

Model-based and Unsupervised Machine-learning Approaches for the Characterization of Responder Profiles for Cardiac Resynchronization Therapy

Marion Taconné¹, Virginie Le Rolle¹, Alban Gallard¹, Kimi P Owashi¹, Adrien al Wazzan¹, Elena Galli¹, Jens-Uwe Voigt², Jurgen Duchenne², Otto Smiseth³, Erwan Donal¹, Alfredo Hernandez¹

¹ Univ Rennes, CHU Rennes, Inserm, LTSI-UMR 1099, F-35000, Rennes, France; ² Department of Cardiovascular Disease, KU Leuven, Leuven, Belgium; ³ Center for Cardiological Innovation and Department of Cardiology, Oslo University Hospital, Oslo, Norway

Abstract

The objective of this study is to improve the interpretability of a previous unsupervised clustering analysis of the CRT response through a physiological model-based approach. The developed clustering approach was applied on 250 CRT candidates based on clinical, original and classical echocardiographic features. Patient-specific computational models were proposed for patients associated of each cluster barycenter in order to provide an explainable analysis in relation with physiological mechanisms. Five phenogroups were identified from the clustering approach with response rates ranging from 50% to 92.7%. Concerning the model-based approach, a match was observed between the 16 experimental and simulated myocardial strain curves pattern with a mean RMSE of 3.97% (± 1.74) on the five patients. Moreover, the identified model parameters provide us information about the mecano-electrical coupling and tissue properties. The gain of information provides by the parameters model identification, added to the clinical and classical echocardiographic features is promising for an understanding of LV mechanical dyssynchrony and the identification of patients suitable for CRT.

1. Introduction

Despite the recognized clinical benefits of Cardiac Resynchronization Therapy (CRT) [1] for patients suffering from systolic heart failure (HF) and a bundle branch block, around 30% of implanted patients do not respond to this therapy. An important research subject in this field is thus to better identify those patients that may benefit from CRT, before the implantation of the therapy. To this end, along with the exploitation of electrophysiological data, our team has proposed the analysis of cardiac mechanical dyssynchrony, by processing regional cardiac strain curves observed through pre-operative echocardiography

[2–4]. We have also proposed a multiparametric clustering method, integrating clinical and echocardiographic data, to phenogroup 250 CRT candidates, with respect to their response to therapy and outcome [5]. This clustering approach, led to the identification of specific subgroups of CRT response and provided information about how cardiac regional deformations, quantified through strain integrals, may be related to CRT response.

Although the automatic quantitative analysis of longitudinal strain curves provides a good discriminating value between clusters, the physiological interpretation of the modifications observed in regional strain morphologies remain a major challenge, since these strain curves reflect complex mechanisms associated with electrical conduction delay, mechanical cardiac activity and inter-regional interactions. In this context, physiological model-based methods appear as a promising tool to increase the interpretability of the analysis, since most of the parameters provide a direct physiological meaning [6, 7].

The objective of this paper is to propose a method to improve the interpretability of the unsupervised clustering method proposed in our previous works, through a patient-specific physiological model-based approach. Strain features and clinical data from 250 HF patients were analyzed. Model parameter identification for patients of each cluster (CRT responder or not) were performed and parameters reflecting physiological mechanisms were analyzed.

2. Methods

2.1. Clinical data and echocardiography

The prospective database includes 250 patients (from different centers in Europe) who were eligible on the basis of clinical grounds for CRT implantation. The study was carried out in accordance with the principles outlined in the Declaration of Helsinki and was approved by the local ethical committee of each center.

