

Master thesis on Computational Biomedical Engineering

Universitat Pompeu Fabra

Prospective clinical study for the pre-operative identification of sites of origin in outflow tract ventricular arrhythmias with electrophysiological simulations

Itziar Busquets Sáez

Supervisor: Óscar Cámara Rey

Co-supervisor: Diego Penela Maceda, Rubén Doste

July 2021





Master thesis on Computational Biomedical Engineering

Universitat Pompeu Fabra

Prospective clinical study for the pre-operative identification of sites of origin in outflow tract ventricular arrhythmias with electrophysiological simulations

Student: Itziar Busquets Sáez

Supervisor: Óscar Cámara Rey

Co-supervisor: Diego Penela Maceda, Rubén Doste

July 2021





# Table of Contents

1.	Introduction .....	1
1.1	Cardiac anatomy and physiology.....	1
1.2	Clinical problem background.....	3
1.2.1	Cardiomyocyte orientation .....	3
1.2.2	Outflow tracts anatomy .....	4
1.3	ECG analysis techniques .....	7
1.4	ECG specific features to determine the site of origin .....	8
1.5	Radiofrequency ablation procedure .....	11
1.6	State of the art of computational pace-mapping techniques .....	12
1.7	Objectives .....	15
2	Methods.....	16
2.1	Computational pipeline.....	16
2.2	Pseudo-ECG based in-silico pace-mapping pipeline.....	17
2.2.1	Modelling dataset .....	17
2.2.2	Building the biventricular mesh from cardiac CT images.....	18
2.2.3	Mesh labelling and determining fiber orientation .....	19
2.2.4	Sites of origin considered in this work .....	21
2.2.5	Simulations input files .....	22
2.3	Machine learning approach.....	25
2.3.1	Machine learning algorithms evaluation datasets.....	25
2.4	Prospective implementation of the pseudo-ECG based in-silico paced mapping and the machine learning algorithms. ....	28
3	Results .....	30
3.1	Pseudo-ECG based in-silico paced mapping .....	30
3.1.2	Qualitative pseudo-ECG validation with clinical literature .....	34
3.2	Machine learning algorithms for differentiating between LVOT or RVOT origin. ....	38
3.2.1	Results evaluating on the ECGs used for the modelling approach.....	38
3.2.2	Results evaluating on the new extracted ECG datasets.....	39
3.3	Prospective evaluation of the machine learning and the in-silico pace-mapping pseudo-ECG based pipelines. ....	40
4	Discussion .....	41
5	Conclusions .....	43

6	Bibliography.....	45
	Appendix: Simulated pseudo-ECGs.....	50



## Acknowledgements

I would first like to thank my supervisor, Oscar Camara, for his support and guidance during all these months. Also, I want to thank my co-supervisor Diego Penela for his help in this project.

Special thanks to Rubén Doste and Nuria Moya, for always answering my doubts and sharing all the material with me.

I would like acknowledge my collaborators Hospital Teknon-Quirón, and Barcelona Supercomputing Center, for giving me the opportunity to work with them and giving me the tools to develop this project.

Also, I want to thank Guillermo Jiménez, for his counsel and help when analyzing and processing the ECG signals.

Finally, I would like to thank my PhySense classmates, for being so supportive and always willing to help during this whole year.



## Abstract

Outflow tract ventricular arrhythmias (OTVAs) are a subset of ventricular tachycardias which occur in patients with no apparent structural heart disease. Although they are classically benign, in some patients they can be highly symptomatic in patients resistant to drug therapy. In those cases, radiofrequency ablation therapy is usually prescribed. During this intervention, cardiac electroanatomic mapping is performed, by recording the patient's electrical myocardium activity, and performing pace mapping, which consists in stimulating specific heart regions to provoke an arrhythmia.

Doste et al.<sup>1</sup> developed a complete modelling pipeline for non-invasively diagnosing the site of origin in patients with OTVAs. Later, Moya<sup>2</sup> adapted this work by building a more user friendly pipeline by lowering the computational cost. These highly detailed frameworks acquired promising results, although they have not been tested in a clinical environment. In the present work, the aforementioned pipelines are going to be analyzed, and some of the steps are redefined in order to match the requirements to be implemented in the clinical workflow, such as the definition and localization of the sites of origin, cardiac segmentation time shortening, or the quantitative electrocardiogram (ECG) analysis.

Computational models are used to generate patient specific pseudo-ECGs from different stimulated sites of origin. Also, machine learning tools are used to classify patient real ECG OTVA cases from real-world clinical databases. Both approaches are compared and analyzed. The detailed electrophysiological simulations are highly time consuming, so they are not easily implementable into the clinical daily basis. Thus, it has been concluded that machine learning methods are more suitable tools to be implemented in the clinical workflow in terms of acquisition time and necessary computational resources. In this work, some machine learning based solutions combining real and simulated ECGs are presented as an alternative to non-invasively diagnose the ventricular chamber origin of OTVAs.

**Keywords:** idiopathic; outflow tract ventricular arrhythmia; electrocardiogram; premature ventricular contraction; electrophysiological simulation; site of origin; machine learning.



# 1. Introduction

Idiopathic ventricular tachycardia (VT) is recognized as a ventricular arrhythmia (VA) in patients without an apparent structural heart disease and accounts for approximately 10% of all patients referred for an evaluation of VT<sup>3</sup>. Idiopathic VTs are most often presented as premature ventricular contractions (PVCs) with a focal origin, and they commonly occur in the outflow tract regions<sup>4</sup>, which are a portion of the left or right ventricle of the heart from which the blood flows in order to enter to the aortic or pulmonary arteries. The activation from a PVC origin propagates through the healthy myocardium, and exhibit characteristic electrocardiograms (ECGs) which can be analyzed in order to have an approximation of where the ectopic foci of the arrhythmia is located<sup>4</sup>.

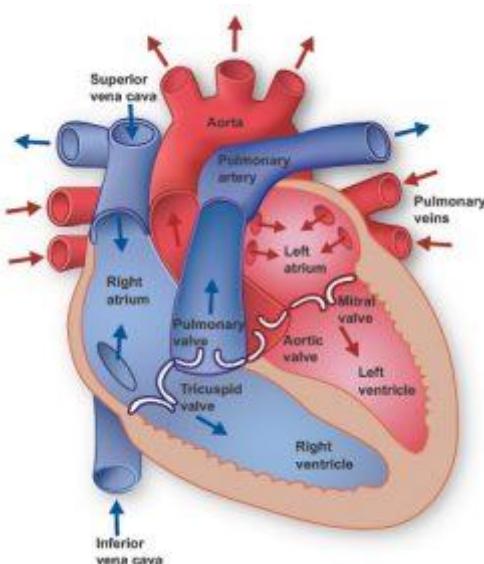
Treatment for drug resistant arrhythmic patients consists in radiofrequency catheter ablation, which requires the previous identification of the site of origin (SOO) of the arrhythmia by performing invasive catheter electroanatomical mapping. The complex anatomy of these OTs makes the identification of the exact SOO of these arrhythmias very challenging, since the OTs are anatomically close to each other.

## 1.1 Cardiac anatomy and physiology

The heart is a muscular pump that collects deoxygenated blood from all parts of the body, which is carried to the lungs to be oxygenated and releases carbon dioxide. In order to survive, tissues need a continuous oxygen supply, nutrients and metabolic waste products need to be removed.

The heart is divided into four chambers: the upper left and right atrium, and the lower left and right ventricles. The right atrium and ventricle are separated by the tricuspid valve, whereas the left atrium is isolated from the left ventricle by the mitral valve. Both ventricles are separated by the interventricular septum, which is a myocardial wall.

As observed in Figure 1, the pulmonary artery is inferiorly bounded by the pulmonary valve, arising from the right ventricle, and the aorta artery arises from the left ventricle after the aortic valve. The pulmonary valve is positioned anterior, leftwards, and 1-2cm upper relatively to the aortic root<sup>5</sup>.



*Figure 1. Heart basic anatomy. Obtained from <https://www.texasheart.org>*

The superior and inferior cava veins return oxygen-deprived blood to the right atrium, which passes to the right ventricle during the filling phase, and is sent to the lungs through the right ventricle outflow tract during ejection. Then, the pulmonary veins return oxygenated blood from the lungs to the left side of the heart so it will be pumped through the whole body. For this reason, the left ventricle wall is thicker than the right ventricle wall, since it needs generate higher pressure to eject blood to the whole body.

The electrical conduction system of the heart is represented in Figure 2. The heart rhythm is normally controlled by the sinoatrial (SA) node, which acts as a natural pacemaker. It produces electrical impulses that trigger the normal heartbeat. The SA node sends a depolarization wave through the atrium until arriving to the atrioventricular node, where the impulse is delayed in order to give atrial contraction enough time to fill the ventricles<sup>6</sup>. Then, it is transmitted rapidly along the His-Purkinje system and depolarizes the ventricles in a systematic way. These impulses cause them to contract and pump blood to

the lungs and the rest of the body. Cardiac arrhythmias are the abnormalities or perturbations in the normal activation or beating of the heart myocardium<sup>7</sup>.

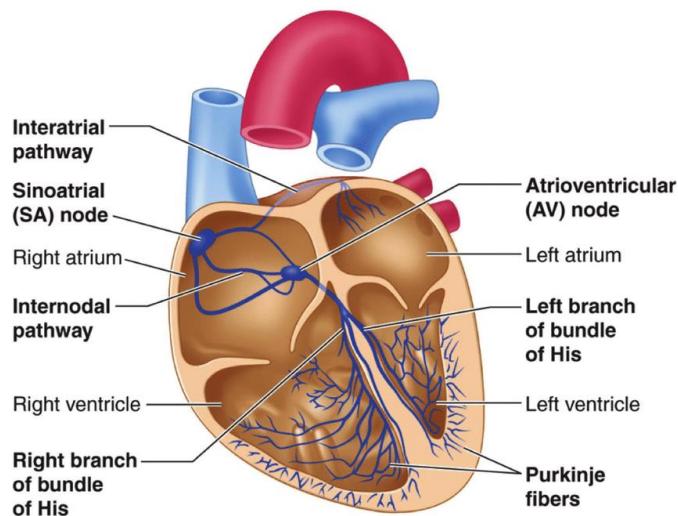


Figure 2. Heart electrical conduction system. Obtained from <sup>40</sup>.

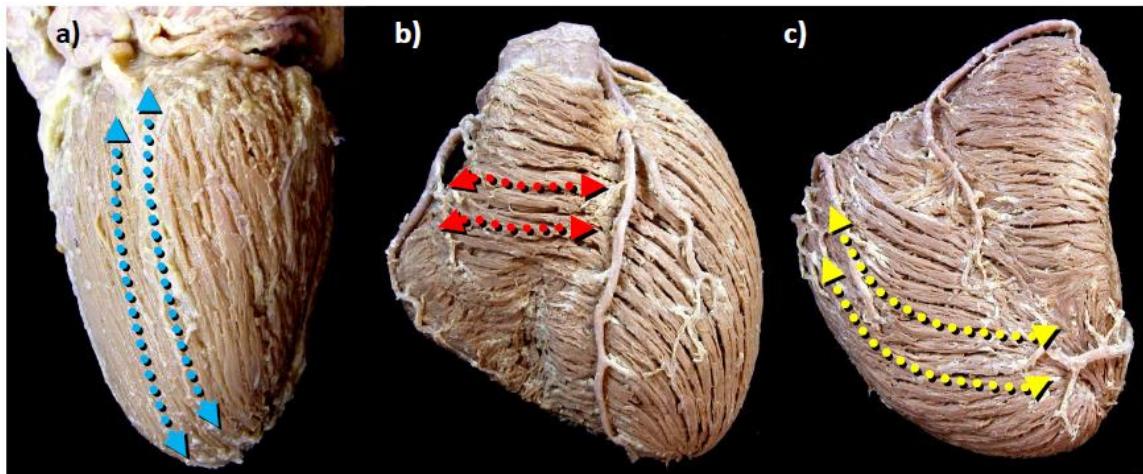
## 1.2 Clinical problem background

Understanding cardiac anatomy is essential for the diagnosis and treatment of arrhythmias. Recent investigations have unraveled anatomic features, architectural aspects, and histologic details of certain components of the heart, of interest to understanding the substrates of tachycardias and their ablation<sup>8</sup>.

### 1.2.1 Cardiomyocyte orientation

The cardiomyocyte is the individual contractile cell of the myocardium. These cardiomyocytes are arranged in a way that its distinct preferred direction is visually easy to identify<sup>9</sup>. The term fibers describes the macroscopic display of cardiomyocyte arrangement. Cardiomyocyte organization influences cardiac electrical propagation and mechanical contraction, since its fiber orientation defines the preferable electrical conduction pathway<sup>1,10</sup>.

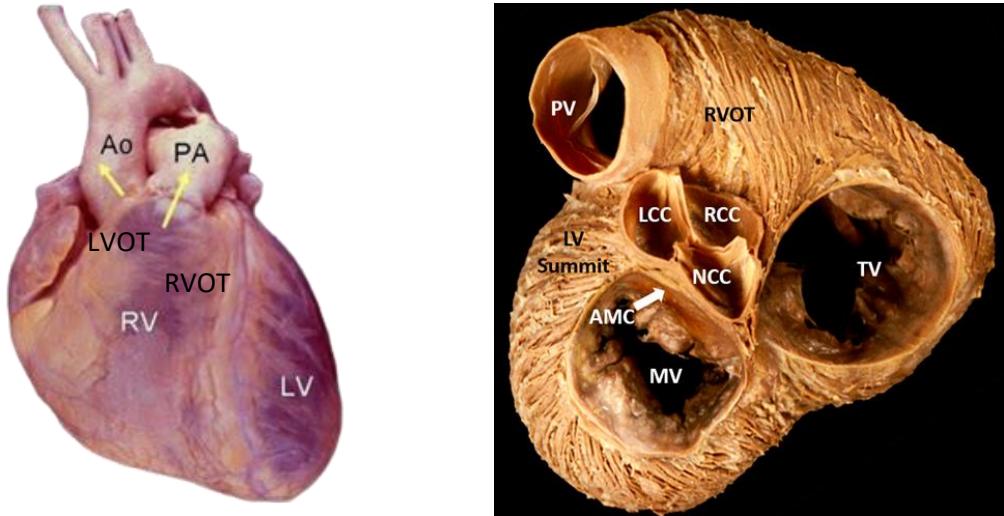
As observed in Figure 3, the complex fiber orientation is heterogeneous. Three distinct directions are usually defined by anatomists: a longitudinal direction (parallel to the apico-basal direction), a circumferential direction (parallel to the short axis direction of the heart), and an oblique direction (any other angle)<sup>1,8</sup>.



*Figure 3. Fibers oriented in the longitudinal direction (a), in the circumferential direction (b), and in an oblique direction (c). Image courtesy of Damián Sánchez Quintana.*

### 1.2.2 Outflow tracts anatomy

The close anatomic proximity of the OTs and the complex overlay of the RVOT with the LVOT makes clear separation and the application of morphological characteristics at each site challenging<sup>5,8,11–13</sup>. The right ventricular outflow tract (RVOT) passes anteriorly over the left ventricular outflow tract (LVOT), which is directed rightward, resulting in a crossover relationship between the two tracts, in fact most of the RVOT is to the left of the LVOT (Figure 4). Therefore, the posterior part of the RVOT is adjacent to the anterior part of the LVOT. It is superiorly bounded by the pulmonary valve, and caudally by the tricuspid valve<sup>8</sup>. The anterior part of the RVOT is the right ventricular free wall. Three semilunar leaflets are distinguished in the RVOT: the anterior, left and right pulmonary cusps. A more clear representation of the outflow tracts anatomy and orientation is observed in Figure 5.



*Figure 4. In the left image, the right and left ventricles (RV and LV), the right and left outflow tracts (RVOT and LVOT) and the aortic and pulmonary arteries (Ao and PA) are displayed. An anterior and left sided orientation of the RVOT from the LVOT is observed. Image obtained from <https://drsvenkatesan.com>. Right image: Short axis view from atria side. Pulmonary and aortic valves displayed. MV: mitral valve, TV: tricuspid valve, AMC: atrioventricular continuity, LV summit: left ventricular summit, LCC: left coronary cusp, RCC: right coronary cusp, NCC: non-coronary cusp , PV: pulmonary valve. Image obtained from<sup>8</sup>.*

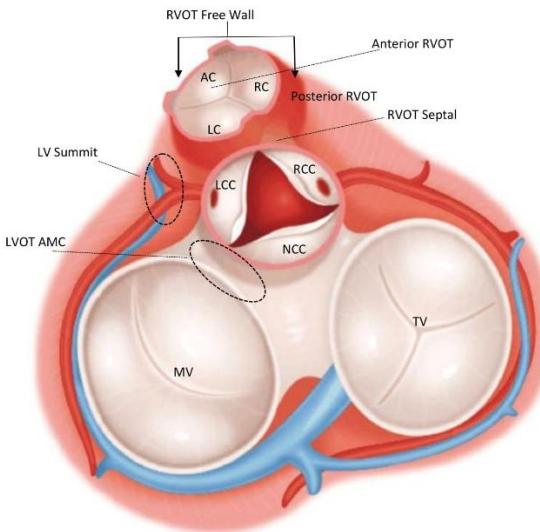
The LVOT is superiorly bounded by the aortic valve and is characterized by an aortomitral continuity (AMC). The three cusps of the aortic valve, as seen in Figure 4, are the left coronary cusps (LCC), the right coronary cusp (RCC) and the non-coronary cusp (NCC). It is important to highlight that the more posterior, septal part of the RVOT is adjacent to the LCC and RCC, while the NCC is next to the interventricular septum.

