper characterization of HCM phenotypes capitalizing on recently acquired digital ECG in conjunction with CMR data which is unattainable in larger retrospective databases. We hypothesized that detailed quantification of QRS morphology and T wave abnormalities using high fidelity ECGs would identify important features of the underlying electrophysiological and anatomical substrate, to enable improved phenotypic characterization of HCM patients.

## MATERIALS AND METHODS

### Ethics and Study Population

This prospective study was approved by the National Research Ethics Committee (REC ref 12/LO/1979) and informed written consent was obtained from each participant. Eighty-five patients with HCM were recruited from the University of Oxford Inherited Cardiac Conditions clinic, John Radcliffe Hospital, Oxford, UK. HCM diagnosis was made by presence of a pathogenic mutation in a known sarcomeric gene or, in the absence of an identified mutation, HCM was defined as LVH ( $\geq 15$  mm) not originating from another cause. Gene positive patients without hypertrophy (G+LVH-) were included in the study as a number of SCDs have been reported in this patient cohort (Varnava et al., 2001a; Pasquale et al., 2012) and the consideration of LVH alone has limitations (Sen-Chowdhry et al.,

**Abbreviations:** AHA, American Heart Association; BP, blood pressure; BPR, blood pressure response; CMR, cardiovascular magnetic resonance; ECG, electrocardiogram; G+LVH-, genotype positive HCM with normal left ventricular wall thickness; HCM, hypertrophic cardiomyopathy; ICD, implantable cardioverter-defibrillator; LAD, left axis deviation; LV, left ventricle; LVH, left ventricular hypertrophy; LVOT, left ventricular outflow tract; MYBPC3, myosin-binding protein C; MYH7, beta-myosin heavy chain; NSVT, non-sustained ventricular tachycardia; NYHA, New York Heart Association; SCD, sudden cardiac death; TWI, T wave inversion.

2016); of note, five of the G+LVH- subjects had abnormal ECGs with voltage criteria for LVH. Thirty-eight age- and gender-matched healthy volunteers were non-smokers without cardiovascular disease, hypertension, diabetes, or family history of cardiomyopathy or SCD.

The methodology is summarized in **Figure 1** and detailed methods can be found in Supplementary Material 1.

## Electrocardiographic Measurements

Participants underwent 12-lead ambulatory Holter monitoring (sampling frequency 1,000 Hz, H12+, Mortara Instrument, Milwaukee, WI, USA) for 24 h and a standard 12-lead resting digital electrocardiograph (Burdick 8500, Glasgow, UK) (Macfarlane et al., 1990). Standard measurements of ECG axes, amplitudes and intervals were obtained for both global (reported in **Tables 1, 3, 4**) and single-leads from the Burdick ECG XML files.

## Cardiovascular Magnetic Resonance (CMR) Imaging

CMR imaging was performed at 3 tesla (TIM Trio, Siemens) in all participants, except in 10 patients with ICD at the time of enrollment (CMR prior to ICD insertion performed for clinical care was evaluated in these patients). Analysis was performed using cmr42© (Circle Cardiovascular Imaging, Calgary, Canada). The extent, and morphology, of hypertrophy was identified on short axis images and categorized into 4 subtypes: no hypertrophy—genotype positive patients with wall thickness  $\leq 12$  mm (G+LVH-); septal hypertrophy—basal and/or mid septal wall thickness  $> 12$  mm in genotype positive patients or  $\geq 15$  mm in genotype negative patients; apical hypertrophy—apical wall thickness  $\geq 15$  mm below papillary muscle level; mixed hypertrophy—coexisting hypertrophy in septal and apical segments.

## Clinical Data Collection

Genetic results and the conventional risk factors (Elliott et al., 2000) were obtained as part of the patient's routine clinical care. The five conventional risk factors were defined as: NSVT (three or more consecutive ventricular beats at a rate of  $\geq 120$  bpm lasting  $< 30$  s on clinical 3-lead 24- to 48-h Holter monitoring), unexplained syncope ( $\geq 1$  episode of unexplained syncope), family history of SCD (history of SCD in  $\geq 1$  first degree relative  $\leq 40$  years old or SCD in a first degree relative with confirmed HCM at any age), massive LVH (LV wall thickness in any myocardial segment of  $\geq 30$  mm on short axis CMR images) and abnormal exercise blood pressure (BP) response (rise in systolic BP  $< 20$  mmHg or a fall of  $> 10$  mmHg from baseline to peak exercise in patients  $\leq 40$  years old). HCM Risk-SCD score (2014 ESC guidelines) was calculated using 7 disease variables as in O'Mahony et al. (2014) (Supplementary Material 1.4). A  $\geq 6\%$  risk of SCD at 5 years is classified as high risk and ICD implantation is recommended, 4–6% is intermediate risk and ICD may be considered, and  $< 4\%$  is low risk.

## Holter ECG Pre-processing

The first 30-min ECG excerpt was used to analyze the 8 linearly independent leads (I, II, V1-6) with custom-built software using MATLAB (Mathworks, MA, USA). A wavelet based delineator (Martínez et al., 2004) extracted the peaks and boundaries of the ECG waveforms. High frequency noise was removed using a low-pass Butterworth filter with cut-off frequency of 45 Hz, baseline drift was removed by a cubic spline method and a notch filter rejected the 50 Hz mains power artifact. The first twenty beats with maximal ST segment-T wave signal-to-noise ratio were then considered for analysis and were aligned with respect to the QRS complex by Woody's method (Sörnmo and Laguna, 2005) (Supplementary Material 1.5). A sensitivity analysis showed that analyzing different 30-min excerpts in the recording yielded the same results. Thirty minutes also provided enough data to screen the excerpts and avoid changes in beat morphology due to changes in heart rate. Average QRS and STT waveforms were then computed from these 20 beats.

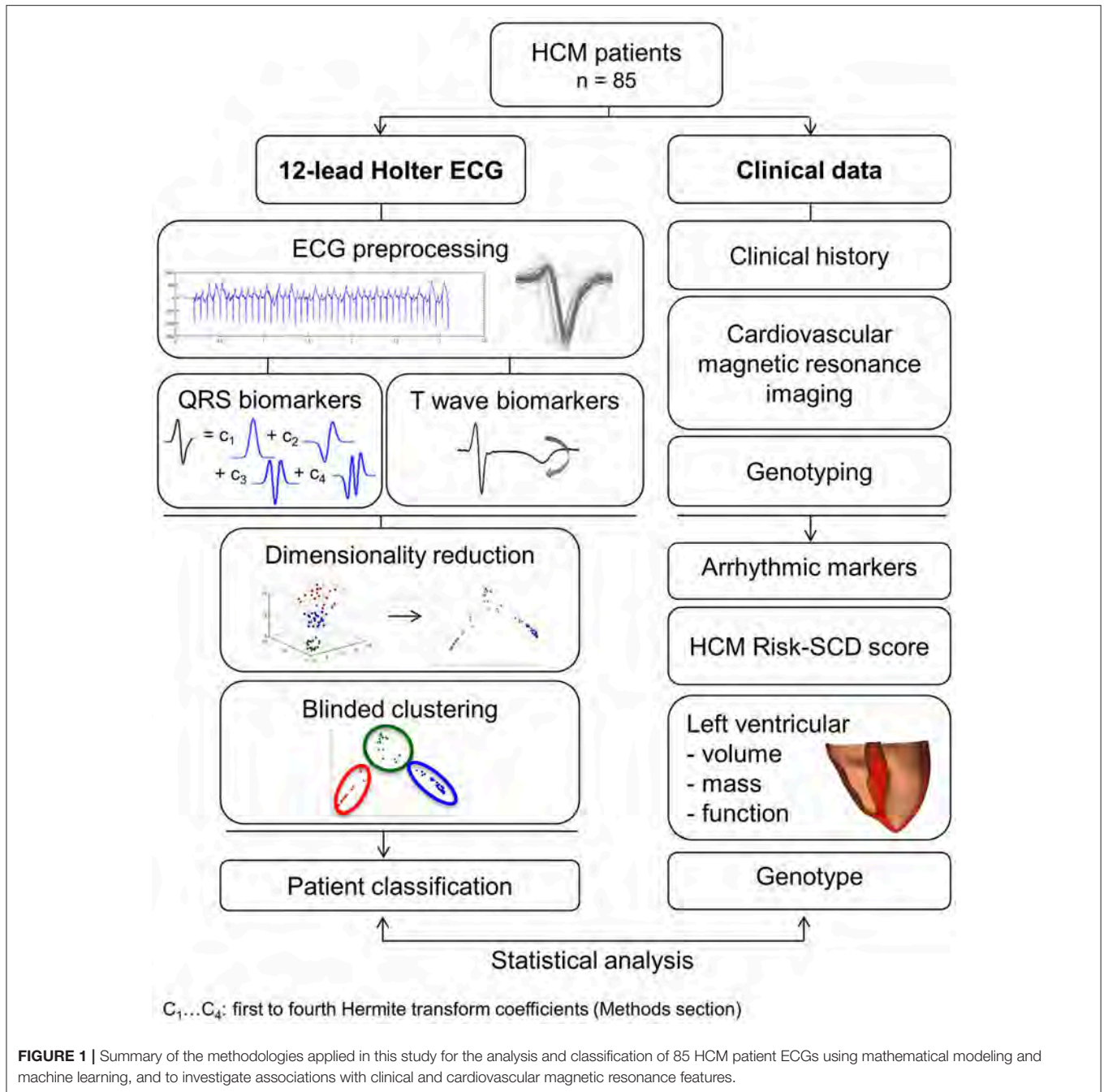
## Extraction of QRS and T Wave Biomarkers