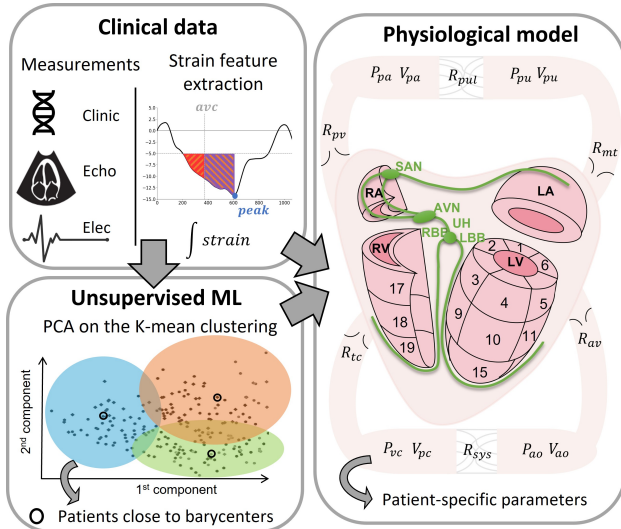


Figure 1: Overview of the proposed methods: a) Database and feature extraction, b) Unsupervised Machine Learning (ML): clustering visualized by a PCA and c) the patient-specific cardiovascular model.

Clinical, electrocardiographic, and echocardiographic data were collected. Before implantation of the CRT device, patients were imaged by transthoracic echocardiography (ViVid E9, S70 or E95, General Electric Healthcare, Horten, Norway) in order to extract regional myocardial strain curves.

All patients received a follow-up at-rest echocardiography at 6 months. Responders were defined as having a $\pm 15\%$ decrease in LV end-systolic volume at the 6-month follow-up, compared with baseline.

2.2. Unsupervised clustering for CRT response estimation

All the details of the feature extraction and clustering phases are described in a previous paper of our team [5]. Feature extraction was performed from clinical, electrocardiographic, and echocardiographic data, leading to a set of 70 features per patient. Concerning specifically echocardiography, integrals were calculated for strain signals of each segment from the beginning of QRS to the strain peak and to the aortic valve closure. The set of all features was clustered by applying the K -Means method [8] using the Sklearn [9] library. The optimal number of clusters was determined using a Silhouette score. In order to visualize clustering results, a principal component analysis (PCA) was performed.

2.3. Physiological model description

Four main sub-models, based on previous works of our team were coupled (Fig.1 *right panel*): i) cardiac electrical system, ii) right and left atria, iii) multi segment representation of the right and left ventricle and iv) systemic and pulmonary circulations. This model has previously been

described in detail in [7] and has been validated on data from 3 LBBB HF patients [6].

Cardiac electrical system: The cardiac electrical activity was represented by a set of interconnected automata, adapted from [10].

Right and left atria: The right and left atrial pressures were defined as linear functions of instantaneous volumes and elastances representing the elastic properties.

Right and left ventricle: The left ventricle wall was divided into 16 segments according to the standardized segmentation of the AHA. The right ventricle wall was divided into three segments. Each segment s can be separated into active and passive components: $T_s = T_{s,pass} + T_{s,act}$. $T_{s,pass}$ and $T_{s,act}$ are described by non-linear relations that include K_{pass} and K_{act} , which are parameters related to passive stiffness and myofiber contractility. An electro-mechanical driving function (EMDF) $f_{a,s}$ was defined to represent, in a simplified manner, the complex processes involved in the electro-mechanical coupling at the tissue-level [10].

Systemic and pulmonary circulations: The cardiovascular system model integrated the pulmonary arteries, capillaries and veins, and the systemic arteries and veins. Each compartment is represented by resistance and compliance relations.

2.4. Parameter identification

The identification process was applied to the patient located at the smallest distance from the barycenter of each cluster. Based on our previous study [6], a set of parameters are selected for patient-specific model identification. This identification was implemented with an evolutionary algorithm (EA) [11]. This type of algorithm consists of making evolve a population of a set of parameter values X in order to minimize an error function J . The evolution of the solution is performed through the application of a set of transformations of the set X , mimicking biological evolution processes such as selection, crossover and mutation. Function J is defined as the error between the 16 strain curves from experimental measurements (2-, 3- and 4CH views) and the corresponding simulated strain curves, obtained by the model using the set of parameters X :

$$J_{error} = \sum_{s=1}^{16} J_s \quad (1)$$

$$J_s = \frac{1}{T_c} \sum_{t_e=0}^{T_c-1} |\varepsilon_s^{exp}(t_e) - \varepsilon_s^{model}(t_e, X)| + |\varepsilon_s^{exp} - \varepsilon_{s,min}^{model}| \quad (2)$$

where ε_s^{exp} and ε_s^{model} are the myocardial strain signals obtained from available clinical data and simulated outputs respectively.