In the more superior part of the left ventricle is found the left ventricular summit. It is a triangular region located within the bifurcation of the left anterior descending and the left circumflex coronary arteries<sup>14</sup>.

The plane of the pulmonary valve is 1-2 cm higher than the aortic valve and nearly horizontal, whereas that of the aortic valve is tilted lower and at an angle of at least 45° from the median plane<sup>8</sup>.

Originating from the ventricles of the heart, a premature ventricular complex (PVC) causes an extra, or abnormal, heartbeat that occurs earlier than it should. Ventricular

tachycardia (VT) manifested by three or more consecutive PVCs are seen at a rate of 100 beats per minute (bpm) or higher<sup>15</sup>.



*Figure 5. Sub anatomic structures from the outflow tracts. LVOT: left ventricle outflow tract, RVOT: right ventricle outflow tract, MV: mitral valve, TV: tricuspid valve, AC: anterior cusp, RC: right cusp, LC: left cusp, LCC: left coronary cusp, RCC: right coronary cusp, NCC: non coronary cusp. Image obtained from<sup>16</sup>.*

The left and right ventricle outflow tracts are common sites of origin (SOO) of VT, with clinical presentation including single extrasystoles, repetitive nonsustained or sustained VT, and triggered ventricular fibrillation<sup>5</sup>. Idiopathic VTs, defined as those occurring in patients without heart structural disease, are usually originated in the outflow tract regions<sup>14,17,18</sup>. These ones most commonly arise from the RVOT (up to 80% of them<sup>19</sup>) than from the LVOT<sup>13</sup>. However, the availability of advanced electroanatomical mapping techniques have allowed the access to more complex cardiac structures, and then LVOT origin is currently diagnosed in nearly 50% of the cases<sup>11,19</sup>. Although the arrhythmogenic focus is often endocardial, both intramural and epicardial sites of origin are described<sup>15</sup>. OTVAs are classically benign, but some patients can be highly symptomatic and resistant to medical therapy. Moreover, frequent ectopy can progress to a PVC-induced cardiomyopathy, impaired quality of life, incessant VT and sudden cardiac death<sup>15</sup>. In these cases, catheter ablation (CA) can be considered. CA has a low procedural complication risk (<1%) that is outweighed generally, by a high success rate with long-term probability of cure. Before CA, it is essential to have an appreciation of the OTVA

origin. Distinguishing between right or left OT SOO preoperatively is useful in terms of guiding access for the ablation, counseling patients on specific risks and may even involve referral to a specialist unit if localized to a high-risk or difficult position<sup>20</sup>.

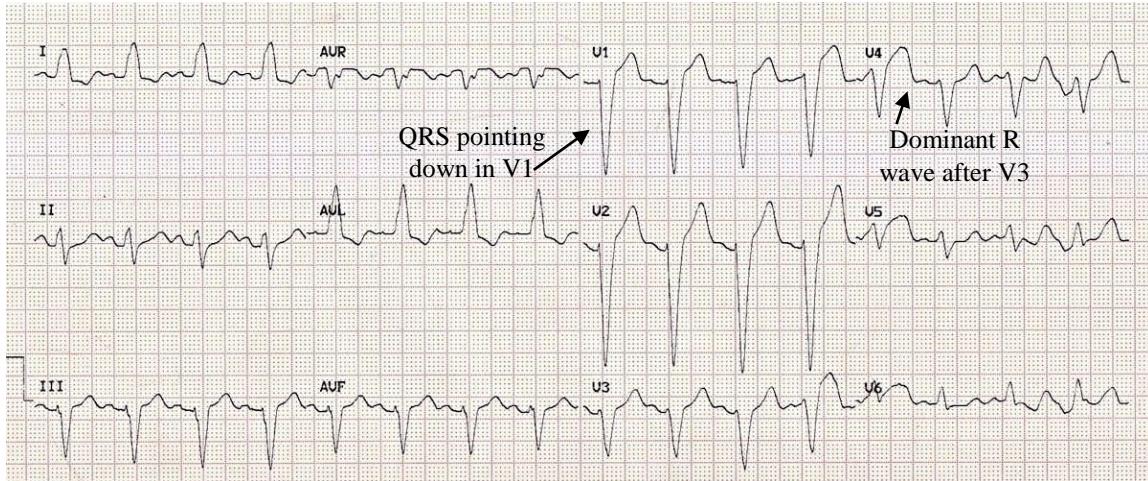
### 1.3 ECG analysis techniques

Electrocardiographic records (ECG) can be analyzed for determining features to distinguish between a left or right origin of the OTVA. However, it requires advanced skills in ECG analysis since their ECG characteristics are similar because of their anatomical proximity<sup>5,11,15</sup>.

Due to its anterior location, VTs originating from the RVOT usually present a left bundle branch block (LBBB) configuration (a downwardly deflected QRS complex in V1) with rare exceptions<sup>2,5,15</sup>, while more posterior structures, far from the anterior chest, will produce right bundle branch block pattern (RBBB). OTVAs arising from the LVOT may have a RBBB as well as an atypical LBBB morphology depending on the exact SOO in the LVOT. For example, the RCC is immediately posterior the septal RVOT, and usually displays a LBBB morphology.

Regarding the 6 precordial leads (V1 to V6), the term precordial transition represents the first lead with a dominant R wave<sup>2</sup>. A precordial transition occurring before V3 is associated with a LVOT origin, and a later precordial transition than V3 is very likely to have an origin in the RVOT<sup>2,4</sup>. However, when a LBBB morphology and a V3 precordial transition are found, it is uncertain to differentiate a RVOT against a LVOT tachycardia origin<sup>4,11,15</sup>.

Both of these features can be observed in Figure 6. A LBBB is observed since the QRS complex points down in V1, and the precordial transition occurs in V4. Both of these characteristics suggest a RVOT origin, according to literature.



*Figure 6. 12-lead ECG recording, where LBBB can be observed by looking at V1 (QRS pointing down), and a late precordial transition (occurring in V4), suggesting RVOT origin. LBBB: left ventricular bundle branch block, RVOT: right ventricular outflow tract. Image obtained from <http://acuclinic.com.au>*

## 1.4 ECG specific features to determine the site of origin

Some ECG characteristics found in the literature are useful in order to distinguish a left against a right ventricular OT VT origin considering different specific sites. These characteristics can be observed in Figure 7<sup>20</sup>. Some features from different SOO in the RVOT are the following.

**Pulmonary lateral (PL):** due to its anterior location, it usually presents a later precordial transition than septal origins (>V4). Usually present notching on the QRS complex in lead I and in the inferior leads<sup>4,15,20</sup>. QRS duration in inferior leads is longer than in PVC originating from the septum, and also deeper S waves in the inferior leads<sup>4,20</sup>.

**Septal origin:** smaller S wave amplitudes in leads V1 to V3 than from the PL and taller R wave amplitude<sup>4</sup>. Leads II, III and aVF tend to be taller and narrower for septal compared versus PL sites within the RVOT<sup>14</sup>. Moving from more posterior to anterior direction, lead I is more positive in posteroseptal origins, and it decreases until being negative in anteroseptal origins.

- **Septal posterior (SPP):** usually shows a larger R wave in lead I compared to a more anterior origin, and presents notching in the inferior leads<sup>4,20</sup>.

- **Septal medium (SPM):** in lead I, an intermediate not positive nor negative qrs, and a more positive R wave in V1 but still LBBB<sup>14</sup>.
- **Septal anterior (SPA):** LBBB pattern, shows an rS (small r wave and wider S wave) morphology in V1, a dominant R wave in V6, and a deeper S wave in lead I<sup>4,20</sup>.

On the other hand, OTVAs arising from specific sites of the LVOT are listed below.

**Aortomitral continuity (AMC):** usually presents an absence of S waves in lead V6, a RBBB pattern, and an rS pattern in lead I and qR pattern in lead V1<sup>4,20</sup>. Also, it usually presents a “rebound transition pattern” that is defined as a negative QRS polarity in V2 but positive in V1 and V3<sup>2</sup>.

**Right coronary cusp (RCC):** as it is positioned posterior to the septal RVOT, they display similar morphologies. It usually exhibits a LBBB pattern with an rS wave in V1 and V2 (small R wave and dominant S wave), and transition in V3<sup>4,20</sup>. It also presents more positive QRS complex in lead I than LCC origins<sup>2</sup>.

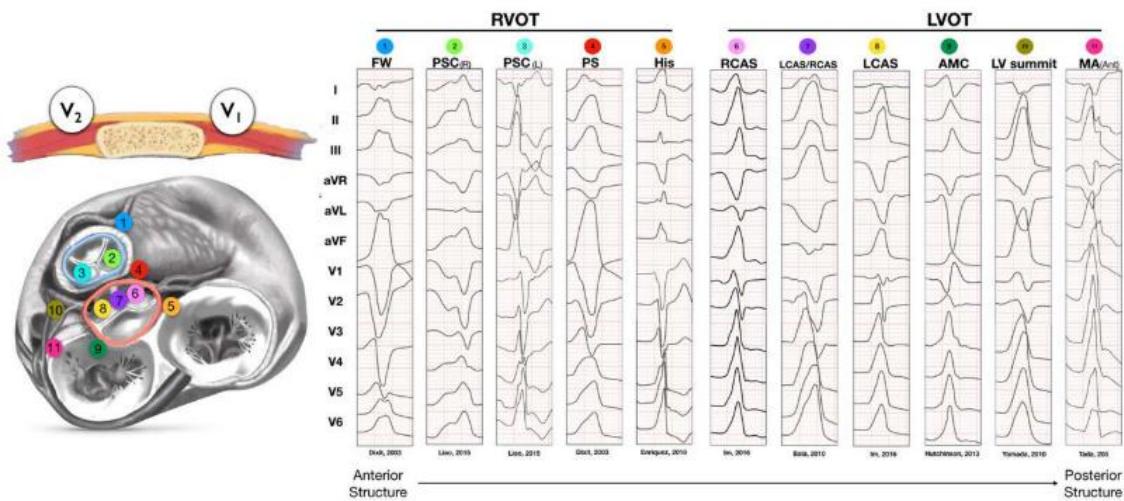
**Left coronary cusp (LCC):** early precordial transition in V1 or V2, earlier than RCC origins. R waves in V1 and V2 are significantly tall and broad. It also has a characteristic multiphasic M or W pattern in V1<sup>4,20</sup>. More negative QRS in lead I than RCC, increase R wave amplitude in the inferior leads. Always exhibits a RBBB pattern<sup>19</sup>.

**Left and right coronary cusps commissure (LCCRCC):** it frequently exhibits a QS morphology in V1 with prominent notching in the downward deflection and precordial transition in V3<sup>14,20</sup>. Rarely exhibits an RBBB pattern.

**Epicardium (LV summit):** usually presents a LBBB morphology. Also a V2 pattern break, which means that the R wave in V2 is less positive than in V1 and V3. Other significant features found in the literature are dominant R wave in V6 and an rS morphology in lead I<sup>20</sup>.

**Intramural:** it is challenging to identify the exact sites of PVC origins in those regions by the current mapping technique and technologies because of the anatomical barriers such as an intramural region. However, according to literature OTVAs arising from intramural origins usually present a left inferior axis QRS, and they rarely present an

RBBB pattern<sup>4</sup>. Also, the QRS duration of intramural SOO is usually shorter than from epicardium origins.



*Figure 7. Typical 12-lead ECG morphologies from common outflow tract sites, distinguishing between a RVOT and a LVOT origin. AMC: aortomitral continuity; FW: free wall; LCAS: left coronary aortic sinus; LV: left ventricular; LVOT: left ventricular outflow tract; MA<sub>Ant</sub>: anterolateral mitral annulus; MV: mitral valve; PS: posteroseptal; PSC<sub>L</sub>: left pulmonary sinus cusp; PSC<sub>R</sub>: right pulmonary sinus cusp; RCAS: right coronary aortic sinus; RVOT: right ventricular outflow tract; and TV: tricuspid valve.*

Numerous algorithms for predicting the origin of OTVAs using 12-lead ECGs have been reported, although accurate localization of the OTVA is limited<sup>21</sup>. In 2019, Penela et al. presented an easy-to-use algorithm in order to differentiate between a left or right sided OTVA origin by integrating clinical baseline data from the patient, such as his age and previous diseases, with his ECG, concluding that a LVOT origin was related with older age men suffering hypertension<sup>11</sup>.

An ECG algorithm to localize idiopathic VTs using the bundle branch block pattern, the precordial transition, and ECG characteristics of leads I and V1/V6<sup>3</sup> concluded that despite limitations related to complex anatomy, the 12-lead ECG analysis remains an indispensable tool for improving ablation success. Another ECG algorithm for retrospectively predicting the SOO in RVOTs performed in 39 patients achieved an overall positive prediction value of 77.3%, considering 7 different specific sites from the

RVOT, while this value increased until 88.2% when performing a prospective study in 13 more patients<sup>22</sup>. Finally, Betensky et al. developed a V2 transition ratio based algorithm to differentiate the ventricular origin in patients with precordial transition in V3, defined as the most challenging case since it can correspond to a left or right VA origin. It consisted in comparing the R/S amplitude of the PVC with the R/S ratio of a sinus ECG and concluded that this is a reliable measure for predicting the chamber origin in patients with V3 transition<sup>18</sup>.

Despite the successes of these techniques, some of the drawbacks related to individual variability are the inability to correct cardiac rotation, which is commonly found in patients, the variation in electrode placement, specific conduction properties, aorta deformities, chest wall thickness or medication consumption<sup>20</sup>.

In this work, a Machine Learning (ML) pipeline developed in collaboration with Dr. Lozano from Universitat de Valencia is going to be tested using real world ECG databases, for the differentiation of the ventricular chamber origin in OTVAs. ML is an application of artificial intelligence that provides systems the ability to automatically learn and improve learning from data, using statistical methods.

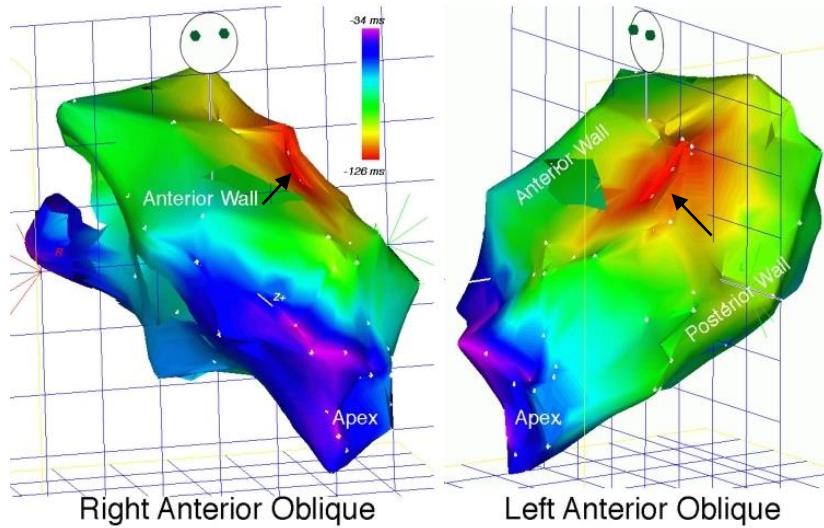
## 1.5 Radiofrequency ablation procedure

Radiofrequency (RF) catheter ablation is a widely acceptable reliable approach for patients with drug resistant arrhythmias. RF ablation uses the principle of RF generated heating by applying its current flow on the diagnosed SOO<sup>23</sup>. Focused energy delivery causes a high tissue temperature rise and produces tissue thermal injury. RF catheter ablation can be performed using unipolar or bipolar ablation. Unipolar ablation consists in applying RF from the catheter tip to a ground electrode located at the patient's skin, while bipolar ablation consists in applying RF flow from a catheter tip to another's one, close to the first one, and allowing deeper transmural injuries, while minimizing the risk of adjacent organ injury<sup>23</sup>.

Before performing RF ablation, electroanatomical mapping (EAM) is traditionally performed during the intervention in order to obtain an activation map of the cardiac structure and identifying the SOO. EAM allows physicians to record intracardiac

electrical activation in a cardiac chamber of interest during arrhythmia mapping, using a multielectrode mapping catheter.

Electrical activation maps allow the representation of the arrhythmia propagation along the cardiac structure. In Figure 8, an activation map of left ventricular tachycardia is observed. Right anterior oblique and left anterior oblique views are displayed. The points of the same color represent isochrones activation zones. The red isochrones indicate earliest ventricular activation, and therefore, it is diagnosed as the tachycardia site of origin<sup>24</sup>. Tachycardia spreads along the left ventricle according to the color-coded time scale.