All biomarkers were calculated per lead. QRS biomarkers are listed and illustrated in Supplementary Material 1.6, Figure S1. In addition to standard QRS biomarkers computed from signal processing, the QRS shape (morphology) was quantified by mathematically modeling the QRS waveform using a combination of Hermite functions, with well-established ability to provide a compact representation of the QRS complex (Laguna et al., 1996) (Supplementary Material 1.7, Figure S2). Indeed, three Hermite functions (Supplementary Material 1.7) enable to recover 98% of the QRS energy in control subjects (Sörnmo et al., 1981). However, four Hermite functions were needed in HCM due to greater QRS heterogeneity. Any particular QRS morphology such as an RSR' pattern can thus be generated as the sum of these scaled shapes.

## Identification of Subgroups in HCM Using QRS and T Wave Biomarkers

Each patient was assigned a vector of morphological QRS and T wave biomarkers. Seven significant features were selected using the Multi-cluster feature selection method described in Cai et al. (2010), which is an unsupervised feature selection algorithm to reduce the number of variables under consideration. In short, the method consists in a combination of spectral analysis of the data with a L1-regularized least squares optimization method. This method was chosen for its ability to preserve the multi-cluster structure of the dataset. The 7 features were then reduced to two dimensions for visualization purposes, using Laplacian eigenmaps dimensionality reduction, as described in Belkin and Niyogi (2003). This method preserves the local geometrical properties of the dataset by computing the eigenvalues and eigenvectors of the graph Laplacian generalized eigenvector problem (Belkin and Niyogi, 2003).

A density-based clustering algorithm, DBSCAN (Ester et al., 1996), was then applied on this low-dimensional representation of the dataset to extract subgroups. This algorithm identified clusters by maximizing the density in each of the clusters. The minimum number of individuals in a cluster was set to  $n/25 = 3$ ,



with  $n = 86$  HCM patients. The distance between neighboring individuals was evaluated using the Euclidean distance. The same results (i.e., same patients' subgroups) were obtained using a different clustering algorithm (k-means). Clustering analysis based only on QRS morphology was performed, and then repeated with the addition of T wave biomarkers. Clustering was performed by AL who was blinded to clinical data. To assess the effect of G+LVH- patients on the results, a further clustering analysis was performed excluding the G+LVH- patients.

### Statistical Analysis

Data are expressed as mean  $\pm$  standard deviation or median and range. Normally-distributed data were compared using *t*-tests or analysis of variance. Non-normally distributed data were compared using the Mann-Whitney *U*-test or Kruskal-Wallis test. Categorical data were compared with Chi-square or Fisher's exact tests. Statistical significance was assumed when  $p < 0.05$  (after Bonferroni adjustment for multiple comparisons, where appropriate). Statistical analysis was performed with IBM SPSS Statistics, version 20.0 (IBM Corp, Armonk, NY, USA).

**TABLE 1** | Characteristics of healthy volunteers and HCM patients.

	Healthy volunteers (n = 38)	HCM patients (n = 85)	p-value
Age, years	47 ± 15	45 ± 14	0.54
Male	25 (66)	58 (68)	0.84
Body mass index, kg/m <sup>2</sup>	24 ± 4	26 ± 4	0.06
Systolic BP, mmHg	117 ± 14	118 ± 14	0.76
Diastolic BP, mmHg	70 ± 10	72 ± 12	0.33
<b>CMR DIMENSIONS AND FUNCTION</b>			
LV end-diastolic volume, ml	155 ± 35	154 ± 33	0.88
LV end-systolic volume, ml	50 ± 17	41 ± 16	<b>0.007</b>
LV ejection fraction, %	68 ± 5	74 ± 7	<b>0.00003</b>
LV mass index, g/m <sup>2</sup>	54 ± 11	72 ± 25	<b>0.00002</b>
Maximal LV wall thickness, mm	11 ± 1	20 ± 6	<b>2 × 10<sup>-14</sup></b>
Left atrial diameter, mm	33 ± 7	39 ± 7	<b>0.00004</b>
<b>ECG FEATURES</b>			
Heart rate, bpm	58 ± 11	58 ± 10	0.81
QRS axis, °	40 ± 34	13 ± 43	<b>0.0007</b>
QRS duration, ms	91 ± 9	98 ± 15	<b>0.02</b>
R wave amplitude	683 ± 200	764 ± 341	<b>0.46</b>
QRS amplitude, mV	1441 ± 414	1825 ± 652	<b>0.002</b>
QRS ascending slope	74 ± 22	89 ± 32	<b>0.025</b>
QRS descending slope	-118 ± 51	-150 ± 57	<b>0.001</b>
Pathological Q wave	0	20 (23)	<b>0.0004</b>
T wave axis, °	34 ± 17	69 ± 54	<b>0.0001</b>
Abnormal T wave axis	0	28 (35)	<b>9 × 10<sup>-6</sup></b>
T wave amplitude, mV	356 ± 141	170 ± 257	<b>0.0001</b>
T wave inversion	0	26 (30)	<b>0.00002</b>
Giant T wave inversion	0	6 (7)	0.18
T peak to T end interval, ms	81 ± 14	85 ± 18	0.16
ST segment displacement, mV	42 ± 45	26 ± 49	0.14
QTc interval, ms	411 ± 17	440 ± 27	<b>8 × 10<sup>-8</sup></b>
JTc interval, ms	300 ± 43	354 ± 99	<b>0.00003</b>

Mean ± standard deviation, or number of participants (%). BP, blood pressure; CMR, cardiovascular magnetic resonance; LV, left ventricular. Bold values mean p-value significant ( $p < 0.05$ ).

## RESULTS

### Study Population Characteristics

The study population characteristics are summarized in **Table 1**. HCM patients ( $n = 85$ ) had a more leftward QRS axis, larger QRS duration and amplitude, steeper QRS slopes and abnormal Q waves compared with healthy volunteers ( $n = 38$ ) (all  $p < 0.03$ ). They also showed lower T wave amplitude, T wave inversion (TWI), abnormal T wave axis and prolonged QTc (all  $p < 0.001$ ). **Table 2** describes the clinical characteristics for the HCM patients. Patients were mainly asymptomatic (median NYHA functional class = 1) and were low risk for SCD (median HCM Risk-SCD score = 2.5%; median total risk factor = 1). Nineteen patients had an ICD implanted for primary prevention with a median follow-up of 3 years. Only 1 patient (5%) received an appropriate shock, which is in keeping with primary prevention discharge rates seen in a previous study (Maron et al., 2000).

**TABLE 2** | Clinical and genotype characteristics of HCM patients.

NYHA class, [median (range)]	1 (1–3)
NYHA III/IV	5 (6)
LVOT obstruction (gradient $\geq 30$ mmHg)	11 (13)
Implantable cardioverter-defibrillator	19 (22)
Appropriate ICD shocks	1
ICD follow-up, years [median (range)]	3 (0–12)
<b>HYPERTROPHY MORPHOLOGY</b>	
No LVH (G+LVH-)	9 (11)
Septal LVH	58 (68)
Apical LVH	4 (5)
Mixed septal & apical LVH	14 (16)
HCM Risk-SCD score, % [median (range)]	2.5 (0.8–11.0)
<b>CONVENTIONAL RISK FACTORS</b>	
NSVT	23 (27)
Syncope	10 (12)
Family History SCD	17 (20)
Abnormal exercise BP response	7 (8)
Massive LVH $\geq 30$ mm	1 (1)
Total number of risk factors [median (range)]	1 (0–3)
0 Risk factors	34 (40)
1 Risk factor	45 (53)
$\geq 2$ Risk factors	6 (7)
<b>GENOTYPE</b>	
Gene negative	27 (32)
MYBPC3	33 (39)
MYH7	24 (28)
Troponin I	1 (1)

Number of patients (%). NYHA, New York Heart Association; LVOT, left ventricular outflow tract; ICD, implantable cardioverter-defibrillator; LVH, left ventricular hypertrophy; G+LVH-, genotype positive HCM with normal wall thickness; SCD, sudden cardiac death; NSVT, non-sustained ventricular tachycardia; BP, blood pressure; ESC, European Society of Cardiology; HCM, hypertrophic cardiomyopathy; MYBPC3, myosin-binding protein C; MYH7, beta-myosin heavy chain.