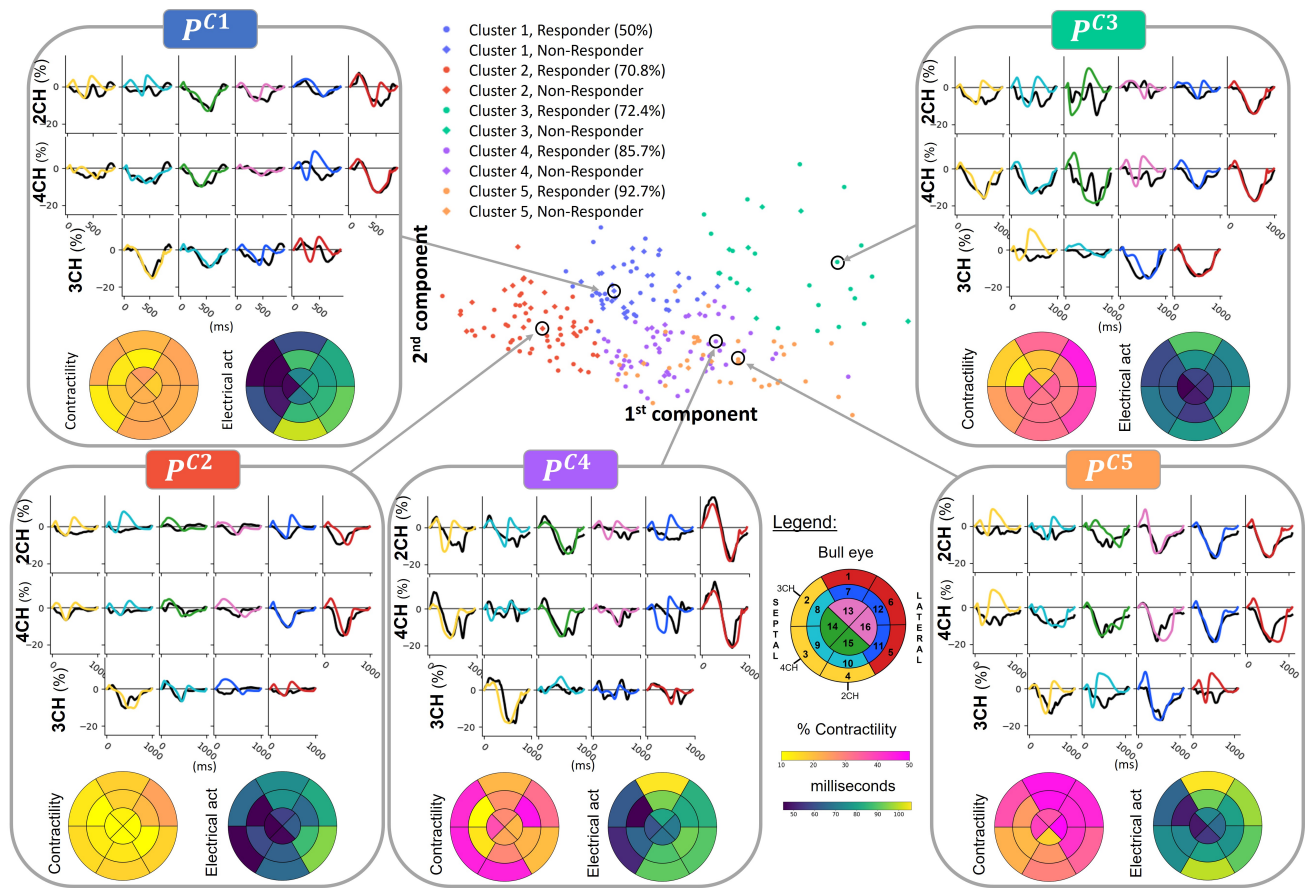


Figure 2: Simulation results of the 5 patients (circled in black) of each cluster visualized by the two first components of a PCA (center). Experimental (black) strain curves of the 2-, 3- and 4CHamber view and their equivalent simulated strain curves (colored) with the bull eyes of identified parameters: contractility and electrical activation delay for each of the 16 segments. The color strain legend of the 3 views is gather on a bull eye (Basal: 1.anterior, 2 anteroseptal, 3.inferoseptal, 4.inferior, 5.inferolateral, 6.anterolateral; Mid: 7.anterior, 8 anteroseptal, 9.inferoseptal, 10.inferior, 11.inferolateral, 12.anterolateral; Apical: 13.anterior, 14.septal, 15.inferior, 16.lateral)

Cluster	Resp rate	Significant features
PC^1	50%	No septal flash and no apical rocking, low lateral integral, low integral difference between lateral and septal wall, low electrical delay and contractility
PC^2	70.8%	High Global Longitudinal Strain, low Global Constructive Work, low septal integral, low mean strain avc, low electrical delay and contractility
PC^3	72.4%	High contractility, high and medium septal integrals, low mean strain peak, low electrical delay
PC^4	85.7	High electrical delay and contractility, high septal and mean minimum strain time, medium septal efficiency (E), female gender
PC^5	92.7	High electrical delay and contractility, high lateral integrals and integral difference

Table 1: Response rate for each identified cluster and their most significant features, colored by relative value: High (green), medium (orange) and low (red)

3. Results

Clustering: The optimal number of clusters k obtained using the silhouette method was $k = 5$ [5]. Tab 1 gathers the responder rate and the most significant features of each cluster. The center of Figure 2 shows the clustering results

of [5]. The two first principal components of the PCA analysis are represented with the CRT response of each patient. In most phenogroups, the quantitative analysis of strain curves, especially associated with the lateral wall, was more discriminative than markers usually associated with response to CRT (apical rocking and septal flash).

