

Bayesian Classification Applied to Strain in Arrhythmogenic Left-Ventricle Cardiomyopathy

Yolanda Vives-Gilbert¹, Begoña Igual², Santiago Jiménez-Serrano¹, Jorge Sanz³, Raquel Cervigón⁴, Antonio Cebrián¹, Jose M Santabárbara², José Millet¹, Esther Zorio³, Francisco Castells¹

¹Instituto ITACA, Universitat Politècnica de València, Valencia, Spain

²ERESA Grupo Médico, Valencia, Spain

³Hospital Universitari i Politècnic La Fe, Valencia, Spain

⁴Universidad de Castilla-la Mancha, Cuenca, Spain

Abstract

Arrhythmogenic cardiomyopathy (AC) is a rare disease associated with ventricular arrhythmias and sudden cardiac death. While AC of the right ventricle has been more extensively studied, exclusive left-ventricle involvement needs to be better characterized. Myocardial strain, obtained by feature tracking, provide insight into its biomechanical behavior. To characterize it, multivariate classifiers can be applied.

The sample consisted of 13 AC-LV and 13 non-carriers of the mutation. The feature tracking algorithm of Circle cvi⁴² was applied to the cardiac magnetic resonance of each patient. A Naïve Bayes classifier with a feature subset selection method was applied to the parameters of peak strain, strain rate, displacement and velocity.

We obtained an accuracy of 90% in NB and we arrived to 93% for CFS-NB. The strain parameters selected by the FSS algorithm were three: longitudinal peak strain and peak systolic and diastolic velocities. In all the selected features, AC-LV patients had smaller values as controls.

In conclusion, myocardial strain is affected in AC-LV patients. Naïve Bayes classifiers allow obtaining a good discriminating accuracy among groups.

1. Introduction

Arrhythmogenic cardiomyopathy (AC) is a disease characterized by a progressive replacement of the ventricular myocardium by fibrofatty tissue and it is associated with ventricular arrhythmias and risk of sudden cardiac death. Arrhythmogenic right-ventricle cardiomyopathy has been more extensively reported in the literature [1], while isolated or predominantly left ventricular (LV) involvement has been less studied [2].

The diagnosis of arrhythmogenic LV cardiomyopathy (AC-LV) is challenging, and its prevalence is likely to be underestimated [3].

Cardiac imaging has helped in the recent years to diagnose and follow the progression of AC, being cardiovascular magnetic resonance (CMR) one of the most valuable methods due to the high-spatial resolution and its non-invasive nature [4]. Myocardial strain is the result of the complex interaction of intrinsic and extrinsic forces applied to the heart, a tissue with variable elastic properties, so measurements of local strain patterns (deformation) provide insights into the biomechanical behaviour of the myocardium. If myocardial tissue is replaced by fat, as in AC-LV, stiffness changes and the pattern and the magnitude of the myocardial deformation can be affected [5].

Feature tracking algorithms allow obtaining many strain-related parameters, like peak strain, strain rate and strain velocity among others [6,7]. Due to the large number of derived variables, it is difficult to make a general characterization. Supervised machine learning algorithms could be useful tools to study: (1) if AC-LV patients have myocardial strain affected when compared to normal controls and in that case (2) to identify which strain parameters are the most discriminative.

In the recent years, there has been a growing interest in applying machine learning techniques using strain features obtained from cardiac images, mainly echocardiographic [8]. Bayesian classifiers [9] are simple classifiers that perform pretty well and have been applied before to other medical fields. These classifiers can be combined with feature subset selection (FSS) methods to automatically identify the most discriminative strain parameters.

The objective of this work is to study strain in AC-LV by applying machine learning decision support systems like bayesian classifiers and combine them with FSS. While different studies have been performed before to characterize strain in AC-RV [7], as far as we know, this

is the first study that characterizes strain in AC-LV.

2. Materials and methods

2.1. Sample

The sample consisted of 26 patients split into two groups: 13 patients diagnosed with AC-LV and 13 relatives non-carriers of the pathogenic mutation of the proband (“non-carriers” by the following). There were not significant differences in sex and age between groups. AC-LV was diagnosed when definitive or borderline Task Force Criteria for AC were met plus typical patterns of LV involvement were identified, like late gadolinium enhancement (LGE) at CMR and electrocardiogram (ECG). All individuals in the “non-carriers” group had normal ECG, physical examination and CMR study and were free from disease with potential impact on the cardiovascular system.

The study was approved by the local research ethics committee, and all subjects gave their informed consent to participate.