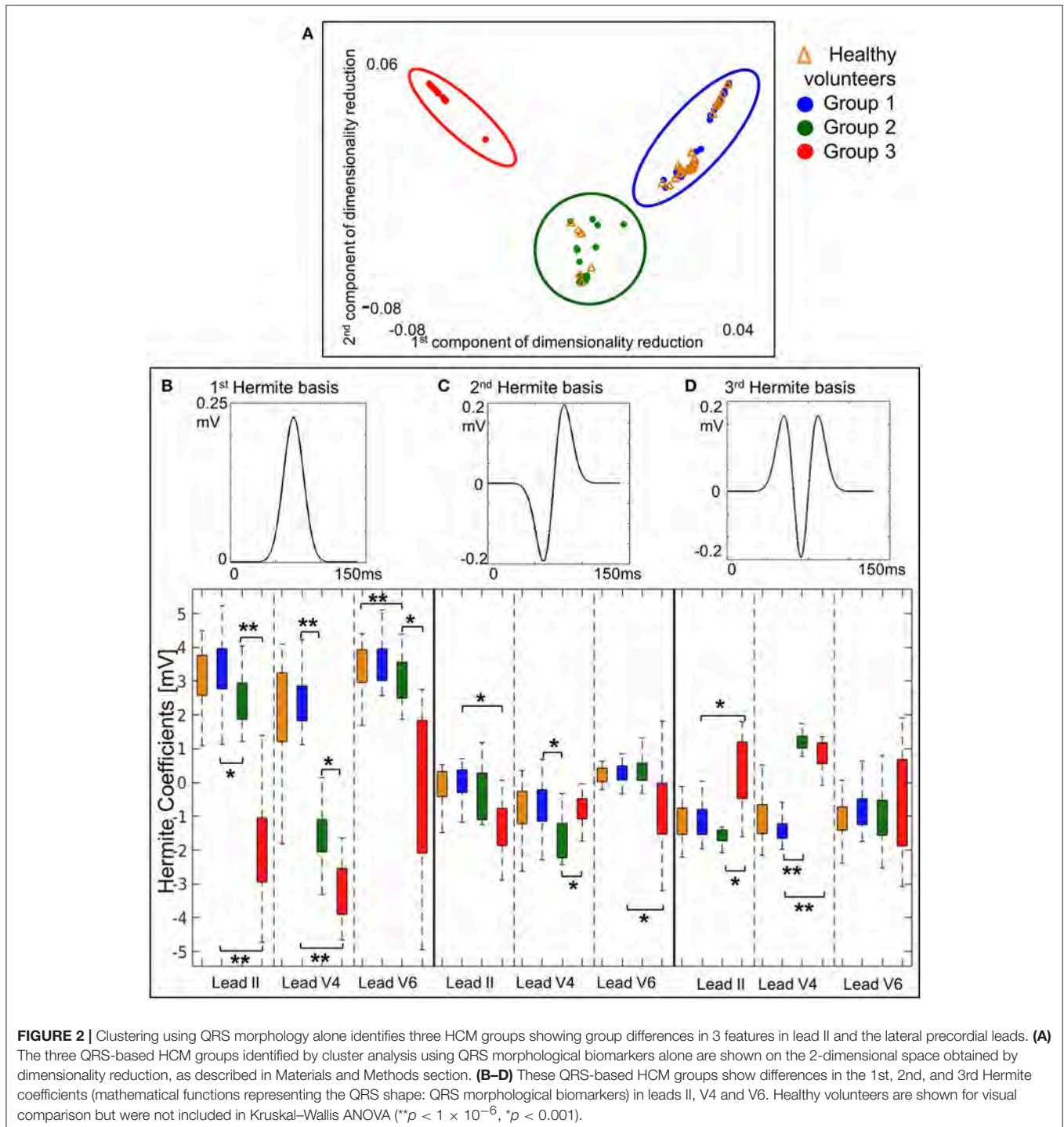
Eleven percent of patients were G+LVH- and the majority had isolated septal hypertrophy (68%).

### Clustering Using QRS Morphology Only

Based on QRS morphology alone, three HCM groups were obtained (**Figure 2A**), with the main differences in first, second and third Hermite basis in lead II and lateral precordial leads (**Figures 2B–D**; Supplementary Material 2.2, Figure S3). The ECG features for the QRS-based groups can be found in Supplementary Material 2.2, Table S1.

Group 1 was the largest with 52% of the patients, displaying normal QRS morphology (**Figures 2, 4A**). Group 2 (22% of patients) showed differences in V4 (in the first three Hermite bases) compared to healthy volunteers and Group 1 but no difference in V6 (**Figure 2**, all *post-hoc*  $p < 0.001$ ). Thus, lead V4 in Group 2 displayed shorter R wave duration (2: 38 ms, 1: 47 ms; 2 vs. 1,  $p = 0.003$ ) and deeper S waves (2:  $-1,170 \mu\text{V}$ , 1:  $-568 \mu\text{V}$ ; 2 vs. 1,  $p = 0.0003$ ) compared to Group 1 (**Figure 4A**; Supplementary Material 2.2, Table S1).

Group 3 (26% of patients) exhibited vast differences in lead II and V4-6 (in the first three Hermite bases) compared to



the other HCM groups and healthy volunteers (**Figure 2**, all *post-hoc*  $p < 6 \times 10^{-6}$ ). Differences in lead II were manifested as abnormally shifted QRS axis toward left axis deviation (LAD) (3:  $-37^\circ$ , 1:  $30^\circ$ , 2:  $29^\circ$ ; 3 vs. 1 and 3 vs. 2, all *post-hoc*  $p < 2 \times 10^{-7}$ ). In V4, the R wave duration was shorter than Group 1 and amplitude was shorter than Groups 1 and 2, S wave duration was longer than Group 1, and S amplitude was deeper than Group 1 (3 vs. 1 and 3 vs. 2, all *post-hoc*  $p <$

0.003). In V6, R wave duration and amplitude were shorter, and S wave duration and amplitude larger than both the other groups (3 vs. 1 and 3 vs. 2, all *post-hoc*  $p < 0.01$ ; **Figure 4A**).

Although ECG features were significantly different between the QRS-based groups, clinical features and markers of arrhythmic risk were not (**Table 3**), suggesting that QRS biomarkers alone, may not be useful for risk stratification.

**TABLE 3** | Clinical features for QRS-based HCM groups.

	Group 1 (n = 44)	Group 2 (n = 19)	Group 3 (n = 22)	p-value (group comparison)
Age, years	47 ± 15	44 ± 12	43 ± 13	0.34
Male	30 (68)	12 (63)	16 (73)	0.79
Body mass index, kg/m <sup>2</sup>	27 ± 5	25 ± 4	25 ± 4	0.16
Systolic BP, mmHg	120 ± 13	117 ± 13	114 ± 15	0.28
Diastolic BP, mmHg	74 ± 11	73 ± 13	68 ± 12	0.13
HCM Risk-SCD score, %	2.6 (1–11)	2.1 (1–6)	2.5 (1–9)	0.98
<b>CONVENTIONAL RISK FACTORS</b>				
NSVT	14 (32)	4 (21)	5 (23)	0.69
Syncope	6 (14)	2 (11)	2 (9)	0.91
Family History SCD	7 (16)	5 (26)	5 (23)	0.60
Abnormal exercise BPR	5 (11)	0	2 (9)	0.44
Massive LVH ≥30 mm	0	0	1 (5)	–
<b>HYPERTROPHY</b>				
LV mass index, g/m <sup>2</sup>	74 ± 28	63 ± 17	73 ± 26	0.30
Max LV wall, mm	19 ± 6	19 ± 5	21 ± 6	0.22
Hypertrophy morphology				<b>0.006</b>
No LVH (G+LVH-)	8 (18)	1 (5)	0	
Septal LVH	23 (51)	16 (85)	20 (90)	
Apical LVH	2 (4)	1 (5)	1 (5)	
Mixed septal & apical LVH	12 (27)	1 (5)	1 (5)	
<b>OTHER CLINICAL FEATURES</b>				
LV end-diastolic volume, ml	152 ± 28	155 ± 33	156 ± 41	0.87
LV end-systolic volume ml	40 ± 14	43 ± 14	41 ± 21	0.50
LV ejection fraction, %	74 ± 8	72 ± 8	74 ± 7	0.61
Left atrial diameter, mm	40 ± 7	37 ± 6	38 ± 8	0.25
LVOT gradient, mmHg	7.0 (4–111)	6.7 (2–110)	6.8 (3–92)	0.92
<b>GENOTYPE</b>				
Gene positive	28 (64)	13 (68)	17 (77)	0.59

Mean ± standard deviation, median (range) or number of patients (%). HCM, hypertrophic cardiomyopathy; SCD, sudden cardiac death; NSVT, non-sustained ventricular tachycardia; BP, blood pressure; BPR, blood pressure response; LVH, left ventricular hypertrophy; LV, left ventricular; G+LVH-, genotype positive HCM with normal wall thickness; LVOT, left ventricular outflow tract. Bold values mean p-value significant ( $p < 0.05$ ).