Identifications and simulations: From the identification process, we obtained patient-specific simulated strain curves for the 5 patients associated with each cluster barycenter (circles in Fig.2) on the 16 LV segments compared to the experimental curves. Although, for some patients the strain morphologies are not completely reproduced for all the 16 curves, a close match was observed between experimental and simulated curves pattern. The mean RMSE on the five patients is $3.97\% (\pm 1.74)$. Figure 2 also shows the identified contractility and electrical parameters identified for each patient and each LV cardiac segment, represented through bull-eyes diagrams. Clusters 4 and 5 show an early shortening of the septal wall. The patients of cluster 3, 4 and 5 (responders), which are associated with a more elevated rate of responders, show higher contractility values in all their segments with a mean contractility value of 29.2%, 29.2% and 35.1% respectively, compared to the patients in clusters 1 and 2 (non-responders) with 20.5% and 14.2%. LV electrical delays, associated with clusters 4 and 5, are also slightly

higher than the other clusters with a maximum electrical delay of 106 ms for these two cluster against 101, 95 and 90 ms for the cluster 1, 2 and 3 respectively.

4. Discussion

The main contribution of this work concerns the proposal of an original pipeline that combines two different approaches: an unsupervised clustering and the identification of a patient-specific physiological model. This study is a first step to the use of identified parameters in a CRT-response prediction process. The model-based approach provides explainable results, since patient-specific parameters provide a direct physiological interpretation.

Characterization of responder profiles: Results from the clustering phase provides groups of different response rates, which could be analyzed using identified parameters. Groups with below-average rates (cluster 1 and 2) are globally associated with low strain integral values and reduced myocardial contractility, as illustrated on bull-eyes obtained from the model. The other clusters (3, 4 and 5) present higher strain integral values and more elevated contractility. Concerning specifically cluster 5, that could be described as the super responder group, strain morphology shows a typical LBBB activation pattern with early stretching of lateral wall and early shortening of septal wall. This particular contraction patterns could be explained by model parameters, which show a preserved contractility and elevated electrical activation delays, known as pure electrical dyssynchrony.