*Figure 8. Left ventricular activation map where the arrhythmia spreads along the left ventricle. Earliest activation site is found in red color where the black arrow is pointing. Image from <sup>24</sup>.*

In case there are unsustained arrhythmias, pace mapping technique allows to stimulate concrete cardiac regions in order perform arrhythmias using an electrocatheter as well, so physicians can identify the SOO.

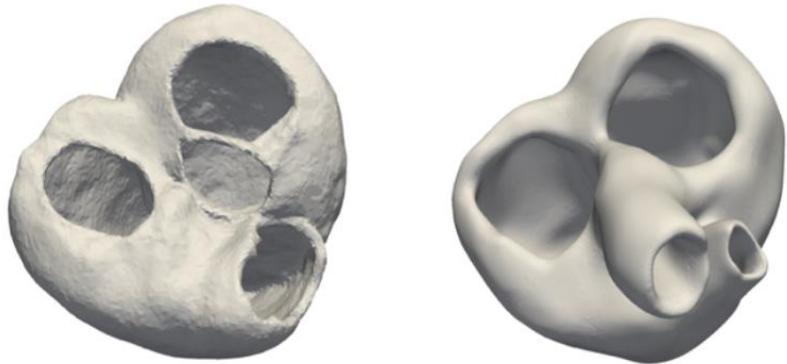
## 1.6 State of the art of computational pace-mapping techniques

Doste et al.<sup>1,12</sup> developed a complete framework in order to perform cardiac arrhythmia simulations for the in-silico pace-mapping prediction of the SOO in OTVAs. From cardiac computed tomography (CT) images, a 3D biventricular heart model is designed,

whose fiber orientation is assigned by a rule based method algorithm<sup>25</sup>. Simulated ECGs from cell to body level were computed allowing to simulate patient specific ECGs from 12 different sites of origin (in-silico pace mapping). Finally, these predictions were retrospectively validated quantitatively comparing the obtained fully simulated ECGs with patients' preoperatively acquired ECGs using correlation techniques and also machine learning (ML) tools in order to assign the most likely SOO. This study was performed in 11 patient specific geometries, acquiring 12 ECG simulations for each one. It was found that the modelling pipeline predicted right or left ventricle origin in 10/11 patients, but not as high accuracy was found when predicting the exact SOO (54% of accuracy). The results using ML techniques showed high accuracy (89%) when using a k-NN (k nearest neighbors) algorithm for differentiating a left vs. right ventricle origin as well. A large simulated ECG database was built to use as a training dataset. One of the biggest concerns of this work was that the region above the outflow tracts was not considered, and also that this was performed in a small sample of patients. Also, that the pipeline had a very high computational cost and ECG simulations took hours to run.

Unlike fully simulated ECGs, pseudo-ECGs do not consider the tissue volume and characteristics of the torso between the heart and the leads, instead, they approximate the gradient of the transmembrane voltage at distinct points of the whole surface of the myocardium, providing an approximate measure of the electrical activation of the cardiac system at the tissue level<sup>2,26</sup>. This allows lowering the computation time.

Moya<sup>2</sup> adapted Doste's work by simulating pseudo-ECGs from 16 distinct SOO instead of fully simulated ECGs, which allowed minimizing the computational cost and building a more user-friendly pipeline, and also considered the regions above the pulmonary and aortic valves. In Figure 9 there is a comparison between the biventricular model obtained in Doste's work (left) and the one from Moya, where it can be seen that the aortic root and the pulmonary artery are considered.



*Figure 9. Comparison of the computational biventricular heart models extracted from the same patient CT images. Left is from Doste<sup>1</sup> and right is from Moya<sup>2</sup>.*

Finally, the results of this work showed that by looking at one single pseudo-ECG lead, the algorithm correctly predicted the SOO in 70-85% of cases. However, further generated pseudo-ECG validation must be performed, since in Moya's thesis only two patient specific pseudo-ECGs were quantitatively compared with the patient's real ECG. Also, the SOO considered in Moya's work are redefined in this work as suggested by Dr. Penela, an expert electrophysiologist, by defining less but more characteristic points, since with these simulations significant differences between very close points are not observable.

The implementation of these frameworks into clinical practice has not been evaluated yet. Accurate patient geometry modelling, long computational times or data acquisition during patient intervention add difficulties at the time to implement these techniques into the clinical environment. For this reason, the main objective of this work is to see how this can be used as a clinical support tool during patient's intervention, by shortening the whole framework's execution time while still obtaining accurate results.

## 1.7 Objectives

The main objective of this project is:

- To test in a clinical environment the in-silico modelling and machine learning approaches developed for the preoperative identification of the site of origin in OTVAs.

In-silico simulations will be run in the Nord3 cluster of the Barcelona Supercomputing Centre, which has kindly provided the High Performance Computing resources for the development of this project.

And as sub-objectives:

1. To validate the use of pseudo-ECGs as a lower computationally alternative to fully simulated ECGs.
2. To add more accurate SOO in the in-silico simulations.
3. Testing the machine learning algorithms with different real-world clinical ECG databases.

## 2 Methods

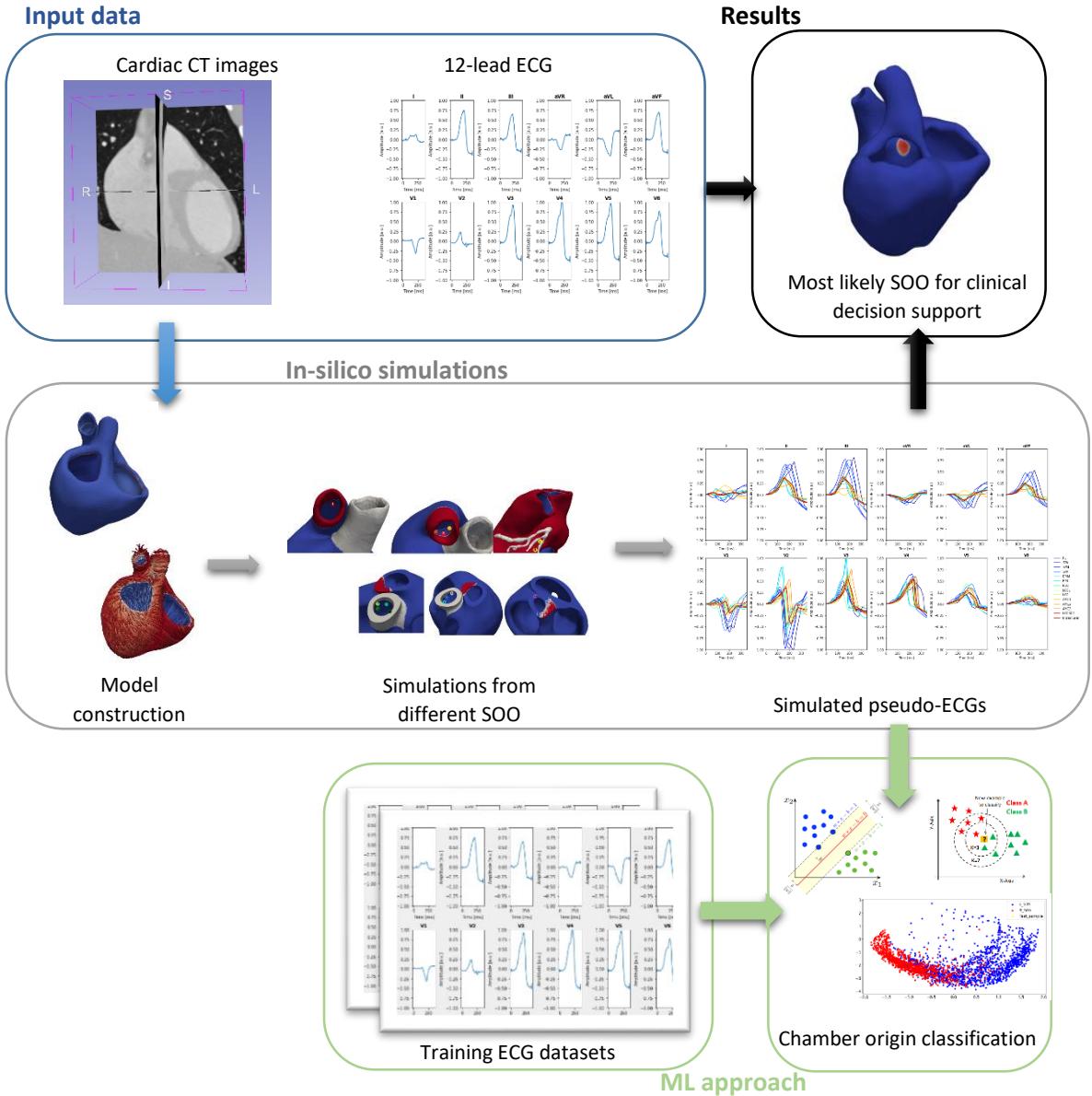
As it has been mentioned above, the previous studies from Doste et al.<sup>1</sup> and Moya<sup>2</sup> were performed retrospectively. In this project, the pipeline was tested closer to the clinical workflow by performing experiments including patients from Hospital Teknon-Quirón to validate its accuracy and implementability into the clinical workflow. The in-silico modelling pipeline and a ML approach were tested analyzing how they could be effectively be used in the clinical environment.

This project has been developed in collaboration with Barcelona Supercomputing Centre, which has provided the High Performance Computing (HPC) resources in order to run the electrophysiological simulations.

### 2.1 Computational pipeline

Initially, a patient specific biventricular model is built from cardiac CT images. The aortic and pulmonary valves specific location and the myofiber orientation were applied to this mesh using computational algorithms. Later, simulations were run from different SOO within the outflow tract regions, and pseudo-simulated ECGs will be obtained. These pseudo-ECGs were compared to the patient previously acquired ECGs performing statistical analysis in order to find which of the pseudo-ECGs from a specific SOO is the most similar to the patient acquired ECG, and thus find which of the considered SOO is the most likely ectopic foci of the arrhythmia.

Regarding the ML approach, a mixed ECG and simulated ECG database was used to train a Support Vector Machine (SVM) and a k-Nearest Neighbors (k-NN) algorithm after applying dimensionality reduction using Principal Component Analysis (PCA), to differentiate between a left or right ventricular origin in patients with OTVAs. Different approaches were tested varying the SVM, k-NN and PCA parameters, as well as the datasets used. This study was performed programming in Python<sup>27</sup>. Both of these pipelines are represented in the Figure below.



*Figure 10. Representation of the computational pipeline, includint the in-silico pseudo-ECG based simulations, and the machine learning approach.*

## 2.2 Pseudo-ECG based in-silico pace-mapping pipeline

### 2.2.1 Modelling dataset

A total of 11 cardiac CT images with their corresponding extracted OTVA were used to generate the pseudo-ECGs applying the in-silico modelling framework. 10 of these patients' data was previously used in the work of Doste<sup>1</sup> and Moya<sup>2</sup>, and an additional case was obtained from Hospital Teknon-Quirón, to evaluate the feasibility of

implementing this framework into the clinical daily basis. Also, 8 of the 10 patient specific geometries were provided by Moya, who kindly sent the biventricular meshes. Without this, all this work would not have been possible since obtaining the detailed patient-specific geometries requires a high amount of time.

### 2.2.2 Building the biventricular mesh from cardiac CT images

From cardiac CT images, an empty biventricular model was created using *3D Slicer*<sup>28</sup>, an open source software platform for medical image informatics, image processing, and three-dimensional visualization. This biventricular model includes the outflow tract regions, and up to 2-3 cm above the pulmonary and aortic valves. This is created using the paint and grow from seeds tools, which allow to manually define the regions that will be segmented and apply an algorithm to grow from these seeds.

This model was then post-processed using *Blender*<sup>29</sup>, another open source software for 3D object creation, which was used to smooth specific zones of the geometry in order to obtain a more specific and smoothed model. Results are displayed in Figure 11.

The resulting meshes consisted of around 6 million nodes and 2 million elements. In order to lower the size of the meshes for subsequent processing in regular computers without a powerful Graphical Processing Unit (GPU), some filters were applied using *Paraview*<sup>30</sup> software. First, an Extract Surface filter was applied, which extracts the polygons forming the outer surface of the model. Then, a Tessellate filter was used, which converts this extracted surface into a simplicial complex with linearly interpolated field values. An Extract Surface filter is applied again, followed by a Decimate filter, which reduces ten times the number of elements of the mesh, and finally a Triangulate filter to obtain a triangular element mesh.

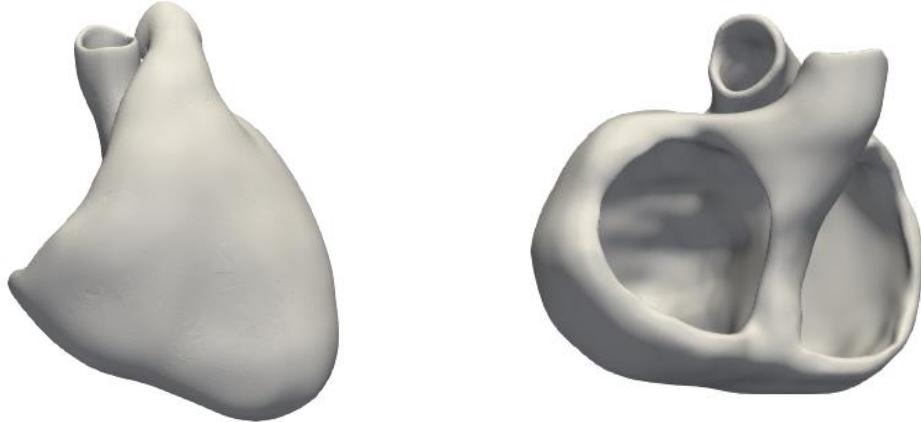


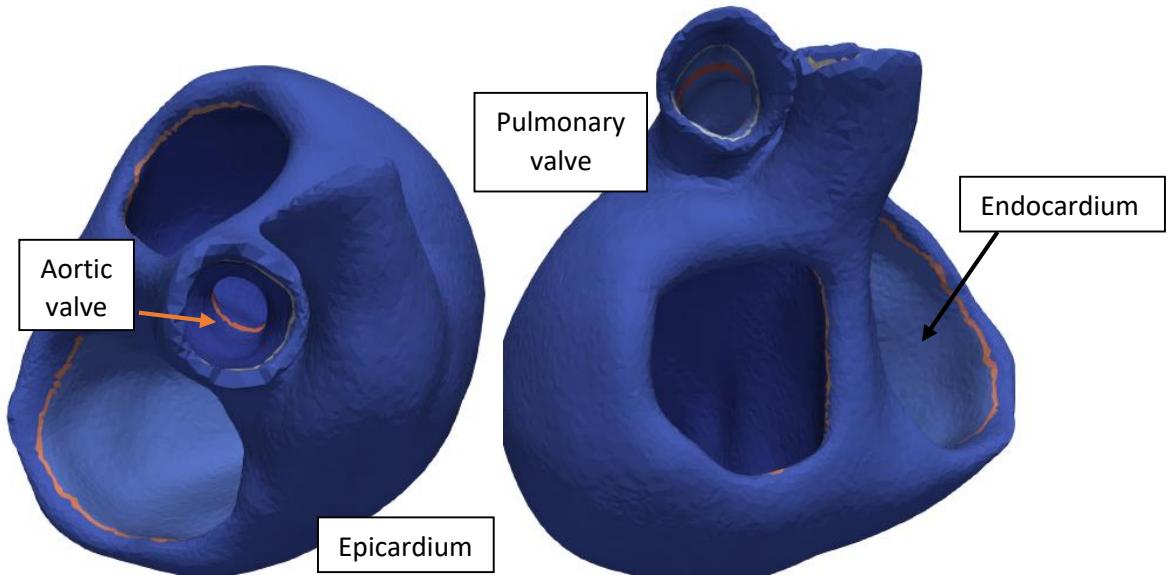
Figure 11. Resulting biventricular meshes from the two first patients.

### 2.2.3 Mesh labelling and determining fiber orientation

Once the biventricular mesh is post-processed, the epicardium and endocardium are labelled following a semi-automatic ray tracing algorithm, also presented in Dr. Doste's PhD<sup>1</sup>. In order to apply this algorithm, an epicardial left ventricular apex point, and the endocardial right ventricular apex point ID must be manually identified. Then, rays are traced from the normal of every face of the surface mesh, and the intersections with other faces are computed. If intersection were found, the face is labelled as endocardium. Either way, it is labelled as epicardium<sup>1</sup>.