## Combined Clustering With QRS Morphology and T Wave Biomarkers Identify Four HCM Phenotypes

With the addition of T wave biomarkers to QRS morphology in the clustering analysis, four HCM groups were identified (Figure 3). TWI of the averaged beat in at least two contiguous leads in V3–6 was the principal T wave biomarker which subdivided Group 1 into two separate clusters. Groups 2 and 3 remained unchanged. Group 1A ( $n = 20$ ) had normal QRS with TWI. Group 1B ( $n = 24$ ) had normal QRS without TWI. Group 2 ( $n = 19$ ) had short R wave duration and deep S wave amplitude in V4. Group 3 ( $n = 22$ ) exhibited short R wave duration and amplitude together with long S wave duration and amplitude in V4–6, and left QRS axis deviation (Figure 4A; Supplementary Material 2.3). Table 4 summarizes

the clinical and standard ECG features for the four HCM groups. Of note, Group 1B had an absence of patients with pathological Q waves (*post-hoc*  $p < 0.0005$ ). There was no difference in demographics, LV volumes or genotype between the groups.

## Differences Between the Four HCM Phenotypes in HCM Risk-SCD Score and Extent of LV Hypertrophy

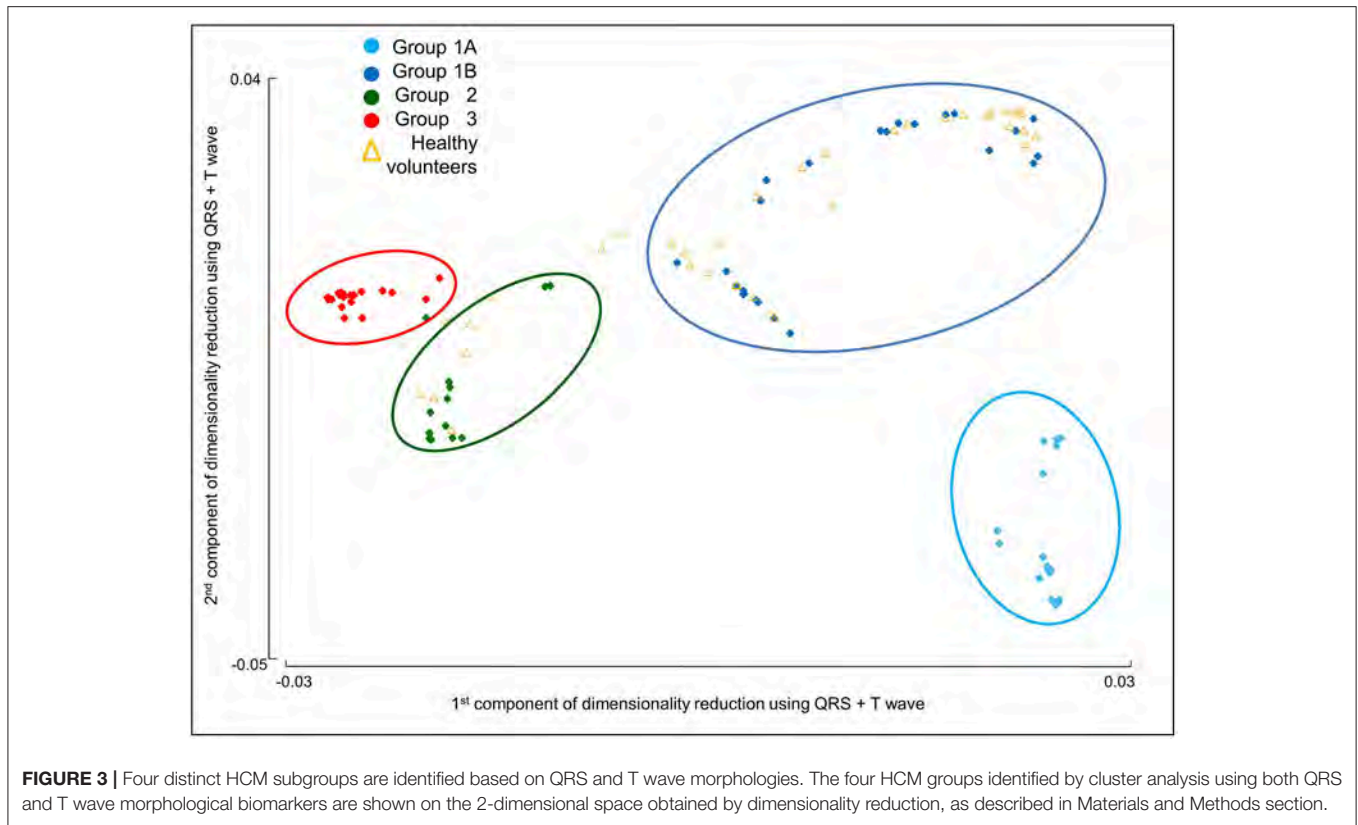
Group 1A had the highest median HCM Risk-SCD score (1A: 4.0%, 1B: 1.8%, 2: 2.1%, 3: 2.5%; 1A vs. 1B, 1A vs. 2, 1A vs. 3, all  $p \leq 0.02$ ; Figure 4C). NSVT differed between the 4 phenotypes ( $p = 0.016$ ) but this was not significant when corrected for the three other risk factors also tested. Group 1A also contained the only ICD patient with an appropriate shock. Group 1B contained 8 of the 9 G+LVH- patients (Table 4, *post-hoc*  $p = 0.0003$ ), of which 5 patients demonstrated ECG voltage criteria for LVH. In spite of their more abnormal QRS morphology, Groups 2 and 3 patients had a lower HCM Risk-SCD score than Group 1A patients (3: 2.5%, 2: 2.1%, 1A: 3.8%; 3 vs. 1A, 2 vs. 1A; *post-hoc*  $\leq 0.02$ ).

Group 1A exhibited a predominance of patients with mixed septal and apical hypertrophy (1A: 11 patients, 1B: 1 patient, 2: 1 patient, 3: 1 patient; *post-hoc*  $p = 4 \times 10^{-6}$ ; Figure 4B). Groups 1B, 2 and 3 had predominantly isolated septal hypertrophy (63, 85, and 90%, respectively).

We also evaluated the effect of excluding the 9 G+LVH-patients on the results (Supplementary Material 2.4, Table S2). The clustering algorithm yielded the same four remaining groups as with G+LVH-patients, with the same differences in QRS and T wave morphologies. Group 1A still exhibited a higher HCM Risk-SCD score compared to Group 1B and to Group 3 (*post-hoc*  $p = 0.006$  and  $0.04$ , respectively), and there was still a predominance of mixed septal and apical LVH in Group 1A.

## DISCUSSION

The main findings of this study are that four HCM phenotypes are identified based on QRS morphology and T wave biomarkers analyzed computationally using high fidelity ECGs, and they show differences in HCM Risk-SCD score and the distribution of LV hypertrophy from CMR. Patients with normal QRS morphology and primary TWI not secondary to QRS abnormalities (Group 1A) had the highest HCM Risk-SCD score and coexistence of septal and apical hypertrophy. Groups 2 and 3 with abnormalities in QRS morphology in V4 and V4–V6, respectively, had predominantly isolated septal hypertrophy. Since the ECG reflects ionic and structural abnormalities which are not captured by current measures within HCM risk stratification, the ECG-based classification proposed here may help to improve risk assessment. Our study shows the benefits of using machine learning methods to effectively dissect HCM heterogeneity, and presents a step forward in improving individual patient management.



## QRS Morphology

QRS morphology reflects the cardiac depolarization sequence and is therefore affected by structural abnormalities such as hypertrophy, cardiac disarray or fibrosis. Here we have quantitatively assessed the whole QRS shape (morphology) mathematically rather than describing discrete features. Cluster analysis with QRS morphology alone identified three groups with unique QRS features (Figure 2) but these QRS variations could not be accounted for by differences in hypertrophy. It is therefore likely that fibrosis (Dumont et al., 2006) and disarray affect depolarization and QRS particularly in patients in Groups 2 and 3. CMR studies using tissue characterization techniques such as late gadolinium enhancement (focal fibrosis), T1-mapping (diffuse fibrosis) and diffusion tensor imaging (disarray) are likely to provide further insights. Our analyses suggest that while QRS features may be informative for diagnosis in HCM (Konno et al., 2004), categorizing HCM groups on this basis showed no differences in association with known markers of risk.