2.1. MRI acquisition

CMR studies were performed using a 1.5-Tesla scanner (Siemens Avanto or GE Signa HDxt). Vertical and horizontal long-axis slices, a stack of contiguous cine short-axis slices from the atrioventricular ring to the apex and para-axial slices from diaphragm to the entire outflow tract were acquired using a steady-state free precession pulse sequence (slice thickness=6 or 8 mm, no interslice gap; repetition time/echo time=56/1.2 ms for Siemens Avanto; repetition time/echo time=3.8/1.7 ms for GE Signa HDxt).

2.2. Feature tracking

All cine CMR studies were analyzed using the feature tracking module of a dedicated software (Circle CVI⁴² v5.5.1, Calgary, Canada). Endocardial and epicardial contours at end diastole were drawn in a semi-automated manner. Afterwards, Circle calculated global strain parameters in three different ways: 2D short axis, 2D long axis and 3D. Strain in 3 different directions are calculated: radial, circumferential and longitudinal (Figure 1).

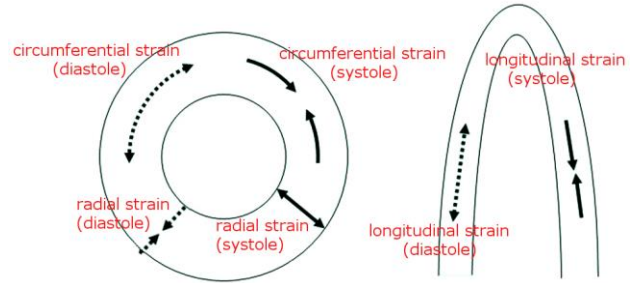


Figure 1. Graphic representation of the three principal myocardial deformations: radial, circumferential and longitudinal.

2D methods include only radial and circumferential parameters in short axis and radial and longitudinal parameters in long axis, while 3D results include radial, circumferential and longitudinal strain parameters, so we took the 3D results in order to have all features obtained in the same way. The following strain parameters in the radial, circumferential and longitudinal deformation axis were used for the classification: peak strain, peak systolic and diastolic strain rate, peak displacement and peak systolic and diastolic velocity.

2.3. Supervised classification

Supervised classification techniques require the definition of a class variable, in which each group (patients and non-carriers) has a different value. Here we compare classifiers to distinguish between patients with AC-LV and relatives without the mutation with the objective of identifying firstly if strain parameters are discriminant enough to distinguish between groups and secondly to study which LV strain features are mostly affected by the disease. Each patient case is defined by a pattern vector of predictive variables representing the different strain parameters. Examinations involve two phases: 1) model training (construction of a model capable of assigning a case into one of the two classes (AC-LV or non-carrier)), and 2) classification (the model assigns a class label to the new cases).

Naïve Bayes (NB) classifier [9] is a simple probabilistic classifier based on strong independence assumptions between the features. This model is a quite straightforward model and perform relatively well in real problems. Naïve Bayes correlation-based feature selection (CFS-NB) is a naïve Bayes classifier built using the predictive variables selected previously using a multivariate filter feature subset selection (FSS) method known as the correlation-based feature selection (CFS) [10]. CFS method evaluates the worth of a subset of attributes by considering the individual predictive ability of each feature along with the degree of redundancy between them which means that identifies features highly correlated with the class variable (and uncorrelated with

each other) and dismiss those irrelevant because of being redundant and do not improve the prediction accuracy of the model. In this work we applied the best-first search [11]. To validate the performance of each model, we applied 10 times stratified k-fold cross-validation with $k = 10$ [12].

To evaluate the performance of these models, the following metrics are used: 1) accuracy: the proportion of cases correctly classified by the model, 2) sensitivity: the proportion of real positive cases correctly classified as positive and 3) specificity: the proportion of real negative cases correctly classified as negative.

3. Results

The performance of NB and CFS-NB classifiers is shown in Table 1.

Table 1. Performance of NB and CFS-NB classifiers.

Classifier	Accuracy	Sensitivity	Specificity
NB	90%	86.67%	90%
CFS-NB	93.33%	86.67%	100%

CFS selected three features as the most predictive (AC-LV vs non-affected, p-value): longitudinal peak strain (-13.1+/-1.4 vs -16.2+/-1.4, $p < 0.001$) and radial peak systolic (27+/-5.5 vs 36.2+/-8.3, $p = 0.003$) and diastolic (-27.8+/-5.9 vs -37.1+/-6.3, $p = 0.001$) velocity (Figure 2).