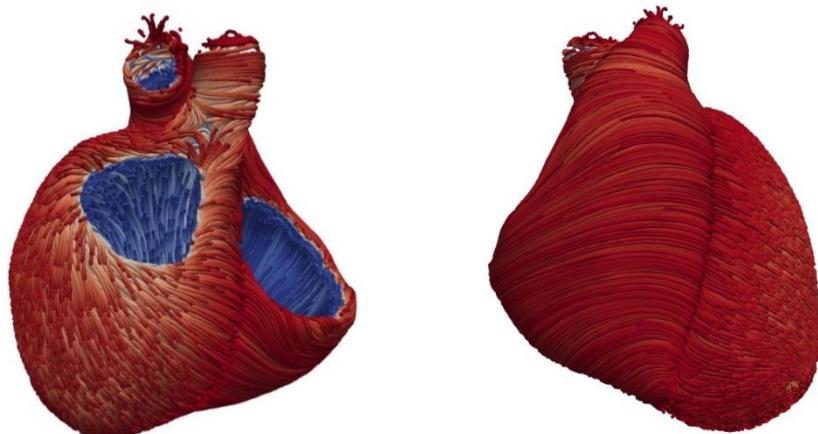
The pulmonary and aortic valves locations is determined manually, by segmenting them using *Slicer*<sup>28</sup> again. Once their location is determined, it is added to the labelled mesh, so we finally have a labelled mesh with the outflow tracts valves location (observed in Figure 12).

The next step is to build a volumetric mesh using *Matlab*<sup>31</sup> iso2mesh tool. This resulting mesh has approximately around 3 million elements.



*Figure 12. Results after adding the valves locations to the model, marked as red and orange lines.*

Cardiomyocyte orientation determination is essential in order to build realistic cardiac models. The fiber orientation algorithm developed by Dr. Doste in his PhD<sup>1</sup>, assigns a unitary 3D vector to every one of the nodes of the biventricular mesh, by applying a rule-based method. This 3D vector is oriented in the preferential direction of electrical wave propagation and myocardial tissue contraction. This method is based on histological observations on ex-vivo hearts which is modelled using a set of mathematical equations. The results after applying this algorithm are displayed in Figure 13. It is noticeable that the orientation of the fibers is opposite in both ventricles, and that it changes following a similar pattern in both ventricles from the epicardium to the endocardium.



*Figure 13. Visualization of the streamlines after applying the fiber orientation rule based method in the first patient geometry.*

## 2.2.4 Sites of origin considered in this work

In this work, 14 different specific SOO were considered as suggested by Dr. Penela, an experienced electrophysiologist and researcher about OTVAs, since he pointed out that these are the most common and characteristic SOO in OTVAs. In the aforementioned work, the number of SOO was excessively large, and in many cases, points were anatomically too close that it is not possible to observe differences using computational approximations between two very close points. For this reason, in order to improve simulation accuracy, these SOO were redefined by defining more separate and characteristic points. Also, in order to obtain a more accurate location of the LV summit, left coronary arteries were segmented using *3D Slicer*, as observed in Figure 14. The red and grey regions represent the regions above the pulmonary and aortic valves respectively, for better visualization.

From the RVOT four SOO were considered, which can be observed in Figure 14:

- 1) Pulmonary lateral (PL)
- 2) Septal anterior (SPA)
- 3) Septal medium (SPM)
- 4) Septal posterior (SPP)

From the epicardium, using the coronary arteries segmentation as reference, two SOO were determined, also seen in Figure 14:

- 5) Summit medium (EPIM)
- 6) Summit lateral (EPIL)

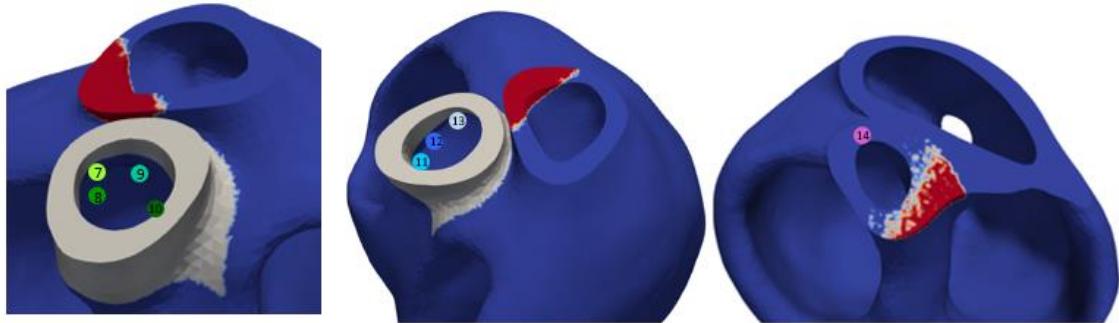


Figure 14. Points considered from the RVOT (1-4), below the pulmonary valve, and from the summit (5,6). RVOT: right ventricular outflow tract.

Then from the LVOT, the following points were considered, displayed in Figure 15.

- 7) Left coronary cusp low (LCCL)

- 8) Left coronary cusp (LCC)
- 9) Right and left coronary cusp commissure (LCCRCC)
- 10) Right coronary cusp (RCC)
- 11) Aortomitral continuity (AMC)
- 12) AMC2
- 13) AMC3
- 14) Intramural point under the LCCRCC



*Figure 15. Sites of origin determined from the left ventricle outflow tract.*

For each patient, and for each site of origin, one ID point is selected in *Paraview* by manually selecting the stimulation region (the SOO defined in this section). The coordinates of this point are introduced in the corresponding input file to run *Alya* simulations, and an activation radius is defined so all the points surrounding the selected one under that radius are activated.

### 2.2.5 Simulations input files

*Alya* is a high performance computational mechanics code to solve engineering coupled problems using numerical methods, developed by BSC. In this work, the electrophysiological *exmedi* module is used to solve the electrical propagation problem as a monodomain problem, obtaining the simulated pseudo-ECGs from different SOO. The O'Hara cellular model<sup>32</sup>, the most state of the art cellular model for simulating human ventricular electrophysiology, is used in these simulations. It allow to simulate arrhythmic

behavior and consists of 50 variables and 41 differential equations, which is the reason why HPC is necessary to apply this model.

The duration of the simulated QRS complexes was 350 ms with a time step of 1 ms. Tissue conductivity was set as constant, at a value of 12 S/m, in the range of cardiac cell conductivity proposed by Stinstra et al.<sup>33</sup>, and the outflow tract regions were set as non-conductive material. Pathological tissue was not considered, as it corresponds to OTVAs. A cycle length of 857 ms was defined, corresponding to 70 beats per minute. Also, a starting transmembrane potential of -87.99 mV was set.

In order to get the pseudo-ECGs generated from each SOO, 13 inputs files are needed as input for Alya. 7 of these files are general for all geometries (except case.dom where the number of elements, boundary elements and nodes are modified manually):

- case.dat: the case name, the simulation duration, the time steps and the modules that will be used are defined.
- case.dom: the number of nodes, elements, and boundary elements are defined, as well as the number of fields and dimensions. In this case there are two fields, Field 1 = fibers (3 dimensions), Field 2 = cell type (unidimensional). The boundary conditions, mesh geometry, materials and sets are defined.
- case.exm: corresponds to the *exmedi* module. In this file, the starting potential from the considered SOO, the electrode coordinates where the pseudo-ECGs are computed, as well as the material properties (membrane potential, cellular model) and the postprocessing of the results are defined.
- case.ker: contains information related to the numerical treatment and the postprocessing.
- case.post: specify data format and simulation parameters.
- Ohr\_New\_framework: parameters for O’Hara Rudy cell model, included in case.exm (in material properties of the conductive tissue).

Then, there are 6 more files specific for each model generated after the fiber orientation algorithm is implemented:

- case.cell: assigns to each element its cell type (epicardium, endocardium and myocardium). Included in case.dom.

- case.defsets: assigns to each node its corresponding set.
- case.fibers: assigns to each element the fiber orientation in 3D.
- case.fix: included in muscle.dom, contains the mesh boundary conditions for each element.
- case.geo: assigns the number of nodes for each element (4 for all since they are tetrahedral) and then, for each element, it assigns the ID of the corresponding 4 nodes. Included in case.dom.
- case.mat: it assigns the material (conductive or non-conductive) to each of the elements. Included in case.dom.

These simulations were run in the Nord3 cluster from the BSC. The output for each simulation is the 12-lead generated pseudo-ECGs from one SOO, and the calculation time is 20 minutes. That is, for each patient, 14 simulations were run, which could be executed simultaneously. In order to determine the coordinates of the electrodes where the potential is calculated, a torso model containing the most relevant organs of the chest, such as the lungs or the ribs, was downloaded from the CoMMLab repository<sup>34</sup>. The biventricular model was manually located within that torso model for having a simplified reference of the electrode location with respect to the heart geometry.

The whole pipeline implementation time has been calculated to analyze how these steps could be implemented into the clinical workflow, and what steps could be accelerated.

- Semi-automatic heart segmentation → 4 – 5 h
- Surface mesh smoothing and filtering → 30 min
- Mesh labelling (epicardium, endocardium and valves) → 30 min
- Creating volumetric mesh → 1 h
- Adding fiber orientation → 1 h
- Preparing input files, running simulations and analyzing results → 1 – 2 h

Approximately, every patient case requires 8 to 10 hours of work following the pseudo-ECG based in-silico pace mapping pipeline. This is a major drawback when trying to bring these sophisticated, time consuming techniques into practice, considering that the patient CT is usually acquired the same day, just before the intervention.

## 2.3 Machine learning approach

### 2.3.1 Machine learning algorithms evaluation datasets

Two different datasets were used to test two ML classification algorithms. The first dataset, contained the simulated ECGs developed by Dr. Doste in his PhD<sup>1</sup>, and the real ECG dataset provided by Zheng et al.<sup>16</sup>, which is called dataset A for simplification. The second dataset, called dataset B, contained the aforementioned dataset by Zheng et al, and the simulated pseudo-ECGs obtained in this work. The number of samples for each dataset and their combinations are displayed in Table 1 and 2. This algorithm was developed in collaboration with the Computer Science department in Universitat de Valencia. The classifier was trained to distinguish between a left or right ventricular origin of the OTVA.

Two algorithms were fitted with each of the datasets: a SVM classifier using a Gaussian kernel, and a k-NN algorithm using 4 neighbors. In both cases, a PCA dimensionality reduction algorithm was used selecting the 3 first principal components in the PCA space. In Figure 16 can be seen the cumulative variance depending on the number of components using PCA with dataset A.

*Table 1: Original datasets used in this work.*

	Samples	LVOT origin	RVOT origin
Doste's simulated ECGs <sup>1</sup>	2496	1456	1040
Zheng et al. dataset <sup>16</sup>	333	77	256
Simulated pseudo-ECGs	154	110	44
Total	2983	1643	1340

Table 2: Dataset combinations used in this work.

Data	Dataset ID	Samples	LVOT origin	RVOT origin
Zheng et al. dataset <sup>16</sup> + Doste's simulated ECGs <sup>1</sup>	A	2829	1533 (0.54)	1296 (0.46)
Zheng et al. dataset <sup>16</sup> + simulated pseudo-ECGs	B	487	187 (0.38)	300 (0.62)

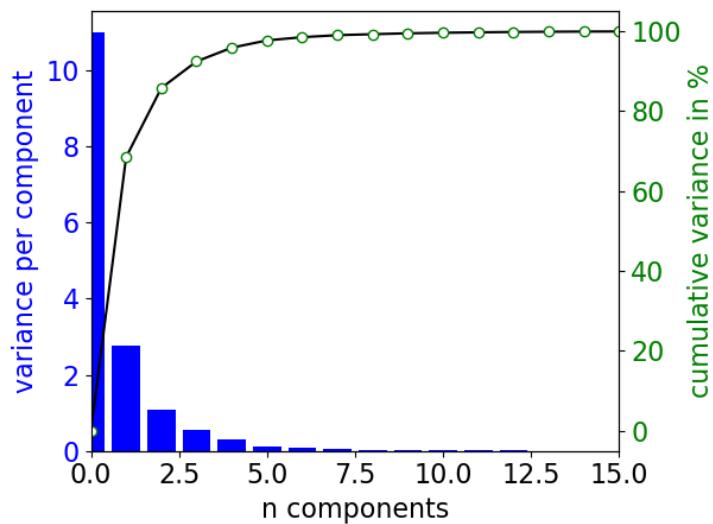


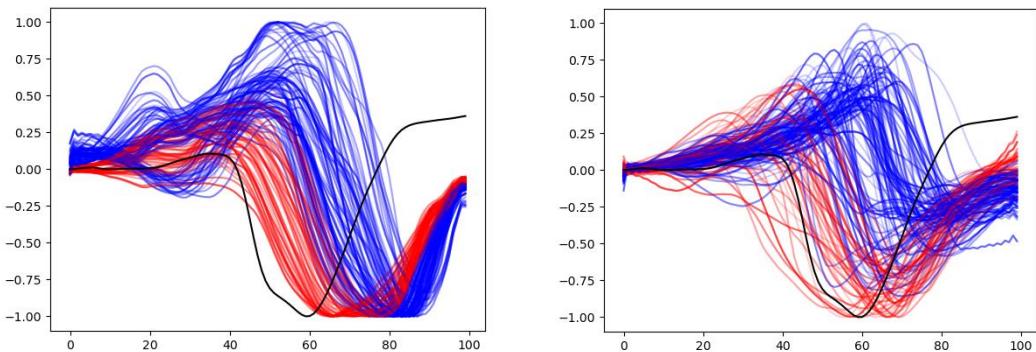
Figure 16. Variance per component and cumulative variance as a function of the number of components using PCA algorithm applied to the ECG training dataset. PCA: Principal component analysis.

### 2.3.2 Machine learning algorithms steps

The first algorithm step is to manually select the R point of the PVC for this to be extracted (Figure 17). Then, the QRS complex is extracted, normalized between -1 and 1, and smoothed. This processed signal can be seen in black in Figure 18 plotted together with training data, both with fully simulated ECGs and simulated pseudo-ECGs. Concretely, the red and blue lines (RVOT and LVOT origin respectively) correspond to simulated ECGs from Dr. Doste's PhD<sup>1</sup> (left), and pseudo-ECGs simulated in this work (right).



*Figure 17.* Sample V2 lead from one patient, concretely with a LVOT PVC origin. The PVC corresponds to the third beat, centered on sample 2000. PVC: premature ventricular contraction.

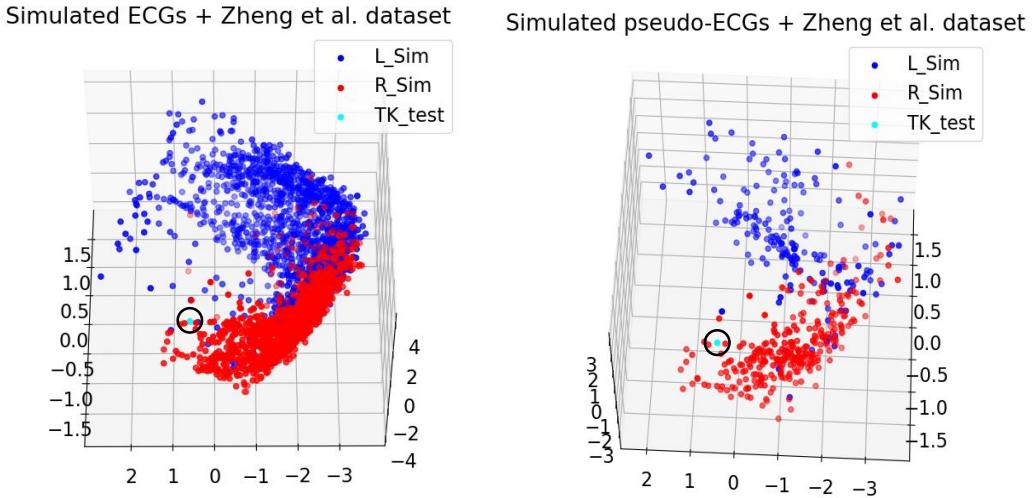


*Figure 18.* Left: visualization of simulated V2 lead QRS complex obtained by Dr. Doste, (red and blue lines, corresponding to al RVOT and LVOT origin respectively), against the patient extracted QRS (black) originated in the LVOT. Right: simulated pseudo-ECGs obtained in this work (RVOT: red lines, LVOT: blue lines) against the patient extracted QRS (black), originated in the RVOT. RVOT: right ventricle outflow tract, LVOT: left ventricle outflow tract.

Results using both methods were compared. In all cases cross-validation was used to obtain more robust classification accuracy using 5 folds.

Precordial lead V2 was used to train and test the classifier since the highest differences between left against right sided originated PVCs were found on this lead, as Dr. Lozano suggested. In Figure 19, a PCA space representation of V2 leads using three principal components shows how the two clusters are easily distinguishable using both dataset

combinations. Blue points correspond to LVOT origins while red points belong to RVOT origins. The cyan spot circled in a black circle corresponds to the classification of the test sample observed in Figure 19, which is located closer to the red points (RVOT origin).