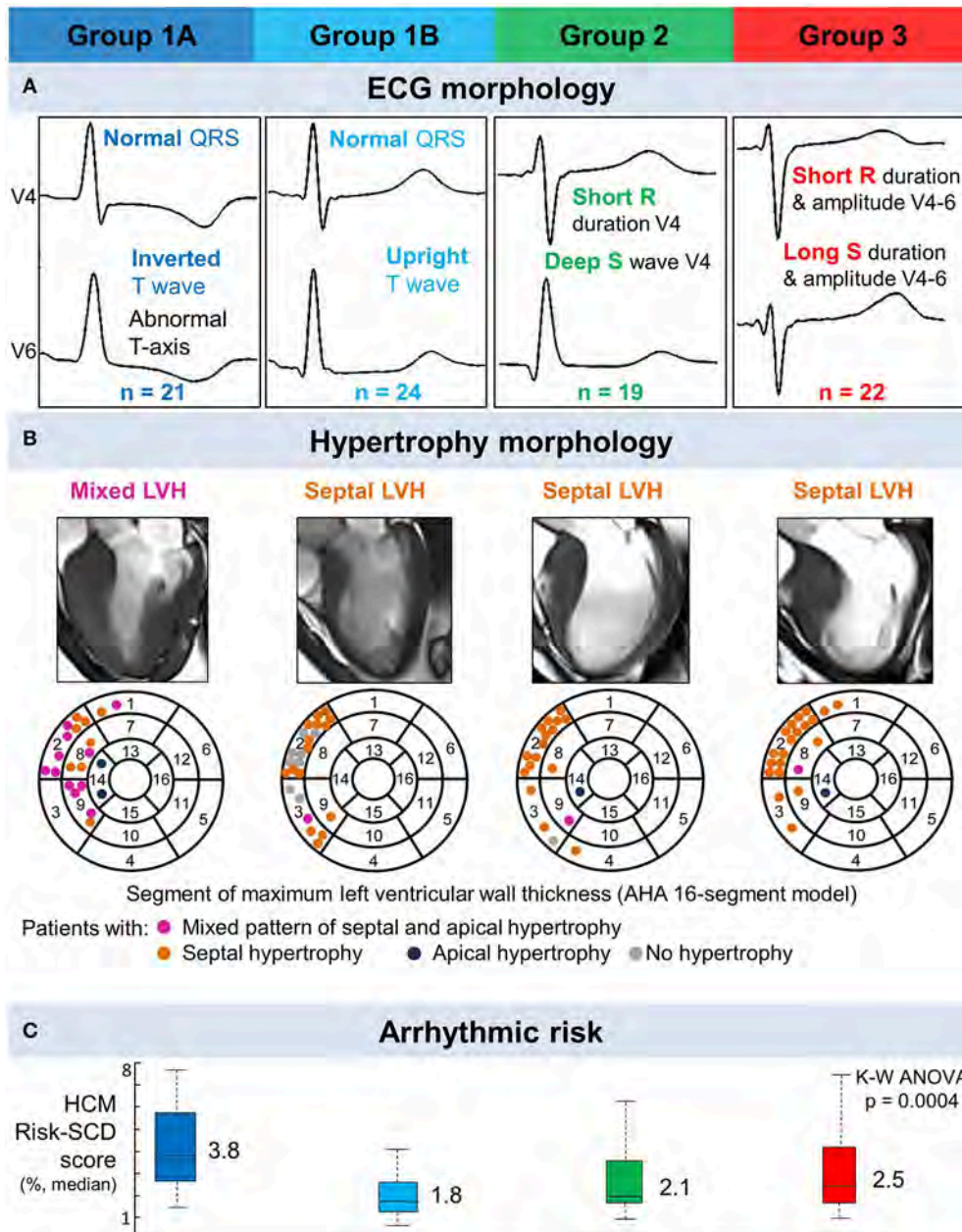
## T Wave Inversion

Following the classification in three groups based on QRS morphology, lateral TWI, a common repolarization abnormality in HCM (Papadakis et al., 2009), as the critical T wave biomarker which separated patients with normal QRS into those with and without TWI. Groups with abnormalities in QRS were not affected by the inclusion of TWI in the clustering. Group 1A patients with TWI, even though with a normal QRS, had

a higher HCM Risk-SCD score, an appropriate ICD shock and a predominance of mixed septal and apical pattern of hypertrophy. TWI can be considered as primary or secondary abnormalities. Primary TWI occurs with altered heterogeneity in action potential duration or morphology without changes in depolarization (normal QRS). Secondary TWI occurs with aberrant depolarization (abnormal QRS) in the context of normal action potential characteristics (Fisch, 1992). Thus Group 1A patients with normal QRS have primary. While TWI in Groups 2 and 3 with abnormal QRS represents secondary or combined primary and secondary TWI (Rautaharju et al., 2009). However only four patients in Group 2 and a single patient in Group 3 had TWI, while all 20 patients in Group 1A displayed TWI. TWI *per se* in HCM has been shown to increase SCD risk in some studies (Kuroda et al., 2002; Ostman-Smith et al., 2010) but not in others (Maron et al., 1982; Sherrid et al., 2009). Our results provide a more specific characterization of the influence of TWI in SCD risk, highlighting the importance of simultaneous normal QRS and TWI for increased risk. These results suggest that it is a primary TWI that increases SCD risk in HCM, rather than TWI secondary to depolarization abnormalities. A larger cohort is needed to confirm whether risk differs between patients with primary and secondary TWI.

TWI may be caused by repolarization abnormalities due to structural and ionic remodeling. In HCM, the overexpression of the L-type  $Ca^{2+}$  current, increased late sodium current and reduction of repolarization currents lead to prolongation and





**FIGURE 4 |** Four distinct ECG phenotypes in hypertrophic cardiomyopathy exhibit differences in hypertrophy morphology and arrhythmic risk. **(A)** Representative ECGs for patients in each of the four groups with distinct ECG morphology, identified by combined clustering with QRS morphology and T wave biomarkers. Group 1A—normal QRS with inverted T wave (primary T wave inversion), Group 1B—normal QRS with upright T wave, Group 2—short R wave duration and deep S wave in V4, Group 3—left axis deviation, short R wave duration and amplitude, and long S wave duration and amplitude in V4 and V6. **(B)** Distribution of hypertrophy illustrated using a representative CMR for each group (top), and the segment of maximum left ventricular wall thickness for each patient (marked as a dot) in each group using the AHA 16-segment model (Cerqueira et al., 2002). Group 1A had a predominance of patients with mixed septal and apical left ventricular hypertrophy (LVH; pink dots). Group 1B had the most gene positive patients with no hypertrophy (gray dots). Group 2 and 3 patients mainly had isolated septal hypertrophy (orange dots). Four patients had apical hypertrophy (navy dots). **(C)** HCM Risk-SCD score for each group. Patients with primary T wave inversion not secondary to QRS abnormalities (Group 1A), had the greatest HCM Risk-SCD score.