Integrating model-based and clustering approaches: The proposed approach improves the interpretability of the clustering analysis by integrating knowledge of physiological mechanisms, related to cardiac contraction, to a phenotyping of HF patients. The model-based approach provides additional information on the regional electrical and mechanical LV function. The area of reduced contractility, identified by the model, is of primary importance because the localization of a potential scar (especially in lateral wall) could be associated with a lower rate of CRT response because of the inefficient stimulation of a necrotic territory. The global approach represents a step forward in the development of personalized LV modelling in the field of CRT, as it can help to disclose the intrinsic complexity of LV mechanics in CRT candidates.

Limitations: Several assumptions have been made to propose the computational model, including simplifications concerning the electrical and mechanical behaviors such as fiber, torsion or a complete mechanical continuity.

5. Conclusion

An original method combining unsupervised and physiological model-based machine-learning methods was pro-

posed for the characterization of responder profiles for cardiac resynchronization therapy. The proposed approach improves the interpretability of the clustering analysis for HF patient phenotyping, since the proposed model explicitly integrates meaningful physiological knowledge. Future work will be focused on the parameter identification and evaluation on the whole dataset, as well as on the proposal of a decision support system for therapy planning based on a patient-specific characterization.

References

- [1] Bleumink GS, et al. Quantifying the heart failure epidemic: Prevalence, incidence rate, lifetime risk and prognosis of heart failure - The Rotterdam Study. *European Heart Journal* 2004;25(18):1614–1619. ISSN 0195668X.
- [2] Bernard A, et al. Impact of cardiac resynchronization therapy on left ventricular mechanics: Understanding the response through a new quantitative approach based on longitudinal strain integrals. *Journal of the American Society of Echocardiography* 2015;28(6):700–708. ISSN 10976795.
- [3] Gallard A, Galli E, et al. Echocardiographic view and feature selection for the estimation of the response to CRT. *Plos One* 2021;16(6):e0252857.
- [4] Hubert A, Gallard A, et al. Left ventricular strain for predicting the response to cardiac resynchronization therapy: twomethods for one question. *European Heart Journal Cardiovascular Imaging* 2021;1–8. ISSN 2047-2404.
- [5] Gallard A, et al. Characterization of Responder Profiles for Cardiac Resynchronization Therapy through Unsupervised Clustering of Clinical and Strain Data. *Journal of the American Society of Echocardiography* 2021;34(5):483–493. ISSN 10976795.
- [6] Owashi K, Taconné M, et al. Desynchronization strain patterns and contractility in left bundle branch block through computer model simulation. *Journal of Cardiovascular Development and Disease* 2022;9(2):53.
- [7] Owashi K, Galli E, et al. Patients-Specific Analysis of Myocardial Strains in Left Bundle Branch Block Based on Computational Models. 2020 Computing in Cardiology Conference CinC 2020;47:1–4. ISSN 2325887X.
- [8] Lloyd SP. Least Squares Quantization in PCM. *IEEE Transactions on Information Theory* 1982;28(2):129–137.
- [9] Pedregosa F, et al. Scikit-learn: Machine learning in python. *the Journal of machine Learning research* 2011; 12:2825–2830.
- [10] Le Rolle V, Hernández AI, et al. Model-based analysis of myocardial strain data acquired by tissue Doppler imaging. *Artificial Intelligence in Medicine* November 2008; 44(3):201–219. ISSN 0933-3657.
- [11] Goldberg DE, et al. Genetic Algorithms and Machine Learning 1988;3(2):95–99. ISSN 15730565.

Address for correspondence:

Marion Taconné (marion.taconné@univ-rennes1.fr)
LTSI-Université de Rennes 1, 35042 Rennes, France