*Figure 19. Representation of the QRS samples in the PCA space using three principal components. Red spots: RVOT origin QRS, blue spots: LVOT origin QRS complexes, cyan spot: sample classified QRS, corresponding to a RVOT origin. Left image: using fully simulated ECGs and Zheng et al. ECG dataset<sup>16</sup> as training data (dataset A). Right image: using simulated pseudo-ECGs and Zheng et al. ECG dataset<sup>16</sup> as training data (dataset B). RVOT: right ventricle outflow tract, LVOT: left ventricle outflow tract.*

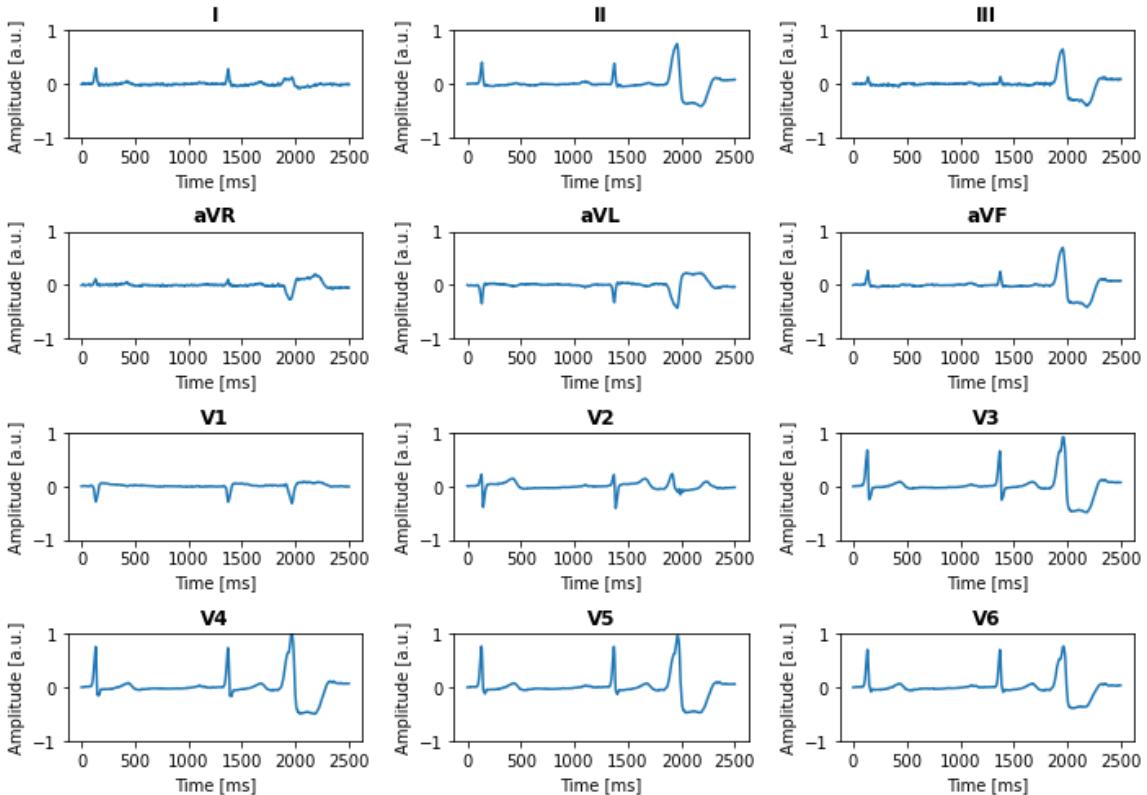
## 2.4 Prospective implementation of the pseudo-ECG based in-silico paced mapping and the machine learning algorithms

One of the main objectives of this project is to analyze how these described pipelines could be implemented into the clinical workflow, overcoming the challenge of adapting these long computational time consuming methods. For this purpose an experiment was performed using data from patient 11. This patient's CT and ECG was obtained from Teknon-Quirón Hospital.

With the objective of reducing the segmentation time, the biventricular model segmentation was performed in Teknon using Adas<sup>35</sup>, an imaging processing platform

intended to help healthcare professionals visualize cardiac structure and fibrosis. However, it was not possible to obtain a closed endocardial and epicardial surface mesh using this software as it was required for the modelling pipeline, since it is more appropriate for obtaining only endocardial surface meshes, so after many attempts, the segmentation was performed using *Slicer* like in the other cases starting over again.

The whole pipeline implementation time from the CT segmentation until obtaining the simulated pseudo-ECGs took approximately 10 hours, and obviously the results were analyzed after the patient's intervention. The extracted ECG of the patient is displayed in Figure 20, where the extra is the third observed beat centered at 2000 ms.



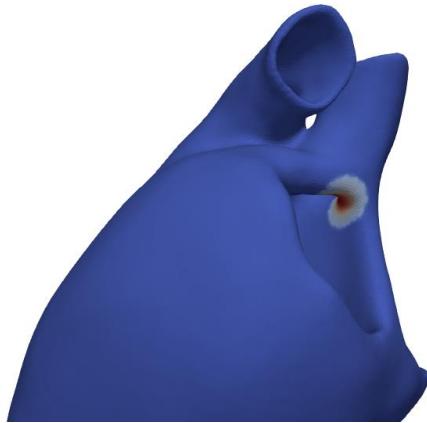
*Figure 20. 12-lead ECG obtained for patient 11, which was used to analyze the feasibility of implementing both pipelines into the clinical workflow. The third beat (around 2000 ms) is the one corresponding to the PVC (premature ventricular contraction).*

## 3 Results

### 3.1 Pseudo-ECG based in-silico paced mapping

In this study, a total of 11 patient specific geometries were analyzed, from which pseudo ECGs were obtained. For each case, 12-lead pseudo-ECGs from 14 different SOO were obtained. That is a total of 154 12-lead pseudo-ECGs.

For each simulation, a total of 350 time-steps with a step size of 10 ms were performed. The stimulus propagation within the geometry can be visualized as observed in Figure 21, where the stimulus clearly propagates from the considered SOO (in this case, from the RVOT), until this potential propagates through the whole biventricular geometry.



*Figure 21. Stimulated point at time step 1 from the aortomitral continuity.*

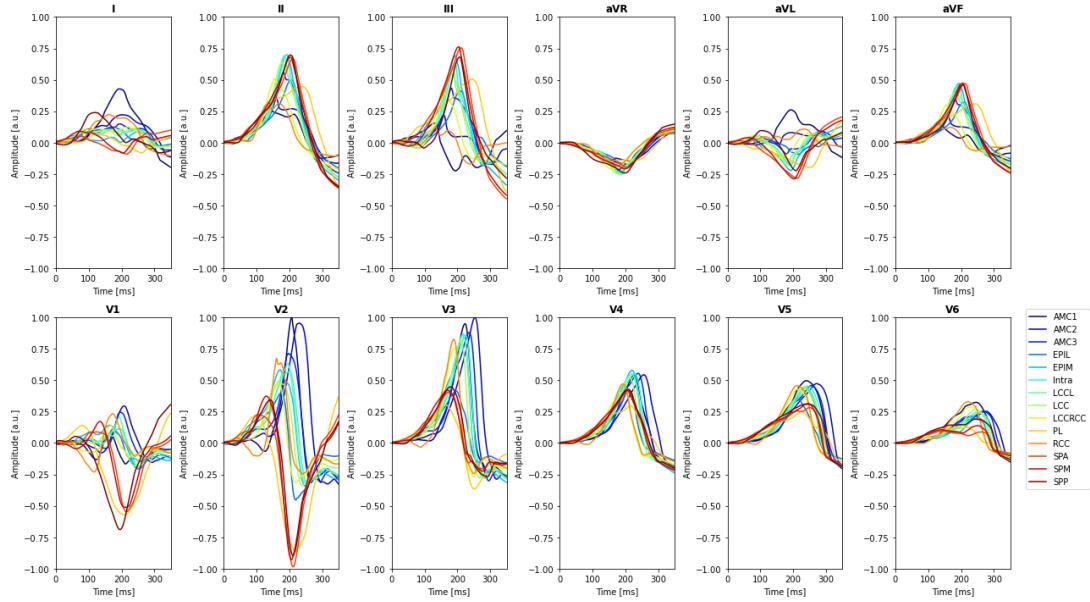
### 3.1 Simulated pseudo-ECGs

In Figure 22 it can be observed the resulting simulated pseudo-ECGs from the 14 different SOO for a patient specific geometry. Cold colors correspond to LVOT SOO, while hot colors correspond to SOO arising from the RVOT. The most noticeable differences between these left or right sided originated pseudo-ECGs are observed in the three first precordial leads (V1-V3).

Both ECG and pseudo-ECG signals were normalized between -1 and 1. The real ECGs were normalized with respect to its 12-lead maximum since its order of magnitude is

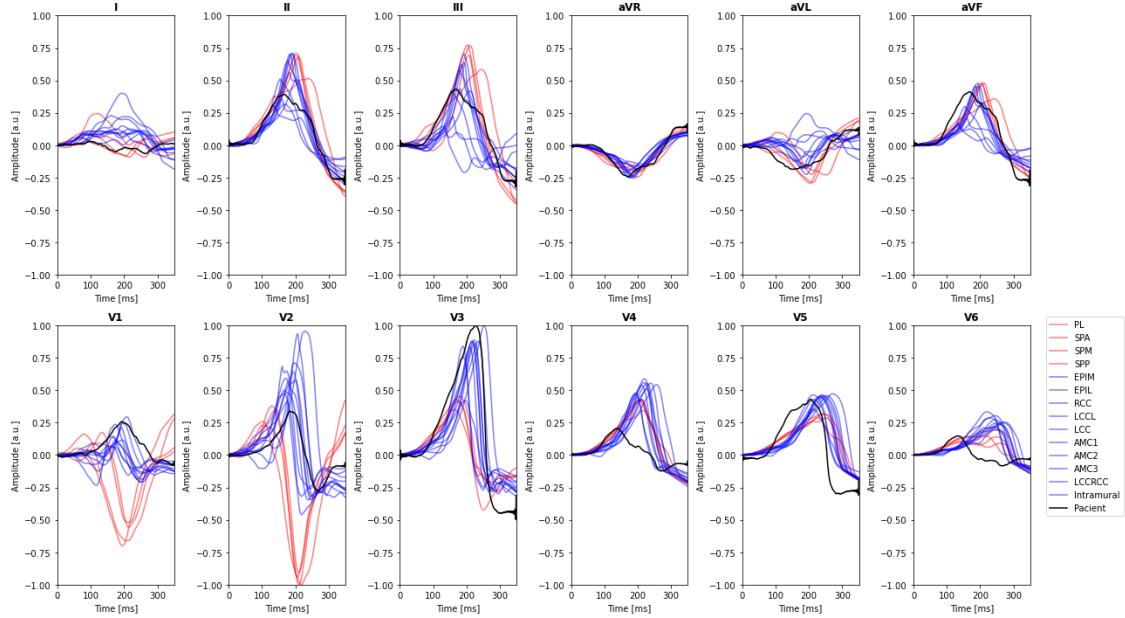
larger than the one of the simulated ECGs. Regarding pseudo-ECGs, for each SOO, they were normalized to its maximum, so the pseudo-ECGs for each SOO are in the range between -1 and 1. Also, real ECGs were preprocessed so the QRS onset starts at 0, since the QRS simulated complexes start at 0 too.

An earlier precordial transition was found on those simulated pseudo-ECGs corresponding to an LVOT origin, while a later precordial transition is found on those origination from the RVOT. This difference is clearly distinguishable in lead V2.



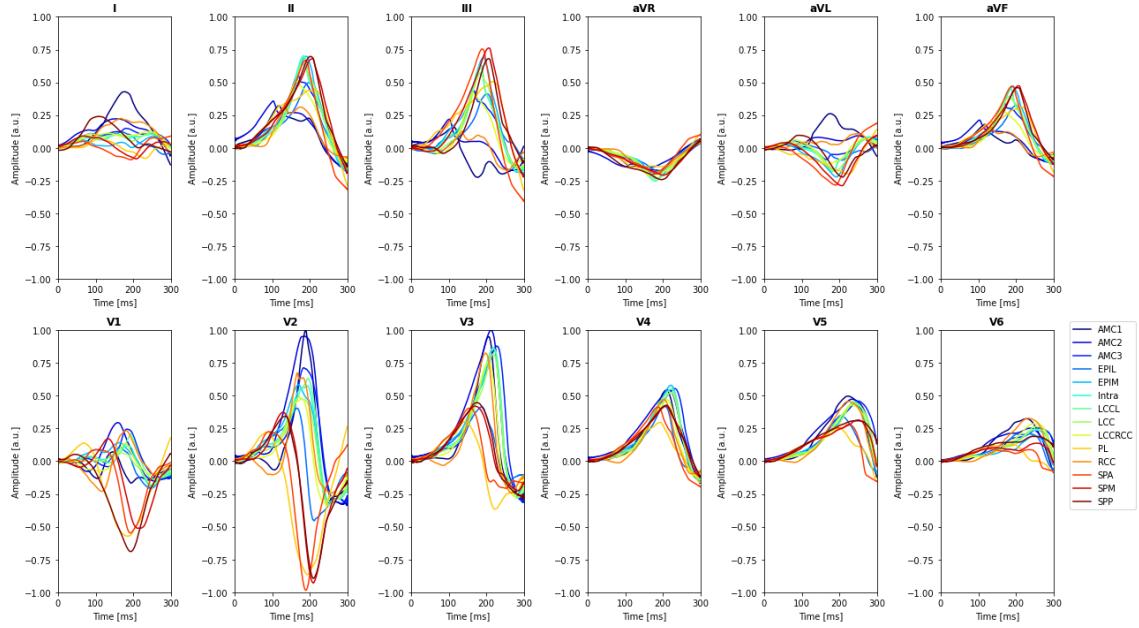
*Figure 22. Simulated pseudo-ECGs from 14 different SOOs. Cold colors refer to RVOTs and hot colors to LVOTs. AMC: aortomitral continuity, LCC: left coronary cusp, LCCRCC: left and right coronary cusp commissure, RCC: right coronary cusp, EPIM: epicardium medium, EPIL: epicardium lateral, PL: pulmonary lateral, SPA: septal anterior, SPM: septal medium, SPP: septal posterior.*

In Figure 23, a clear differentiation between LVOT and RVOT origin pseudo-ECGs can be observed. Also, the patient real ECG is plotted, which allows comparing the similarity between real and simulated ECGs. This concrete patient was diagnosed with a LVOT origin, concretely in the LCC.

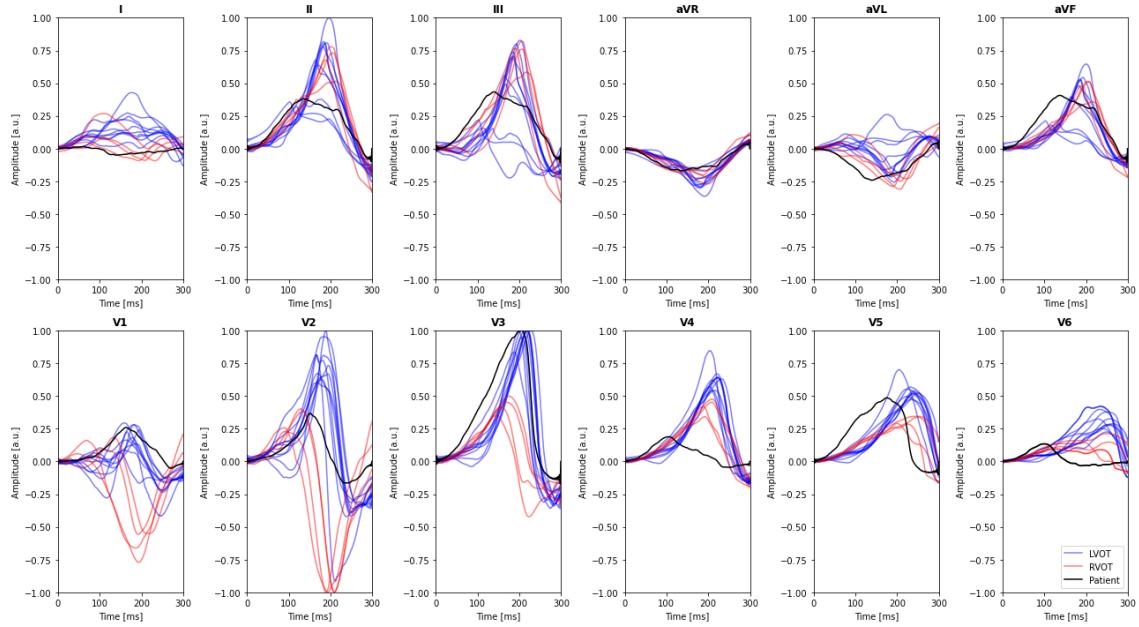


*Figure 23. Comparison of 12-lead LVOT (blue) against RVOT (red) origin generated pseudo-ECGs, against real patient ECG (black) diagnosed with a LCC LVOT origin. LVOT: left ventricle outflow tract, RVOT: right ventricle outflow tract, LCC: left coronary cusp.*

In order to quantitatively compare the QRS complex of the real ECGs and the simulated pseudo-ECGs in an unbiased and objective way, a U-net architecture based Deep Learning (DL) algorithm for automatic ECG delineation and detection was applied<sup>36</sup>. The QRS complex onset and offset of both real and simulated ECGs was extracted. Regarding pseudo-ECGs, the QRS was extracted independently for each SOO. These aligned pseudo-ECGs are displayed in the Figures below.



*Figure 24. Extracted QRS complex of the generated pseudo-ECGs using a deep learning model<sup>36</sup>. It is observed how the QRS complexes are now aligned.*



*Figure 25. Extracted QRS complex of the generated pseudo-ECGs using a deep learning model<sup>36</sup>, compared to the patient real ECG (black) diagnosed with a LCC LVOT origin. LCC: left coronary cusp, LVOT: left ventricle outflow tract*

### 3.1.2 Qualitative pseudo-ECG validation with clinical literature

Finally, in Figure 26 are displayed in separate columns the 12-lead pseudo-ECGs from each stimulated point, and the main features are circled in different colors.