**TABLE 4** | Characteristics of the four HCM phenotypes from combined QRS and T wave clustering.

	<b>Group 1A (n = 20)</b>	<b>Group 1B (n = 24)</b>	<b>Group 2 (n = 19)</b>	<b>Group 3 (n = 22)</b>	<b>p-value (group comparison)</b>
Age, years	47 ± 12	48 ± 18	44 ± 12	43 ± 13	0.53
Male	15 (76)	15 (63)	12 (63)	16 (73)	0.78
Body mass index, kg/m <sup>2</sup>	28 ± 6	26 ± 4	25 ± 4	25 ± 4	0.06
HCM Risk-SCD score, %	4.0 (2–11) <sup>a,b,c</sup>	1.8 (1–4)	2.1 (1–6)	2.5 (1–9)	<b>0.0001</b>
<b>CONVENTIONAL RISK FACTORS</b>					
NSVT	11 (55) <sup>*</sup>	3 (13)	4 (21)	5 (23)	<b>0.016</b>
Syncope	2 (10)	4 (17)	2 (10)	2 (9)	0.88
Family History SCD	5 (25)	2 (8)	5 (26)	5 (23)	0.36
Abnormal exercise BPR	2 (10)	3 (12)	0	2 (9)	0.51
Massive LVH ≥30mm	0	0	0	1 (5)	-
Appropriate ICD shock	1	0	0	0	-
<b>HYPERTROPHY</b>					
LV mass index, g/m <sup>2</sup>	90 ± 27 <sup>a,b</sup>	62 ± 23	63 ± 17	73 ± 26	<b>0.0003</b>
Max LV wall, mm	22 ± 4 <sup>a</sup>	16 ± 5	19 ± 5	21 ± 6	<b>0.002</b>
Hypertrophy morphology					<b>1 × 10<sup>-7</sup></b>
No LVH (G+LVH-)	0	8 (33) <sup>*</sup>	1 (5)	0	
Septal LVH	7 (35) <sup>*</sup>	15 (63)	16 (85)	20 (90)	
Apical LVH	2 (10)	0	1 (5)	1 (5)	
Mixed septal & apical LVH	11 (52) <sup>*</sup>	1 (4)	1 (5)	1 (5)	
<b>OTHER CLINICAL FEATURES</b>					
LV EDV, ml	150 ± 26	154 ± 31	155 ± 33	156 ± 41	0.94
LV ESV, ml	35 ± 10	43 ± 15	43 ± 14	41 ± 21	0.26
LV ejection fraction, %	76 ± 7	72 ± 8	72 ± 8	74 ± 7	0.28
Left atrial diameter, mm	41 ± 5	39 ± 8	37 ± 6	38 ± 8	0.25
LVOT gradient, mmHg	6.5 (5–110)	7.2 (4–111)	6.7 (2–110)	6.8 (3–92)	0.81
<b>GENOTYPE</b>					
Gene positive	9 (45)	19 (79)	13 (68)	17 (77)	0.08
<b>ECG FEATURES</b>					
Heart rate, bpm	58 ± 10	58 ± 8	59 ± 13	58 ± 10	0.66
QRS axis, °	14 ± 30 <sup>c</sup>	44 ± 29 <sup>e</sup>	29 ± 34 <sup>f</sup>	-37 ± 28	<b>8 × 10<sup>-13</sup></b>
QRS duration, ms	94 ± 8	100 ± 21	96 ± 16	102 ± 13	0.23
QRS amplitude, mV	2006 ± 780	1762 ± 621	1807 ± 588	1737 ± 613	0.70
QRS ascending slope	93 ± 36	95 ± 33	92 ± 34	76 ± 23	0.19
QRS descending slope	-160 ± 60	-139 ± 56	-155 ± 60	-149 ± 56	0.47
Pathological Q waves	3 (14)	0 <sup>†</sup>	8 (38)	9 (43)	<b>0.0003</b>
T wave axis, °	156 ± 45 <sup>a,b,c</sup>	42 ± 27	49 ± 47	68 ± 41	<b>2 × 10<sup>-11</sup></b>
Abnormal T axis	17 (81) <sup>†</sup>	0 <sup>†</sup>	6 (33)	5 (25)	<b>7 × 10<sup>-8</sup></b>
T amplitude, mV	-135 ± 202 <sup>a,b,c</sup>	308 ± 144	219 ± 209	257 ± 214	<b>9 × 10<sup>-8</sup></b>
T wave inversion	21 (100) <sup>†</sup>	0	4 (21)	1 (5)	<b>8 × 10<sup>-17</sup></b>
Giant T wave inversion	5 (24) <sup>†</sup>	0	1 (5)	0	<b>0.003</b>
T peak to T end, ms	91 ± 17	80 ± 20	86 ± 16	84 ± 19	0.19
ST segment displacement*, mV	-13 ± 51 <sup>a,b,c</sup>	32 ± 33	38 ± 51	48 ± 38	<b>0.0001</b>
QTc interval, ms	452 ± 22	435 ± 26	429 ± 26	443 ± 29	0.05
JTc interval, ms	359 ± 47	349 ± 87	342 ± 97	366 ± 145	0.42

Mean ± standard deviation, median (range) or number of patients (%). HCM, hypertrophic cardiomyopathy; SCD, sudden cardiac death; NSVT, non-sustained ventricular tachycardia; BPR, blood pressure response; LVH, left ventricular hypertrophy; LV, left ventricular; G+LVH-, genotype positive HCM with normal wall thickness; EDV, end-diastolic volume; ESV, end-systolic volume; LVOT, left ventricular outflow tract.

<sup>a</sup>Group 1A vs. 1B, <sup>b</sup>1A vs. 2, <sup>c</sup>1A vs. 3, <sup>d</sup>1B vs. 2, <sup>e</sup>1B vs. 3, <sup>f</sup>2 vs. 3,  $p < 0.05$  on post-hoc pairwise comparisons ( $p$ -values multiplied by 6 for Bonferroni adjustment of 6 tests). <sup>\*</sup>, <sup>†</sup> $p < 0.05$  on post-hoc contingency table analysis ( $p$ -values multiplied by 16 and 8 for Bonferroni adjustment of 4 × 4 and 4 × 2 combinations, respectively). Bold values mean  $p$ -value significant ( $p < 0.05$ ).

heterogeneity in repolarization (Coppini et al., 2013; Passini et al., 2016). TWI may also be the result of ischaemia from microvascular dysfunction commonly seen in HCM (Petersen et al., 2007). A study has shown that patients with apical HCM and cavity obliteration had increased perfusion defects and NSVT rates as a result of ischaemia from extravascular compression of the coronary artery due to myocardial pressure during cavity obliteration (Matsubara et al., 2003). This may also be in keeping with Group 1A having the greatest number of patients with a mixed pattern of coexisting septal and apical hypertrophy which tend to show cavity obliteration and a worse prognosis (Yan et al., 2012); while Group 3 with 90% septal hypertrophy, were at lower risk. Further imaging studies with novel perfusion assessment may improve our understanding of the mechanisms of ischaemia and in turn, the repolarization abnormalities in HCM.

### Low Risk Phenotypes

Patients in Group 1B with normal QRS morphology and upright T waves were indistinguishable from healthy volunteers based on the extracted ECG features both with and without the inclusion G+LVH- patients. Finding the majority of G+LVH- patients (8 out of 9 patients) in Group 1B showed that we can discriminate these inherently low risk patients solely by the ECG, despite 5 of these patients meeting ECG voltage criteria for LVH. Furthermore, the presence of G+LVH- patients in Group 1B suggests that those patients with hypertrophy within this group are likely to have less severe disease and better prognosis (McLeod et al., 2009). We may also speculate that they will have minimal ionic remodeling, fibrosis, disarray and ischaemia giving rise to relatively normal depolarization and repolarization.

Despite the lack of hypertrophy, one G+LVH- patient was found in Group 2 which had QRS abnormalities in V4. This suggests that the ECG reflects the subtly abnormal myocardium which in this case was not hypertrophied, but may have been affected by disarray or ionic remodeling. Although risk is thought to be significantly lower in G+LVH- than in HCM with hypertrophy, there are still a very small number of SCD in G+LVH- patients (Varnava et al., 2001a; Pasquale et al., 2012). Further studies are needed to assess whether computational ECG phenotyping may aid stratification in this particularly challenging low-risk group.

Group 3 patients had marked QRS abnormalities: LAD with QRS differences in V4 to V6. LAD is well-known to be associated with LVH but other factors such as a degree of left anterior fascicular block could also account for this leftward axis. Group 3 patients mainly had isolated septal hypertrophy, yet QRS abnormalities were seen in the lateral leads suggesting that remodeling may occur distal from the septum causing less uniform electrical propagation in the lateral leads. No genotype association was seen across any group but our sample size may not have had adequate power to assess genotype-phenotype correlations.

### Clinical Implications

This study provides evidence that ECG phenotyping with advanced computational QRS morphology and T wave analysis is a powerful method of characterizing HCM heterogeneity.

Data suggest that HCM patients with a primary TWI (with normal QRS) are at greater risk of arrhythmia and SCD. This risk was associated with the distribution of hypertrophy (greater number of segments involved in mixed septal and apical pattern of hypertrophy) rather than magnitude of hypertrophy (as measured by maximum wall thickness or mass index). A large scale longitudinal study with cardiovascular end-point data will allow robust assessment of ECG phenotyping as an independent tool for accurate risk stratification. Studies involving computational image-based modeling and simulation are also needed to disentangle the relative contribution of structural, ischemic, and ionic factors which are likely to determine the heterogeneity in ECG biomarkers (Dutta et al., 2016). This improved understanding of HCM will eventually contribute to the development of new disease-modifying therapies.

### Limitations

Our study used digital ECG data from 12-lead Holter recorders. This enabled the identification of four distinct phenotypes using novel computational methods, which is not possible with standard paper ECGs collected in large studies. For these novel computational methods to be widely translated to clinical studies and practice, there needs to be drive toward digital ECG acquisition rather than paper print-outs, which require manual digitization before mathematical modeling and machine learning methods can be applied.

Given the large information content gathered for each patient in our study, the database is necessarily limited in the number of patients assembled. We included HCM patients without comorbidities (described in Supplemental Material 1.1) to ensure there were no confounders in our data. Our analysis was however able to identify different patient subgroups and also with differences in risk. Over the next 5–10 years, a large prospective long-term follow-up study such as the multicenter Hypertrophic Cardiomyopathy Registry (HCMR) (2,750 patients) (Kramer et al., 2015) may provide the data to determine whether our findings allow improving current risk stratification for SCD using scanned paper 12 lead ECGs. Our study provides the detailed analysis based on high fidelity ECG recordings that would enable such validation. As a follow-up, a larger dataset would make possible to consider a supervised machine learning approach, such as support vector machines, random forests or neural networks (Lyon et al., 2018), taking as input both ECG biomarkers and risk scores to identify the subgroup at higher risk. It would also allow the use of more complex unsupervised approaches such as self-organizing networks, as proposed in Lagerholm et al. (2000). However, large databases usually do not include the comprehensive set of modalities we include in our study. For example, these big databases of 967 and 2,485 HCM patients (McLeod et al., 2009; Cortez et al., 2017) do not provide high-fidelity recordings and lack CMR data.