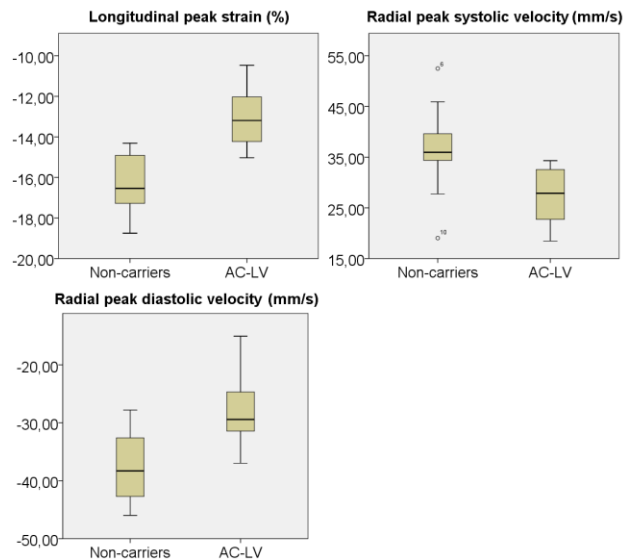


Figure 2. Boxplot of the most discriminative variables selected by CFS-NB classifier: longitudinal peak strain (%), radial peak systolic velocity and radial peak diastolic velocity (mm/s).

Peak strain values of the most discriminative features

are obtained from the graphs in Figure 3. In all the cases AC-LV values are smaller than those of the non-carriers.

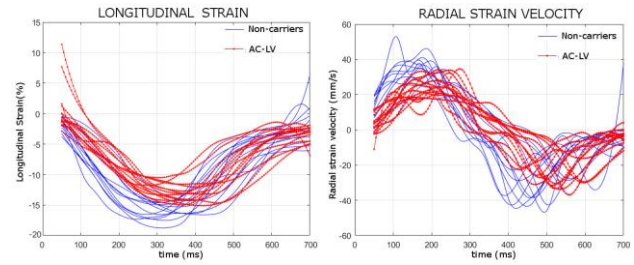


Figure 3. Longitudinal strain and radial strain velocity of AC-LV (red) and “non-carriers” (blue).

4. Discussion

The diagnosis of AC-LV is a clinical challenge because of its low prevalence and the lack of criteria specific for AC patients with dominant LV involvement, being the Task Force criteria [13] and its 2010 revision [1] oriented to AC-RV.

The usefulness of strain imaging to objectively quantify myocardial function in AC has been previously demonstrated in tissue-Doppler and speckle-tracking echocardiography studies [4]. However, with the advances of new computational image analysis tools, conventional cine CMR and tissue tracking algorithms are being used to quantify strain with success [14,15].

Applying classifiers to global strain features allows us having a general overview of the mechanics of the left ventricle. Peak strain, systolic and diastolic strain rate, displacement and systolic and diastolic velocity for the three axis, radial, circumferential and longitudinal, were considered.

CFS-NB was the most appropriate classifier to model strain for two reasons: (1) it obtained the highest accuracy, with 93% and a specificity of 100% and (2) as it uses FSS, it selected the most predictive strain parameters, obtaining a simpler model easily interpretable by clinicians. In all the selected features, AC-LV patients had smaller strain values as “non-carriers”. NB also obtained good accuracy results, with 90%, however specificity was not so high as in CFS-NB.

4.1. Limitations

Taking into account that AC-LV is a rare disease, the sample is quite small. The main problem of applying machine learning techniques with small datasets is the overfitting problems, which is when the model fits the training data too well but its predictions are poor. The same effect can be encountered when applying a feature selection algorithm. Overuse of the accuracy estimates in FSS may cause overfitting in the feature-subset space which means that if there are many feature subsets, it is

likely that one of them leads to a hypothesis that has high predictive accuracy. While the use of the cross-validation algorithm minimizes the overfitting effect and gives an insight on how the model will generalize to an independent dataset, it is important to take into account possible overfitting problems that could have been arisen in this work.

5. Conclusions

Cine CMR and feature tracking algorithms allow obtaining accurate strain parameters. This study shows the reduced global LV myocardial strain in patients with AC-LV when compared to “non-carriers”. When strain is combined with machine learning algorithms like Naïve Bayes, we obtain a predictive model with an accuracy of 93%. Longitudinal peak strain and radial peak velocity (systolic and diastolic) were the most discriminant features.

Acknowledgements

This work has been supported by the Spanish Government, under the grant DPI2015-70821-R received from the research program ‘Retos de la Sociedad’ by the Ministry of Economics and Competitiveness.