#### General features

Analyzing the simulated pseudo-ECGs obtained in this work, some typical OTVA characteristics found in the literature are observed. The ECGs of idiopathic OTVAs are characterized by positive R waves in all inferior leads, and deep S waves in both aVR and aVL leads (almost a QS pattern)<sup>4</sup>. These features are clearly shown in most of the simulated pseudo-ECGs. Also, a LBBB pattern is observed in V1 on those SOO corresponding to an RVOT origin, which is consistent with literature as well<sup>37</sup> due to its anterior location with respect to the LVOT, while RBBB and LBBB patterns are observed on SOO corresponding to the LVOT (see V1 in Figure 25).

#### RVOT origin features

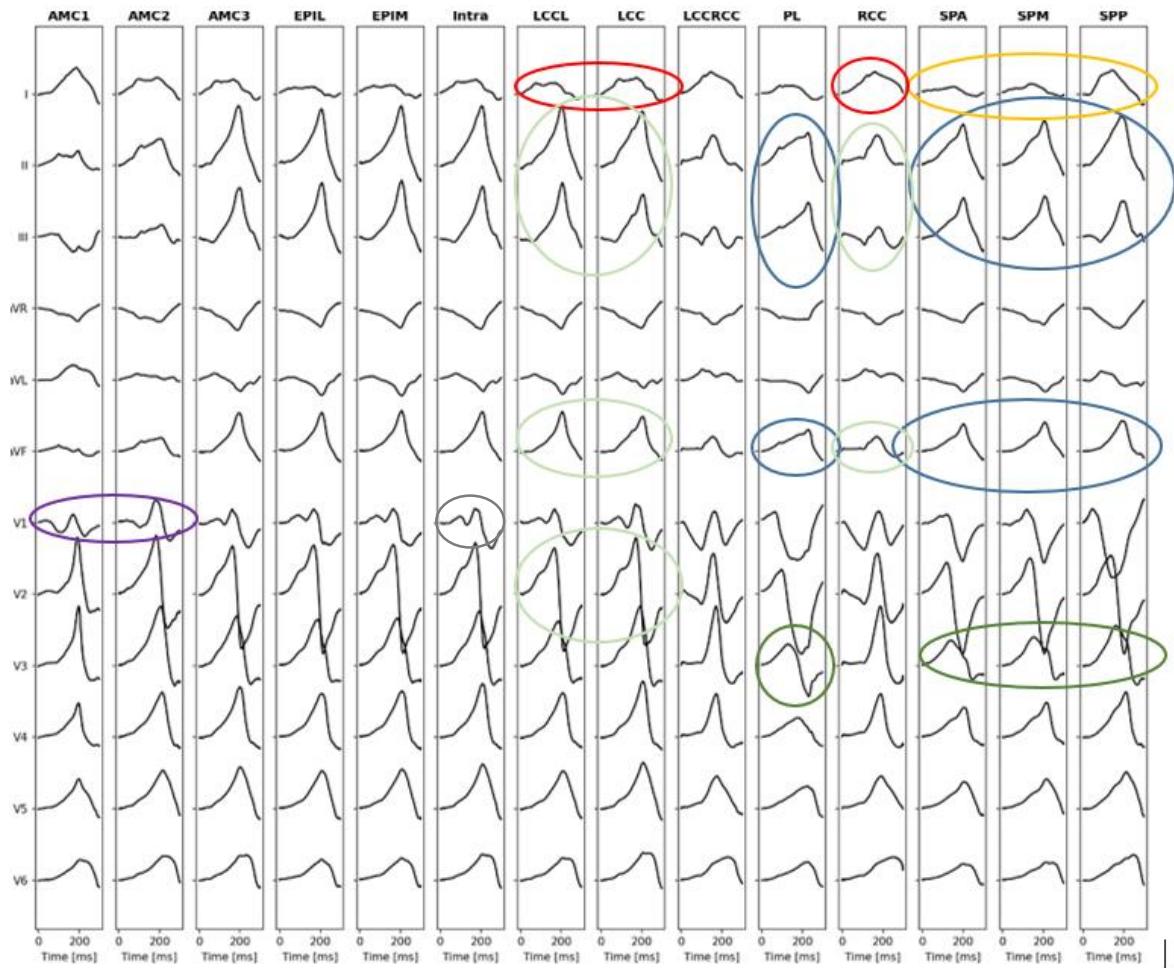
Analyzing more deeply the simulated pseudo-ECGs, within the RVOT origins, taller R waves and more narrow QRS complexes are observed on pseudo-ECGs arising from the **septal** than from the **lateral wall** origins in the inferior leads (II, III, aVF) (marked in blue). Also, a slightly later precordial transition is observed than in septal origins (marked in dark green). However, smaller S waves and taller R waves in septal origins than in PL origins are not observed in leads V1-V3.

Within **septal** origins, from anterior to posterior, posterior origins have more positive lead I than more anterior origins, which is the main difference between SPA, SPM and SPP sites of origin (marked in orange).

#### LVOT origin features

**AMC:** no S wave is found in lead V6as expected, and a qR shape in V1 is observed (marked in purple).

**RCC:** a more positive QRS in lead I than in LCC origins is marked in a red circle. A characteristic LBBB pattern is found in patient 8 (see Appendix).



*Figure 26. Analysis of the pseudo-ECGs (black) in patient 2. Each column corresponds to the pseudo-ECG obtained by stimulating the corresponding SOO. AMC: aortomitral continuity, LCC: left coronary cusp, LCCRCC: left and right coronary cusp commissure, RCC: right coronary cusp, EPIM: epicardium medium, EPIL: epicardium lateral, PL: pulmonary lateral, SPA: septal anterior, SPM: septal medium, SPP: septal posterior.*

**LCC:** an early precordial transition is observed (V2), as well as tall R waves in V2. M wave pattern only seen in LCCRCC origin. Higher R wave amplitude in inferior leads than RCC origin, marked in light green.

**LCCRCC:** very similar characteristics as LCC origins due to its close proximity. Its characteristic notching is not observable.

**Summit:** a dominant R wave is observed in V6, and an rS pattern in lead I is found in patient 11 for example (see Appendix), as well as an LBBB morphology.

**Intramural:** A lower duration of the QRS complex than in summit origins is not observed. An LBBB pattern is observed and marked in grey.

### 3.1.3 Pseudo-ECGs validation metrics

Simulated pseudo-ECGs were quantitatively compared with the patient real ECGs by computing the correlation coefficient ( $\rho$ ) independently for each lead and for each SOO.

In order to align signals where  $\rho$  was maximum, the cross-correlation coefficient was computed between both signals. It measures the correlation between vector  $x$ , and  $n$  lagged copies of vector  $y$ . The maximum cross-correlation value for all lags was taken as the correlation coefficient for each lead between the real and the simulated pseudo-ECG, obtaining a total of 12  $\rho$  values for each SOO. This alignment to obtain the maximum correlation is displayed in Figure 27.

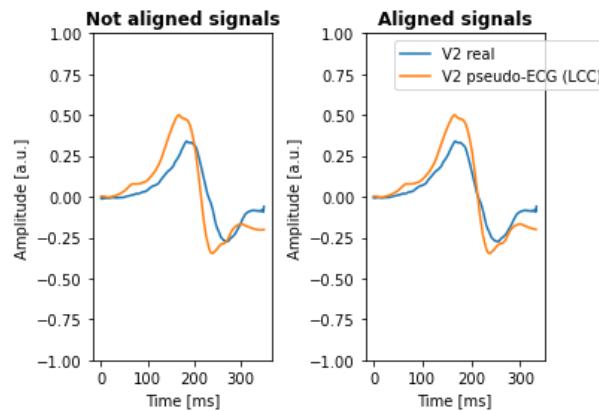


Figure 27. Representation of signal alignment at the point of maximum correlation.

The Fisher transform<sup>38</sup> ( $z'$ ) was used in order to average 12-lead values for each SOO, since this method has been previously used to analyze ECGs<sup>39</sup>. It was computed for each lead using Eq. 1, obtaining 12 values for each SOO, which were averaged ( $z'_\mu$ ). Eq. 2 was applied to the average  $z'$  obtaining the 12-lead correlation value. These transformations are used in order to lower the skew of the distribution.

$$z' = \frac{1}{2} \cdot \ln \left( \frac{1 + \rho}{1 - \rho} \right) \quad Eq. 1$$

$$\rho_{all} = \frac{e^{2z'\mu} + 1}{e^{2z'\mu} - 1} \quad Eq. 2$$

Results are summarized in Table 1 and 2. The predicted SOO and the distinction between a left or right OTVA origin is calculated using all ECG leads and only using V1 to V3 precordial leads. 12-lead predictions only predict the chamber origin in 3/11 cases (27.3%), while only using the three first precordial leads, the chamber origin is correctly predicted in 8/11 cases (72.7%). The three wrongly classified cases correspond to LVOT origins, which were classified as RVOT origins.

*Table 2: Results obtained when differentiating between a LVOT and RVOT origin, by computing the correlation using all ECG leads and using V1-V3 precordial leads.*

	Chamber origin	12-lead prediction	V1-V3 prediction
Patient 1	LV	RV	RV
Patient 2	LV	RV	<b>LV</b>
Patient 3	LV	<b>LV</b>	<b>LV</b>
Patient 4	LV	RV	<b>LV</b>
Patient 5	LV	RV	<b>LV</b>
Patient 6	RV	LV	<b>RV</b>
Patient 7	LV	RV	RV
Patient 8	LV	RV	RV
Patient 9	RV	<b>RV</b>	<b>RV</b>
Patient 10	RV	<b>RV</b>	<b>RV</b>
Patient 11	LV	RV	<b>LV</b>

Regarding results predicting the exact SOO, in RVOT origins, all predicted SOO from the RVOT are considered right since no specific differentiation within RVOT origins is found on the patient real data. And in LVOT origins, LCC and RCC SOO are the real patient SOO. As observed in Table 2, the predicted SOO from the LVOT are sometimes very close anatomically to the real SOO (as in patients 2 and 3).

*Table 3: Results obtained when predicting the exact SOO of the arrhythmia for each patient, by computing the correlation using all ECG leads and using V1-V3 precordial leads.*

	Exact SOO	12-lead prediction	V1-V3 prediction
Patient 1	LCC	SPA	SPA
Patient 2	LCC	SPM	<b>AMC2</b>
Patient 3	LCC	<b>AMC2</b>	<b>AMC2</b>
Patient 4	RCC	SPM	<b>RCC</b>
Patient 5	RCC	SPA	<b>AMC3</b>
Patient 6	RV	LCC	<b>SPA</b>
Patient 7	LCC	SPM	SPP
Patient 8	RCC	SPP	PL
Patient 9	RV	<b>SPP</b>	<b>PL</b>
Patient 10	RV	<b>PL</b>	<b>PL</b>
Patient 11	LCC	SPP	<b>LCC</b>

### 3.2 Machine learning algorithms for differentiating between LVOT or RVOT origin.

These algorithms were tested using the 11 ECG samples previously used in 3.1, which were quantitatively compared with its corresponding patient-specific generated pseudo-ECGs. Also, an ECG OTVA dataset provided by Hospital Teknon-Quirón in Barcelona, and Hospital Puerta del Mar in Cádiz consisting of 89 samples was used to evaluate performance of these algorithms as well.

#### 3.2.1 Results evaluating on the ECGs used for the modelling approach.

In this case, the same overall accuracy (73%) was obtained with all algorithms and training datasets, although some differences were observed between classes. Also, the labelled LVOT samples which were wrongly classified by SVM using dataset B and k-

NN using both training datasets, were the ones corresponding to patients 1, 6 and 7, which were the only ones wrongly classified using the in-silico pseudo-ECG based approach. Results are shown in Table 4.

*Table 4. Machine learning approach results. SVM: support vector machine, k-NN: k nearest neighbors, LVOT: left ventricle outflow tract, RVOT: right ventricle outflow tract.*

Algorithm	Dataset ID	Global accuracy	LVOT accuracy	RVOT accuracy
SVM	A	0.73	0.75	0.67
	B	0.73	0.62	1
k-NN	A	0.73	0.62	1
	B	0.73	0.62	1

### 3.2.2 Results evaluating on the new extracted ECG datasets.

Results showed that k-NN classification algorithm performed a more accurate classification of data, obtaining a global accuracy of 81% using dataset A as training data, and an accuracy of 82% using as training dataset B. On the other hand, using SVM classifier, the global accuracy for dataset A is 53%, and 74% for dataset B. The testing dataset contained 89 samples, from which 57 correspond to a RVOT origin, and 31 to a LVOT origin.

*Table 5: Machine learning approach results. SVM: support vector machine, k-NN: k nearest neighbors, LVOT: left ventricle outflow tract, RVOT: right ventricle outflow tract.*

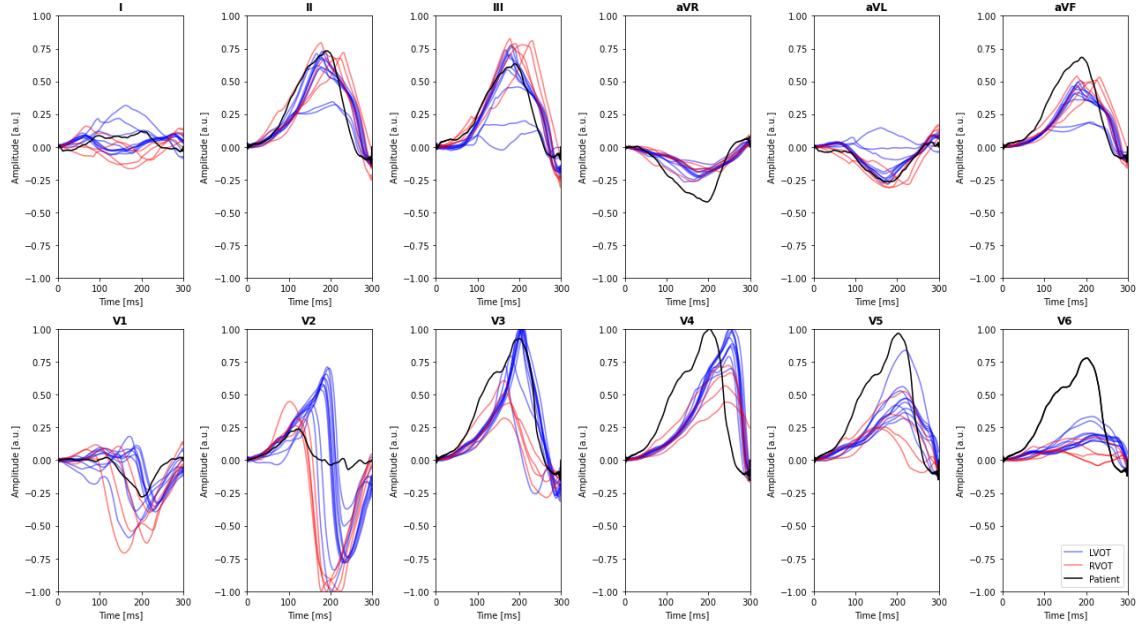
Algorithm	Dataset ID	Global accuracy	LVOT accuracy	RVOT accuracy
SVM	A	0.53	0.94	0.32
	B	0.74	0.9	0.65
k-NN	A	0.81	0.74	0.84
	B	<b>0.82</b>	0.71	0.88

Moreover, by analyzing the prediction accuracy for each class, the SVM classifier showed a higher rate of wrong predictions when testing RVOT origin signals, and high accuracy

when predicting LVOT origins, since it almost always predicted a LVOT origin (especially when using dataset A for training), while using k-NN algorithm, a more balanced accuracy is found when comparing both classes accuracies.

### 3.3 Prospective evaluation of the machine learning and the in-silico pace-mapping pseudo-ECG based pipelines.

In this case, the SOO of the OTVA was correctly predicted with the previously mentioned correlation techniques. The simulated pseudo-ECGs and the real patient ECG are displayed in Figure 28. Also, the ML algorithm considered best correctly predicted a LVOT origin of the OTVA.



*Figure 28. Comparison of 12-lead LVOT (blue) against RVOT (red) origin generated pseudo-ECGs, against real patient ECG (black) diagnosed with a LCC LVOT origin. LVOT: right ventricle outflow tract, RVOT: right ventricle outflow tract, LCC: left coronary cusp.*

## 4 Discussion

Simulated pseudo-ECGs have shown clear differences between right and left sided OTVA origins, especially in the 3 first precordial leads. Deeper S waves are found on simulated pseudo-ECGs from RVOT origins, while taller R waves are found on those corresponding to LVOT origins. The three first precordial leads have been a useful tool to predict the exact location of the arrhythmia, correctly classifying the chamber origin in 8/11 cases. Also, in this work, the QRS complex of pseudo-ECGs and real ECGs has been automatically extracted using DL techniques, which ensures unbiased quantitative analysis, while this was manually performed in the previous work.