Limited accuracy of an ECG criterion can also result from variations in electrode placement especially in precordial electrodes (Kania et al., 2014). However, minimal changes in morphology were observed in leads V4-6. Therefore, criteria based on the lateral leads (which indeed demonstrated the greatest discrimination in HCM) would be robust to the inevitable variability of electrode site placement.

## CONCLUSIONS

Four HCM phenotypes were identified based on QRS morphology and T wave biomarkers using a machine learning approach. Patients with primary TWI not secondary to QRS abnormalities had an increased HCM Risk-SCD score and coexisting septal and apical hypertrophy. These results, and the nature of the underlying processes captured by the ECG, suggest that computational ECG phenotyping has the potential to be a novel and independent factor for risk stratification.

## AUTHOR CONTRIBUTIONS

RA recruited the HCM and Control populations and performed the statistical analysis; AL performed the ECG signal analysis and the computational clustering; RA and AL worked on the writing of the manuscript; MM and EO provided help in the collection of the data; PL and NdF gave input on the computational methods for signal processing and clustering; SN, HW, AM, and BR provided help and guidance on the study design and the writing of the manuscript.

## FUNDING

AL is supported by a scholarship provided by the British Heart Foundation Centre of Research Excellence. RA is supported

## REFERENCES

- Adabag, A. S., Maron, B. J., Appelbaum, E., Harrigan, C. J., Buros, J. L., Gibson, C. M., et al. (2008). Occurrence and frequency of arrhythmias in hypertrophic cardiomyopathy in relation to delayed enhancement on cardiovascular magnetic resonance. *J. Am. Coll. Cardiol.* 51, 1369–1374. doi: 10.1016/j.jacc.2007.11.071
- Belkin, M., and Niyogi, P. (2003). Laplacian eigenmaps for dimensionality reduction and data representation. *Neural Comput.* 15, 1373–1396. doi: 10.1162/089976603321780317
- Cai, D., Zhang, C., and He, X. (2010). “Unsupervised feature selection for multi-cluster data,” in *Proceedings of the 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining [Internet]* (New York, NY: ACM) (cited June 10, 2016), 333–342. Available online at: <http://doi.acm.org/10.1145/1835804.1835848>
- Cerqueira, M. D., Weissman, N. J., Dilsizian, V., Jacobs, A. K., Kaul, S., Laskey, W. K., et al. (2002). Standardized myocardial segmentation and nomenclature for tomographic imaging of the heart. A statement for healthcare professionals from the Cardiac imaging committee of the council on clinical Cardiology of the American Heart Association. *Circulation* 105, 539–542. doi: 10.1161/hc0402.102975
- Christiaans, I., van Engelen, K., van Langen, I. M., Birnie, E., Bonsel, G. J., Elliott, P. M., et al. (2010). Risk stratification for sudden cardiac death in hypertrophic cardiomyopathy: systematic review of clinical risk markers. *Europace* 12, 313–321. doi: 10.1093/europace/eup431
- Coppini, R., Ferrantini, C., Yao, L., Fan, P., Del Lungo, M., Stillitano, F., et al. (2013). Late sodium current inhibition reverses electromechanical dysfunction in human hypertrophic cardiomyopathy. *Circulation* 127, 575–584. doi: 10.1161/CIRCULATIONAHA.112.134932
- Cortez, D., Schlegel, T. T., Ackerman, M. J., and Bos, J. M. (2017). ECG-derived spatial QRS-T angle is strongly associated with hypertrophic cardiomyopathy. *J. Electrocardiol.* 50, 195–202. doi: 10.1016/j.jelectrocard.2016.10.001
- Dumont, C. A., Monserrat, L., Soler, R., Rodriguez, E., Fernandez, X., Peteiro, J., et al. (2006). Interpretation of electrocardiographic abnormalities in

by a British Heart Foundation Clinical Research Training Fellowship. AM and BR are supported by BR's Wellcome Trust Senior Research Fellowship in Basic Biomedical Sciences. PL is supported by project TIN2014-53567-R and TEC2013-44666-R, Spain and Grupo Consolidado BSICoS from DGA, Aragón, Spain. SN and HW acknowledge support from the Oxford NIHR Biomedical Research Centre and the British Heart Foundation. This project has also received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement No. 675451 (CompBioMed project).

## ACKNOWLEDGMENTS

We thank Jane Francis for assisting with CMR scanning, Joanne Sellwood for nursing support, Rachel Given for data input, the Oxford Medical Genetics Laboratories, Oxford University Hospitals NHS Foundation Trust for genetic data, Dr. Sanjay Manohar for helpful discussions and Prof. Martin Farrall for advice with statistics.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fphys.2018.00213/full#supplementary-material>

- hypertrophic cardiomyopathy with cardiac magnetic resonance. *Eur. Heart J.* 27, 1725–31. doi: 10.1093/eurheartj/ehl101
- Dutta, S., Mincholé, A., Zacur, E., Quinn, T. A., Taggart, P., and Rodriguez, B. (2016). Early afterdepolarizations promote transmural reentry in ischemic human ventricles with reduced repolarization reserve. *Prog. Biophys. Mol. Biol.* 120, 236–248. doi: 10.1016/j.pbiomolbio.2016.01.008
- Elliott, P. M., Borger, M. A., Borggrefe, M., Cecchi, F., Charron, P., Hagege, A. A., et al. (2014). ESC guidelines on diagnosis and management of hypertrophic cardiomyopathy: the task force for the diagnosis and management of hypertrophic cardiomyopathy of the European Society of Cardiology (ESC). *Eur. Heart J.* 35, 2733–2779. doi: 10.1093/eurheartj/ehu284
- Elliott, P. M., Poloniecki, J., Dickie, S., Sharma, S., Monserrat, L., Varnava, A., et al. (2000). Sudden death in hypertrophic cardiomyopathy: identification of high risk patients. *J. Am. Coll. Cardiol.* 36, 2212–2218. doi: 10.1016/S0735-1097(00)01003-2
- Ester, M., Kriegel, H., Sander, J., and Xu, X. (1996). “A density-based algorithm for discovering clusters in large spatial databases with noise,” in *KDD-96 Proceedings* (Portland, OR: AAAI Press), 226–231.
- Fisch, C. (1992). “Electrocardiography and vectorcardiography,” in *Heart Disease: A Textbook of Cardiovascular Medicine, 4th Edn.*, ed E. Braunwald (Philadelphia, PA: WB Saunders), 116–160.
- Kania, M., Rix, H., Fereniec, M., Zavala-Fernandez, H., Janusek, D., Mroczka, T., et al. (2014). The effect of precordial lead displacement on ECG morphology. *Med. Biol. Eng. Comput.* 52, 109–119. doi: 10.1007/s11517-013-1115-9
- Konno, T., Shimizu, M., Ino, H., Yamaguchi, M., Terai, H., Uchiyama, K., et al. (2004). Diagnostic value of abnormal Q waves for identification of preclinical carriers of hypertrophic cardiomyopathy based on a molecular genetic diagnosis. *Eur. Heart J.* 25, 246–251. doi: 10.1016/j.ehj.2003.10.031
- Kramer, C. M., Appelbaum, E., Desai, M. Y., Desvigne-Nickens, P., DiMarco, J. P., Friedrich, M. G., et al. (2015). Hypertrophic cardiomyopathy registry: the rationale and design of an international, observational study of hypertrophic cardiomyopathy. *Am. Heart J.* 170, 223–230. doi: 10.1016/j.ahj.2015.05.013
- Kuroda, N., Ohnishi, Y., Yoshida, A., Kimura, A., and Yokoyama, M. (2002). Clinical significance of T-wave alternans in Hypertrophic Cardiomyopathy. *Circ. J.* 66, 457–462. doi: 10.1253/circj.66.457