References

- [1] Marcus FI, McKenna WJ, Sherrill D, Basso C, Bauce B, Bluemke DA, et al. Diagnosis of arrhythmogenic right ventricular cardiomyopathy/dysplasia: Proposed Modification of the Task Force Criteria. *Eur Heart J*. 2010 Apr 1;31(7):806–14.
- [2] Sen-Chowdhry S, Syrris P, Prasad SK, Hughes SE, Merrifield R, Ward D, et al. Left-Dominant Arrhythmogenic Cardiomyopathy. *J Am Coll Cardiol*. 2008 Dec;52(25):2175–87..
- [3] Rachoin R, Saleh BA, Mansour B, Rachwan RJ, AlJaroudi W. Arrhythmogenic ventricular cardiomyopathy and sudden cardiac death: Left or right? *J Nucl Cardiol* [Internet]. 2016 Mar 18 [cited 2017 Apr 5]; Available from: <http://link.springer.com/10.1007/s12350-016-0465-8>
- [4] Haugaa KH, Basso C, Badano LP, Bucciarelli-Ducci C, Cardim N, Gaemperli O, et al. Comprehensive multi-modality imaging approach in arrhythmogenic cardiomyopathy—an expert consensus document of the European Association of Cardiovascular Imaging. *Eur Heart J – Cardiovasc Imaging*. 2017 Jan 9;jew229.
- [5] Sengupta PP, Huang Y-M, Bansal M, Ashrafi A, Fisher M, Shameer K, et al. Cognitive Machine-Learning Algorithm for Cardiac ImagingCLINICAL PERSPECTIVE: A Pilot Study for Differentiating Constrictive Pericarditis From Restrictive Cardiomyopathy. *Circ Cardiovasc Imaging*. 2016 Jun;9(6):e004330.
- [6] Prati G, Vitrella G, Allocca G, Muser D, Buttignoni SC, Piccoli G, et al. Right Ventricular Strain and Dyssynchrony Assessment in Arrhythmogenic Right Ventricular Cardiomyopathy: Cardiac Magnetic Resonance Feature-Tracking Study. *Circ Cardiovasc Imaging*. 2015 Nov;8(11):e003647; discussion e003647
- [7] Vigneault DM, te Riele ASJM, James CA, Zimmerman SL, Selwaness M, Murray B, et al. Right ventricular strain by MR quantitatively identifies regional dysfunction in patients with arrhythmogenic right ventricular cardiomyopathy: RV Strain in ARVC by MR Feature Tracking. *J Magn Reson Imaging*. 2016 May;43(5):1132–9.
- [8] Narula S, Shameer K, Salem Omar AM, Dudley JT, Sengupta PP. Machine-Learning Algorithms to Automate Morphological and Functional Assessments in 2D Echocardiography. *J Am Coll Cardiol*. 2016 Nov;68(21):2287–95.
- [9] Minsky M. Steps toward Artificial Intelligence. *Proc IRE*. 1961 Jan;49(1):8–30.
- [10] Hall M. Correlation-based feature selection for machine learning. [Hamilton, New Zealand]: University of Waikato, Department of Computer Science; 1998
- [11] Pearl J. Heuristics: intelligent search strategies for computer problem solving. Reading, Mass: Addison-Wesley Pub. Co; 1984. 382 p. (The Addison-Wesley series in artificial intelligence).
- [12] Dietterich null. Approximate Statistical Tests for Comparing Supervised Classification Learning Algorithms. *Neural Comput*. 1998 Sep 15;10(7):1895–923.
- [13] McKenna WJ, Thiene G, Nava A, Fontaliran F, Blomstrom-Lundqvist C, Fontaine G, et al. Diagnosis of arrhythmogenic right ventricular dysplasia/cardiomyopathy. Task Force of the Working Group Myocardial and Pericardial Disease of the European Society of Cardiology and of the Scientific Council on Cardiomyopathies of the International Society and Federation of Cardiology. *Br Heart J*. 1994 Mar;71(3):215–8.
- [14] Heermann P, Hedderich DM, Paul M, Schülke C, Kroeger JR, Baeßler B, et al. Biventricular myocardial strain analysis in patients with arrhythmogenic right ventricular cardiomyopathy (ARVC) using cardiovascular magnetic resonance feature tracking. *J Cardiovasc Magn Reson* [Internet]. 2014 Dec [cited 2017 Apr 5];16(1).
- [15] Taylor RJ, Moody WE, Umar F, Edwards NC, Taylor TJ, Stegemann B, et al. Myocardial strain measurement with feature-tracking cardiovascular magnetic resonance: normal values. *Eur Heart J – Cardiovasc Imaging*. 2015 Aug;16(8):871–81.

Address for correspondence.

Yolanda Vives Gilabert.
 Instituto ITACA Edificio 8G
 Universitat Politècnica de València
 Camino de Vera s/n
 46022, Valencia (Spain)
 yovigi@teleco.upv.es