Comparing these results with the ones obtained by Doste et al.<sup>1</sup>, lower classification accuracy is reported in this work, since in Doste's work<sup>1</sup> the ventricular origin was correctly predicted in 10/11 cases in the quantitative analysis of the simulated ECGs against the real patient ECGs. Using the same generic torso model for all patients and considering constant conductivity while ignoring other torso organs are key simplifications which introduce errors in the computation, while simplifying the pipeline. Remind that in Moya's bachelor thesis<sup>2</sup>, the quantitative pseudo-ECG analysis was only performed in 2 patients, so further pseudo-ECG validation has been one of the main objectives of this work.

Some key features found in the literature from specific SOO have been observed in the simulations, as discussed in section 3.1.2, although some of them have not been accomplished. Similar conclusions were extracted in Doste's work<sup>1</sup>, where specific features from more detailed SOO were not observable (AMC, LV summit origins...).

And regarding the ML techniques, k-NN algorithm has obtained highly better accuracy results than SVM using both training datasets. Also, SVM seemed to be biased since when using dataset A as training data since it almost always predicted a LVOT origin of the OTVA. The combination of fully simulated ECGs or pseudo-ECGs with real ECGs has allowed to train larger datasets, showing a satisfactory model generalization when testing the algorithms with real ECG data from clinical databases.

For future improvement, it would be very interesting to perform experiments using ML not only predicting the ventricular chamber origin, but the exact SOO. The closer SOO could be grouped into clusters for simplification. Also, a very interesting idea would be to normalize the precordial transition of the OTVA with respect to the precordial transition of the patient in sinus rhythm, as it has been found in the literature<sup>18</sup>.

It is important to note that the patients' ECGs which were wrongly classified using the ML algorithms were the same which were wrongly predicted during the quantitative comparison of the pseudo-ECGs with the real ECGs (patient 1, 7 and 8). In fact, in these 3 cases, a RVOT is predicted when the samples are labelled with a LVOT origin. However, by analyzing the precordial transition in these samples, a late precordial transition (in V4) is found in patients 1 and 8, while this is a characteristic feature of RV OTVA origins. And in patient 7, whose real label is LCC, a precordial transition in V3 and a LBBB pattern are observed, when characteristic features from LCC origins are an early precordial transition (V1-V2) and an RBBB pattern.

It is very likely too that some of the samples of the testing dataset were wrongly labelled, due to human error, since it is easy to make mistakes during the whole process of extracting, anonymizing, pre-processing and labelling data.

And finally, it is clear that the ML approach is much easier to implement into the clinical workflow, although further experiments must be performed to get more robust results. The pseudo-ECG based in-silico pace mapping approach is a very powerful tool and the obtained pseudo-ECGs have shown promising results. However, it takes a large amount of time to get the simulation output and although some steps such as the segmentation could be sped up using for example a neural network, it requires a lot of time and practice to get reproducible results. In this work, an attempt to shorten the segmentation time using *Adas*<sup>35</sup> was performed, although it was not possible. The fact that the patient's CT images are acquired shortly before the intervention makes this pipeline impossible to be implemented at least for now in a prospective way. However, it would be very useful to use this approach to generate more simulated pseudo-ECGs to create a bigger pseudo-ECG database, like Doste did in his PhD<sup>1</sup>, and to use the new obtained from Hospital Teknon-Quirón and Puerta del Mar to enlarge the real ECG database.

## 5 Conclusions

The main objective of this work is to analyze how these previously developed pipelines can be adapted to move a step forward and implement them in the clinical workflow as a clinical decision support tool. After replicating the pipeline, adding some modifications and analyzing the results, the following conclusions are drawn:

Generating simulated pseudo-ECGs is a very useful technique to generate in a simplified manner patient-specific simulated ECGs. Clear differences have been found between simulated pseudo-ECGs from LVOT and RVOT origins. Considering more specific SOO, some of the key characteristic features found in the literature are appreciated in the simulations. However, less accurate results have been obtained compared to Doste's results<sup>1</sup>, who obtained fully simulated ECGs in his PhD using Elvira software. Nonetheless, these generated pseudo-ECGs can be very useful to train ML or more sophisticated DL algorithms to distinguish between different OTVA origins, as strong generalization has been observed when validating the ML algorithms on the obtained real-world OTVA ECG datasets.

Working closer to the clinical workflow has been one of the main challenges of this work, since communication with clinicians during the patient's intervention is not easy, and long computational acquisition time is impossible to fit within the intervention time at least in these conditions. Nevertheless, the proposed ML algorithms are likely to be included into the clinical workflow due to its short prediction time.

Finally, it is concluded that the combination of ML or DL algorithms can be a very promising approach if they are combined within computational electrophysiological simulations, allowing to shorten the total pipeline time and simplifying some of the manual tasks in order to easily fit in a clinical environment.



## 6 Bibliography

1. Doste Beltrán, R. Computational models of the heart for planning and treatment of outflow tract ventricular arrhythmias. *TDX (Tesis Dr. en Xarxa)* (2020).
2. Moya, N. Design and implementation of a medical education module on in-silico simulations using high-performance computing. (2020).
3. Park, K. M., Kim, Y. H. & Marchlinski, F. E. Using the surface electrocardiogram to localize the origin of idiopathic ventricular tachycardia. *PACE - Pacing Clin. Electrophysiol.* **35**, 1516–1527 (2012).
4. Yamada, T. Twelve-lead electrocardiographic localization of idiopathic premature ventricular contraction origins. *J. Cardiovasc. Electrophysiol.* **30**, 2603–2617 (2019).
5. Hutchinson, M. D. & Garcia, F. C. An organized approach to the localization, mapping, and ablation of outflow tract ventricular arrhythmias. *J. Cardiovasc. Electrophysiol.* **24**, 1189–1197 (2013).
6. Anderson, R. H., Yanni, J., Boyett, M. R., Chandler, N. J. & Dobrzynski, H. The anatomy of the cardiac conduction system. *Clin. Anat.* **22**, 99–113 (2009).
7. Fu, D. guan. Cardiac Arrhythmias: Diagnosis, Symptoms, and Treatments. *Cell Biochem. Biophys.* **73**, 291–296 (2015).
8. Sánchez-quintana, D., Doblado-calatrava, M., Cabrera, J. A., Macías, Y. & Saremi, F. Interventional Electrophysiologist. **2015**, (2015).
9. Stephenson, R. S. *et al.* The functional architecture of skeletal compared to cardiac musculature: Myocyte orientation, lamellar unit morphology, and the helical ventricular myocardial band. *Clin. Anat.* **29**, 316–332 (2016).
10. Whittaker, D. G., Benson, A. P., Teh, I., Schneider, J. E. & Colman, M. A. Investigation of the Role of Myocyte Orientations in Cardiac Arrhythmia Using Image-Based Models. *Biophys. J.* **117**, 2396–2408 (2019).
11. Penela, D. *et al.* An easy-to-use, operator-independent, clinical model to predict the left vs. right ventricular outflow tract origin of ventricular arrhythmias. *Europace* **17**, 1122–1128 (2015).

12. Doste, R. *et al.* In silico pace-mapping: Prediction of left vs. right outflow tract origin in idiopathic ventricular arrhythmias with patient-specific electrophysiological simulations. *Europace* **22**, 1419–1430 (2020).
13. Kim, R. J. *et al.* Clinical and Electrophysiological Spectrum of Idiopathic Ventricular Outflow Tract Arrhythmias. *J. Am. Coll. Cardiol.* **49**, 2035–2043 (2007).
14. Liang, J. J., Shirai, Y., Lin, A. & Dixit, S. Idiopathic outflow tract ventricular arrhythmia ablation: Pearls and pitfalls. *Arrhythmia Electrophysiol. Rev.* **8**, 116–121 (2019).
15. Lavalle, C. *et al.* Electrocardiographic features, mapping and ablation of idiopathic outflow tract ventricular arrhythmias. *J. Interv. Card. Electrophysiol.* **57**, 207–218 (2020).
16. Zheng, J., Fu, G., Anderson, K., Chu, H. & Rakovski, C. A 12-Lead ECG database to identify origins of idiopathic ventricular arrhythmia containing 334 patients. *Sci. Data* **7**, 1–11 (2020).
17. Alcaine, A. *et al.* Automatic activation mapping and origin identification of idiopathic outflow tract ventricular arrhythmias. *J. Electrocardiol.* **51**, 239–246 (2018).
18. Betensky, B. P. *et al.* The V2 transition ratio: A new electrocardiographic criterion for distinguishing left from right ventricular outflow tract tachycardia origin. *J. Am. Coll. Cardiol.* **57**, 2255–2262 (2011).
19. Korshunov, V. *et al.* Prediction of premature ventricular complex origin in left vs. right ventricular outflow tract: A novel anatomical imaging approach. *Europace* **21**, 147–153 (2019).
20. Anderson, R. D. *et al.* Differentiating Right-and Left-Sided Outflow Tract Ventricular Arrhythmias: Classical ECG Signatures and Prediction Algorithms. *Circ. Arrhythmia Electrophysiol.* **12**, 1–15 (2019).
21. Zhou, X., Fang, L., Wang, Z., Liu, H. & Mao, W. Comparative analysis of electrocardiographic imaging and ECG in predicting the origin of outflow tract ventricular arrhythmias. *J. Int. Med. Res.* **48**, (2020).

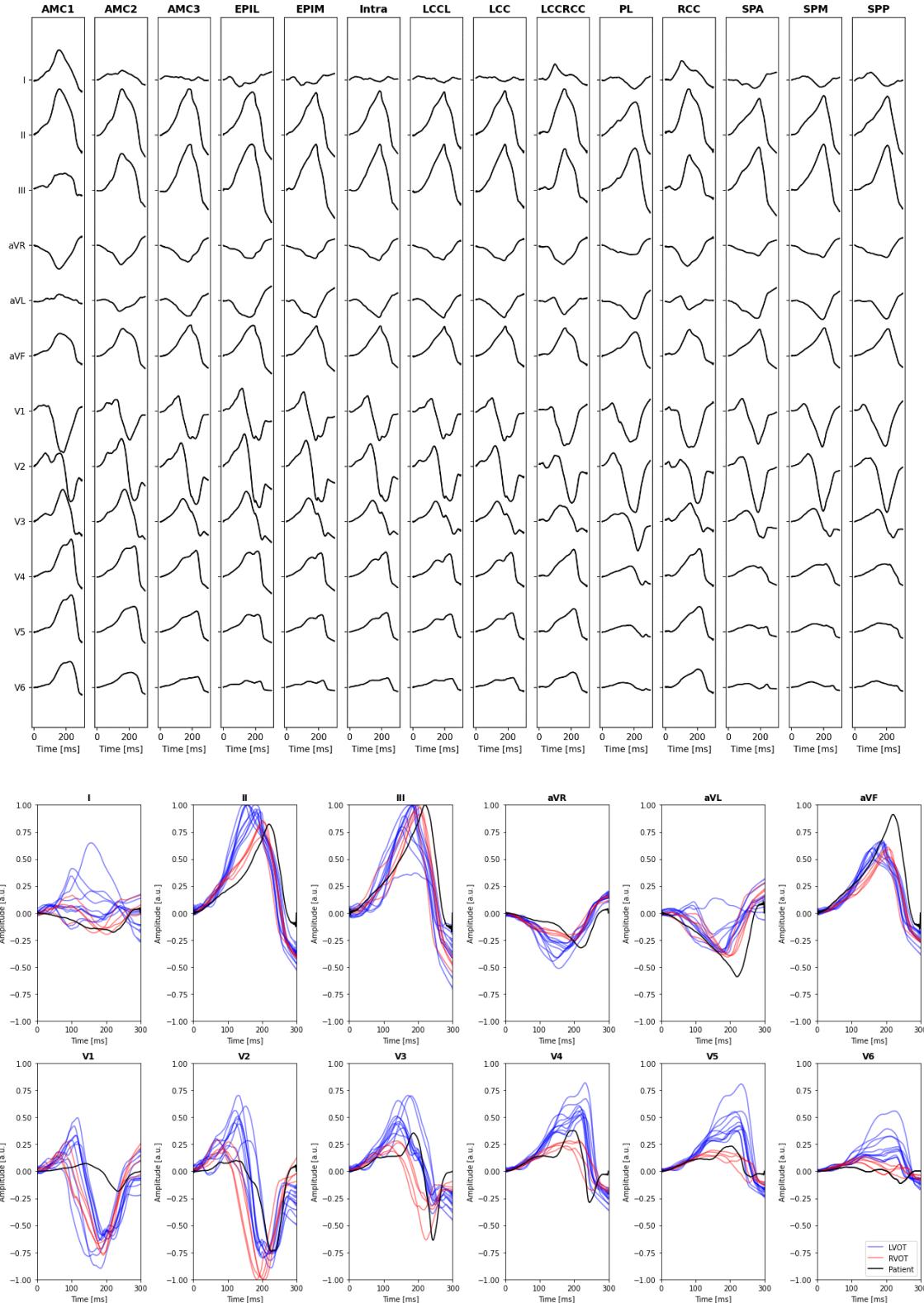
22. Zhang, F. *et al.* Electrocardiographic algorithm to identify the optimal target ablation site for idiopathic right ventricular outflow tract ventricular premature contraction. *Europace* **11**, 1214–1220 (2009).
23. Soucek, F. & Starek, Z. Use of Bipolar Radiofrequency Catheter Ablation in the Treatment of Cardiac Arrhythmias. *Curr. Cardiol. Rev.* **14**, 185–191 (2018).
24. Bhakta, D. & Miller, J. M. Principles of electroanatomic mapping. *Indian Pacing Electrophysiol. J.* **8**, 32–50 (2008).
25. Doste, R. *et al.* A rule-based method to model myocardial fiber orientation in cardiac biventricular geometries with outflow tracts. *Int. j. numer. method. biomed. eng.* **35**, 1–17 (2019).
26. Sacco, F. Quantification of the influence of detailed endocardial structures on human cardiac haemodynamics and electrophysiology using HPC. *TDX (Tesis Dr. en Xarxa)* (2019).
27. Schwarz, J. S., Chapman, C. & McDonnell Feit, E. Welcome to Python. *Python for Marketing Research and Analytics* 3–7 (2020) doi:10.1007/978-3-030-49720-0\_1.
28. 3D Slicer. 3D Slicer Image Computing Platform. <https://www.slicer.org/> (2012).
29. Blender. Blender. <https://www.blender.org/> (2002).
30. Paraview. ParaView. <https://www.paraview.org/> (2005).
31. MathWorks. MATLAB - El lenguaje del cálculo técnico - MATLAB & Simulink. *Natick, Massachusetts* 2 (2018).
32. O’Hara, T., Virág, L., Varró, A. & Rudy, Y. Simulation of the undiseased human cardiac ventricular action potential: Model formulation and experimental validation. *PLoS Comput. Biol.* **7**, (2011).
33. Stinstra, J. G., Hoppenfeld, B. & MacLeod, R. S. On the passive cardiac conductivity. *Ann. Biomed. Eng.* **33**, 1743–1751 (2005).
34. Hasgall, P.A., Neufeld, E., Gosselin, C., Klingenberg, M.A., Kuster, N. DATABASE » IT’IS Foundation. *IT’IS Database for thermal and electromagnetic parameters of biological tissues* (2015).

35. ADAS 3D Fibrosis Imaging for the EP Lab.  
<https://www.galgomedical.com/en/adas3d.html> (2017).
36. Jimenez-Perez, G., Alcaine, A. & Camara, O. U-Net Architecture for the Automatic Detection and Delineation of the Electrocardiogram. *2019 Comput. Cardiol. Conf.* **45**, 2019–2022 (2019).
37. Mariani, M. V. *et al.* Electrocardiographic criteria for differentiating left from right idiopathic outflow tract ventricular arrhythmias. *Arrhythmia Electrophysiol. Rev.* **10**, 10–16 (2021).
38. Corey, D. M. & Burke, M. Averaging Correlations : Expected Values and Bias in Combined Pearson rs and Fisher ' s z Transformations. (2014)  
doi:10.1080/00221309809595548.
39. Okada, J., Sasaki, T., Washio, T., Hisada, T. & Sugiura, S. Patient Specific Simulation of Body Surface ECG. **36**, 309–321 (2013).
40. Ganesan, P., Sterling, M., Ladavich, S. & Ghoraani, B. Computer-Aided Clinical Decision Support Systems for Atrial Fibrillation. *Comput. Technol. - Appl. Eng. Med.* (2016) doi:10.5772/65620.