- Lagerholm, M., Peterson, C., Braccini, G., Edenbrandt, L., and Sornmo, L. (2000). Clustering ECG complexes using Hermite functions and self-organizing maps. *IEEE Trans. Biomed. Eng.* 47, 838–848. doi: 10.1109/10.846677
- Laguna, P., Jane, R., Olmos, S., Thakor, N. V., Rix, H., and Caminal, P. (1996). Adaptive estimation of QRS complex wave features of ECG signal by the Hermite model. *Med. Biol. Eng. Comput.* 34, 58–68. doi: 10.1007/BF02637023
- Lakdawala, N. K., Thune, J. J., Maron, B. J., Cirino, A. L., Havndrup, O., Bundgaard, H., et al. (2011). Electrocardiographic features of sarcomere mutation carriers with and without clinically overt hypertrophic cardiomyopathy. *Am. J. Cardiol.* 108, 1606–1613. doi: 10.1016/j.amjcard.2011.07.019
- Lyon, A., Mincholé, A., Martínez, J. P., Laguna, P., and Rodriguez, B. (2018). Computational techniques for ECG analysis and interpretation in light of their contribution to medical advances. *J. R. Soc. Interf.* 15:20170821. doi: 10.1098/rsif.2017.0821
- Macfarlane, P. W., Devine, B., Latif, S., McLaughlin, S., Shoaib, D. B., and Watts, M. P. (1990). Methodology of ECG interpretation in the Glasgow program. *Methods Inf. Med.* 29, 354–361.
- Maron, B. J. (2002). Hypertrophic cardiomyopathy: a systematic review. *JAMA* 287, 1308–1320. doi: 10.1001/jama.287.10.1308
- Maron, B. J., Casey, S. A., Chan, R. H., Garberich, R. F., Rowin, E. J., and Maron, M. S. (2015). Independent assessment of the European Society of Cardiology sudden death risk model for Hypertrophic Cardiomyopathy. *Am. J. Cardiol.* 116, 757–764. doi: 10.1016/j.amjcard.2015.05.047
- Maron, B. J., Roberts, W. C., and Epstein, S. E. (1982). Sudden death in hypertrophic cardiomyopathy: a profile of 78 patients. *Circulation* 65, 1388–1394. doi: 10.1161/01.CIR.65.7.1388
- Maron, B. J., Shen, W. K., Link, M. S., Epstein, A. E., Almquist, A. K., Daubert, J. P., et al. (2000). Efficacy of implantable cardioverter-defibrillators for the prevention of sudden death in patients with hypertrophic cardiomyopathy. *N Engl. J. Med.* 342, 365–373. doi: 10.1056/NEJM200002103420601
- Maron, B. J., Spirito, P., Shen, W.-K., Haas, T. S., Formisano, F., Link, M. S., et al. (2007). Implantable cardioverter-defibrillators and prevention of sudden cardiac death in hypertrophic cardiomyopathy. *JAMA* 298, 405–412. doi: 10.1001/jama.298.4.405
- Maron, B. J., Wolfson, J. K., Ciró, E., and Spirito, P. (1983). Relation of electrocardiographic abnormalities and patterns of left ventricular hypertrophy identified by 2-dimensional echocardiography in patients with hypertrophic cardiomyopathy. *Am. J. Cardiol.* 51, 189–194. doi: 10.1016/S0002-9149(83)80034-4
- Martínez, J. P., Almeida, R., Olmos, S., Rocha, A. P., and Laguna, P. (2004). A wavelet-based ECG delineator: evaluation on standard databases. *IEEE Trans. Biomed. Eng.* 51, 570–581. doi: 10.1109/TBME.2003.821031
- Matsubara, K., Nakamura, T., Kuribayashi, T., Azuma, A., and Nakagawa, M. (2003). Sustained cavity obliteration and apical aneurysm formation in apical hypertrophic cardiomyopathy. *J. Am. Coll. Cardiol.* 42, 288–295. doi: 10.1016/S0735-1097(03)00576-X
- McLeod, C. J., Ackerman, M. J., Nishimura, R. A., Tajik, A. J., Gersh, B. J., and Ommen, S. R. (2009). Outcome of patients with hypertrophic cardiomyopathy and a normal electrocardiogram. *J. Am. Coll. Cardiol.* 54, 229–233. doi: 10.1016/j.jacc.2009.02.071
- O'Mahony, C., Elliott, P., and McKenna, W. (2013). Sudden cardiac death in Hypertrophic Cardiomyopathy. *Circ. Arrhythm. Electrophysiol.* 6, 443–451. doi: 10.1161/CIRCEP.111.962043
- O'Mahony, C., Jichi, F., Pavlou, M., Monserrat, L., Anastasakis, A., Rapezzi, C., et al. (2014). A novel clinical risk prediction model for sudden cardiac death in hypertrophic cardiomyopathy (HCM Risk-SCD). *Eur. Heart J.* 35, 2010–2020. doi: 10.1093/eurheartj/ehp439
- Ostman-Smith, I., Wisten, A., Nylander, E., Bratt, E. L., Granelli, A., Oulhaj, A., et al. (2010). Electrocardiographic amplitudes: a new risk factor for sudden death in hypertrophic cardiomyopathy. *Eur. Heart J.* 31, 439–449. doi: 10.1093/eurheartj/ehp443
- Papadakis, M., Basavarajiah, S., Rawlins, J., Edwards, C., Makan, J., Firoozi, S., et al. (2009). Prevalence and significance of T-wave inversions in predominantly Caucasian adolescent athletes. *Eur. Heart J.* 30, 1728–1735. doi: 10.1093/eurheartj/ehp164
- Pasquale, F., Syrris, P., Kaski, J. P., Mogensen, J., McKenna, W. J., and Elliott, P. (2012). Long-term outcomes in hypertrophic cardiomyopathy caused by mutations in the cardiac troponin T gene. *Circ. Cardiovasc. Genet.* 5, 10–17. doi: 10.1161/CIRCGENETICS.111.959973
- Passini, E., Mincholé, A., Coppini, R., Cerbai, E., Rodriguez, B., Severi, S., et al. (2016). Mechanisms of pro-arrhythmic abnormalities in ventricular repolarisation and anti-arrhythmic therapies in human hypertrophic cardiomyopathy. *J. Mol. Cell Cardiol.* 96, 72–81. doi: 10.1016/j.yjmcc.2015.09.003
- Petersen, S. E., Jerosch-Herold, M., Hudsmith, L. E., Robson, M. D., Francis, J. M., Doll, H. A., et al. (2007). Evidence for microvascular dysfunction in hypertrophic cardiomyopathy: new insights from multiparametric magnetic resonance imaging. *Circulation* 115, 2418–2425. doi: 10.1161/CIRCULATIONAHA.106.657023
- Rautaharju, P. M., Surawicz, B., and Gettes, L. S. (2009). AHA/ACCF/HRS Recommendations for the standardization and interpretation of the electrocardiogram part IV: The ST segment, T and U waves, and the QT interval: 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.* 53, 982–991. doi: 10.1016/j.jacc.2008.12.014
- Savage, D. D., Seides, S. F., Clark, C. E., Henry, W. L., Maron, B. J., Robinson, F. C., et al. (1978). Electrocardiographic findings in patients with obstructive and nonobstructive hypertrophic cardiomyopathy. *Circulation* 58(3 Pt 1), 402–408. doi: 10.1161/01.CIR.58.3.402
- Sen-Chowdhry, S., Jacoby, D., Moon, J. C., and McKenna, W. J. (2016). Update on hypertrophic cardiomyopathy and a guide to the guidelines. *Nat. Rev. Cardiol.* 13, 651–675. doi: 10.1038/nrcardio.2016.140
- Sherrid, M. V., Cotiga, D., Hart, D., Ehlert, F., Haas, T. S., Shen, W.-K., et al. (2009). Relation of 12-lead electrocardiogram patterns to implanted defibrillator-terminated ventricular tachyarrhythmias in hypertrophic cardiomyopathy. *Am. J. Cardiol.* 104, 1722–1726. doi: 10.1016/j.amjcard.2009.07.056
- Sörnmo, L., and Laguna, P. (eds.). (2005). *Bioelectrical Signal Processing in Cardiac and Neurological Applications*. Burlington: Academic Press (cited October 10, 2014). Available online at: <http://www.sciencedirect.com/science/article/pii/B9780124375529500015>
- Sornmo, L., Ola Borjesson, P., Nygard, M.-E., and Pahlm, O. (1981). A method for evaluation of QRS shape features using a mathematical model for the ECG. *IEEE Trans. Biomed. Eng.* 28, 713–717. doi: 10.1109/TBME.1981.324666
- Spirito, P., Bellone, P., Harris, K. M., Bernabo, P., Bruzzi, P., and Maron, B. J. (2000). Magnitude of left ventricular hypertrophy and risk of sudden death in hypertrophic cardiomyopathy. *N Engl. J. Med.* 342, 1778–1785. doi: 10.1056/NEJM200006153422403
- Varnava, A. M., Elliott, P. M., Baboonian, C., Davison, F., Davies, M. J., and McKenna, W. J. (2001a). Hypertrophic cardiomyopathy: histopathological features of sudden death in cardiac troponin T disease. *Circulation* 104, 1380–1384. doi: 10.1161/hc3701.095952
- Varnava, A. M., Elliott, P. M., Mahon, N., Davies, M. J., and McKenna, W. J. (2001b). Relation between myocyte disarray and outcome in hypertrophic cardiomyopathy. *Am. J. Cardiol.* 88, 275–279. doi: 10.1016/S0002-9149(01)01640-X
- Yan, L., Wang, Z., Xu, Z., Li, Y., Tao, Y., and Fan, C. (2012). Two hundred eight patients with apical hypertrophic cardiomyopathy in china: clinical feature, prognosis, and comparison of pure and mixed forms. *Clin. Cardiol.* 35, 101–106. doi: 10.1002/clc.20995

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Lyon, Ariga, Mincholé, Mahmod, Ormondroyd, Laguna, de Freitas, Neubauer, Watkins and Rodriguez. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.