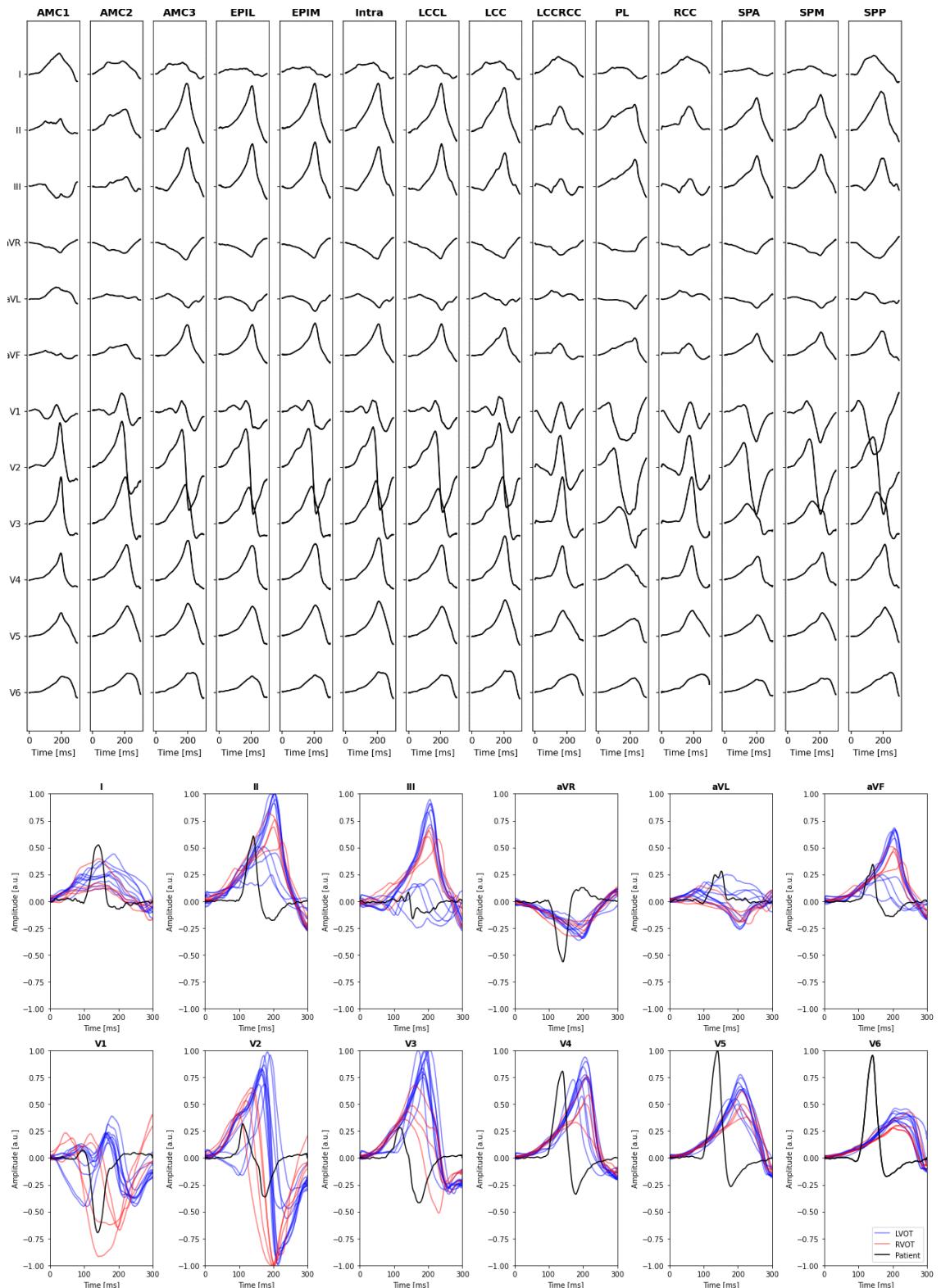


## Appendix: Simulated pseudo-ECGs.

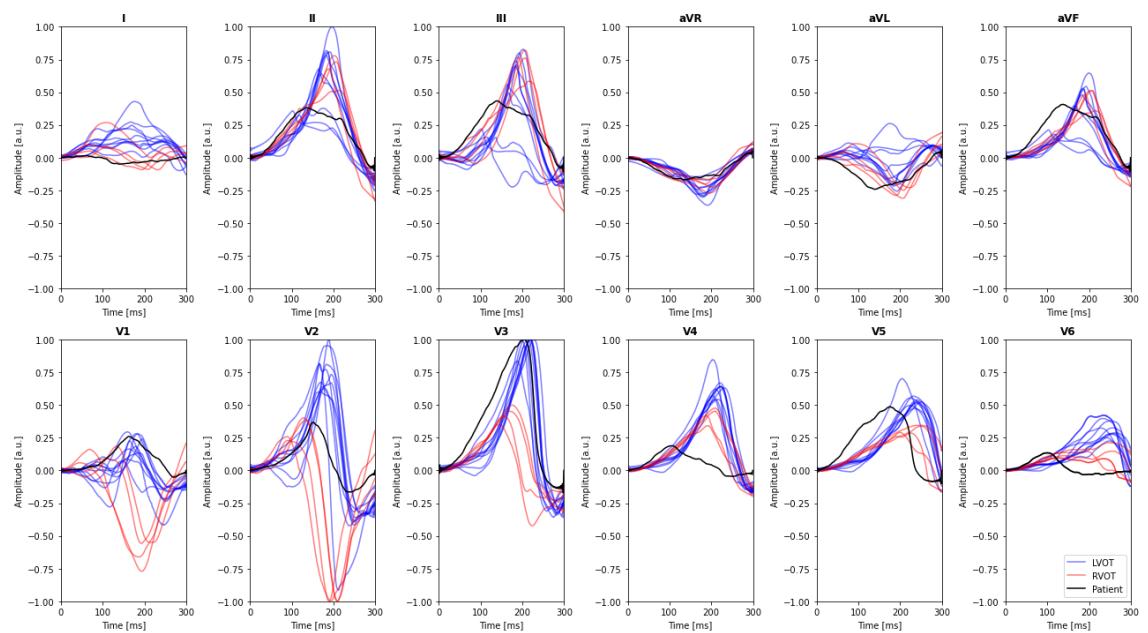
### Patient 1. Real SOO: LCC



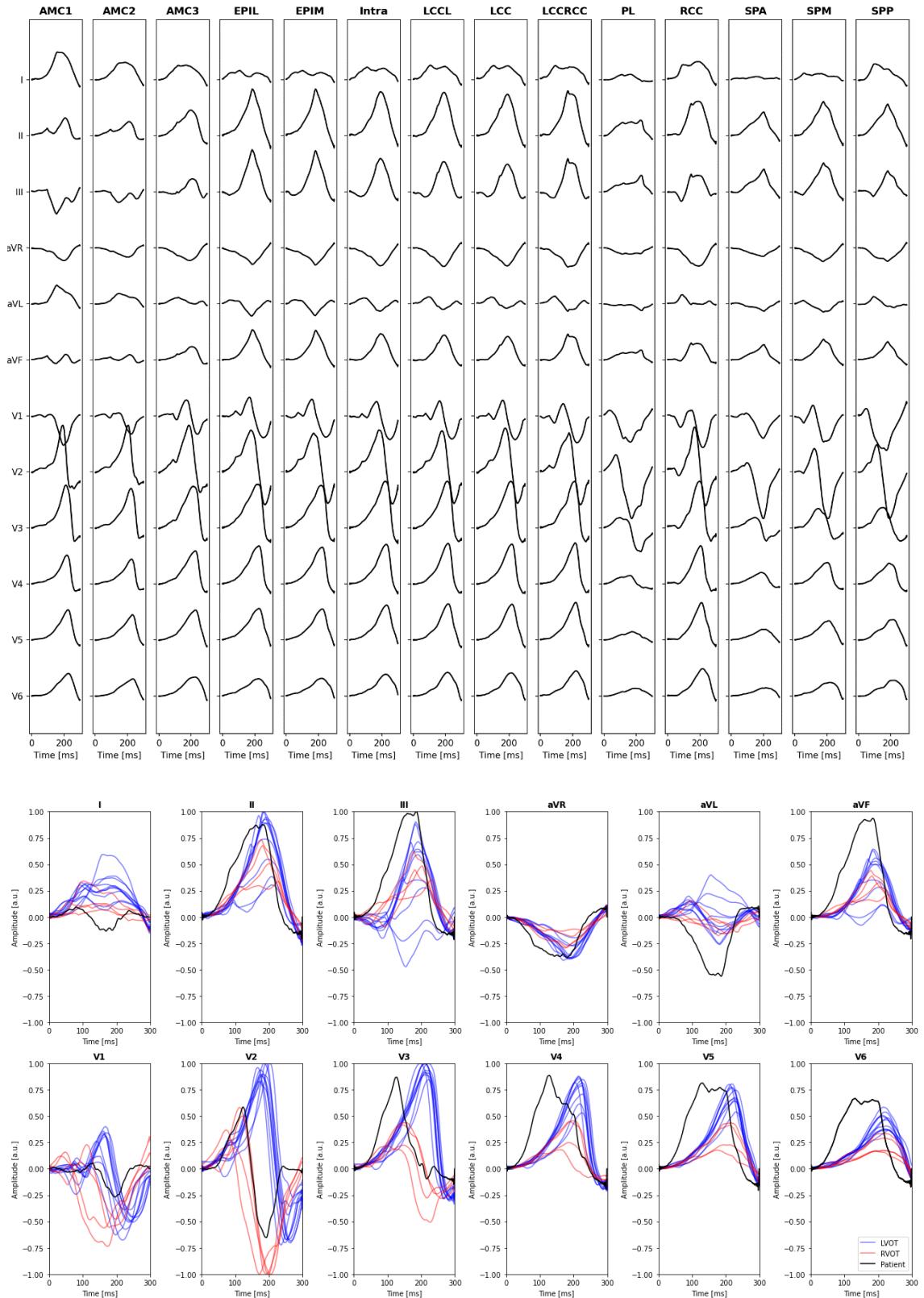
## Patient 2. Real SOO: LCC



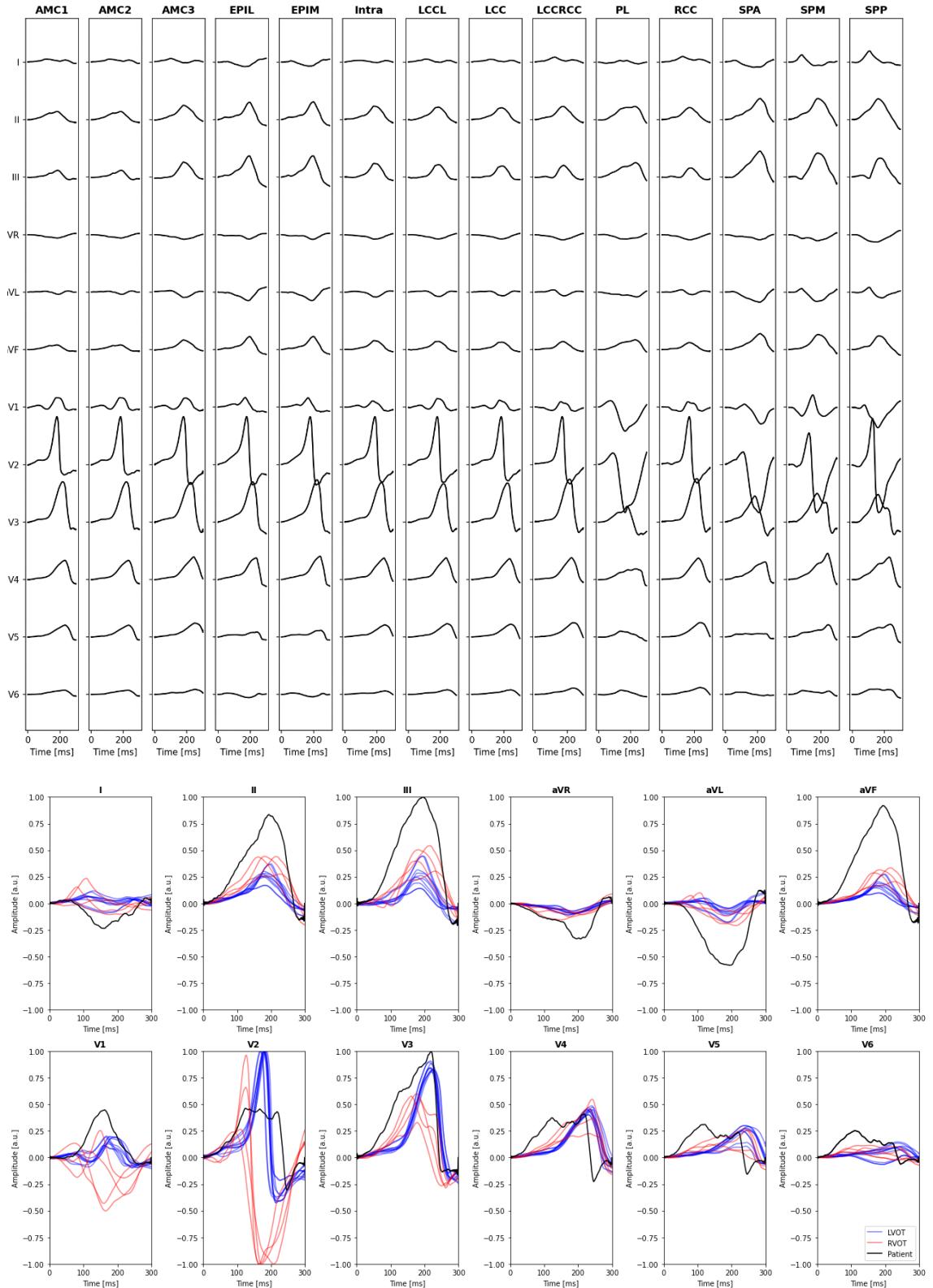
### Patient 3. Real SOO: LCC



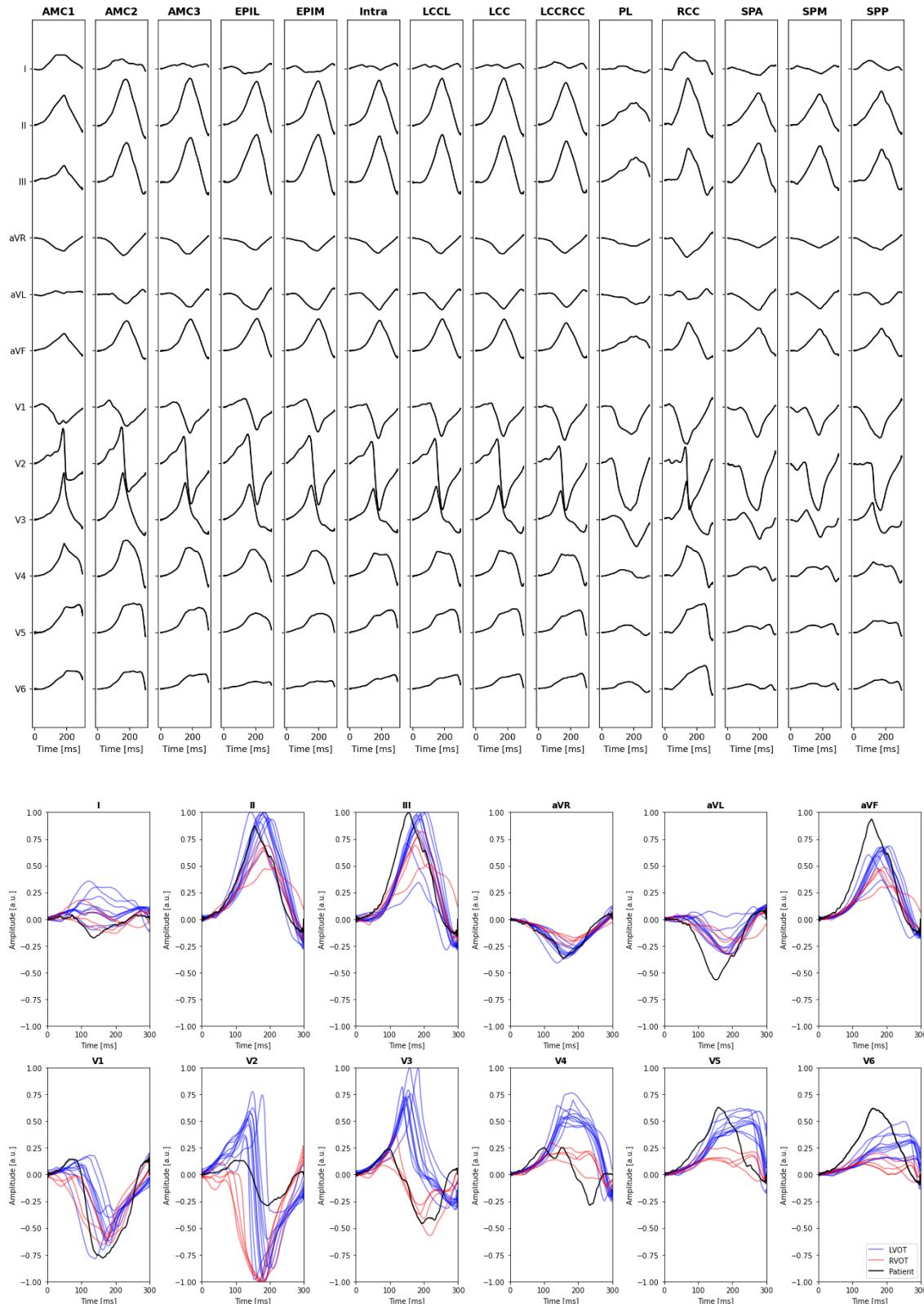
## Patient 4. Real SOO: RCC



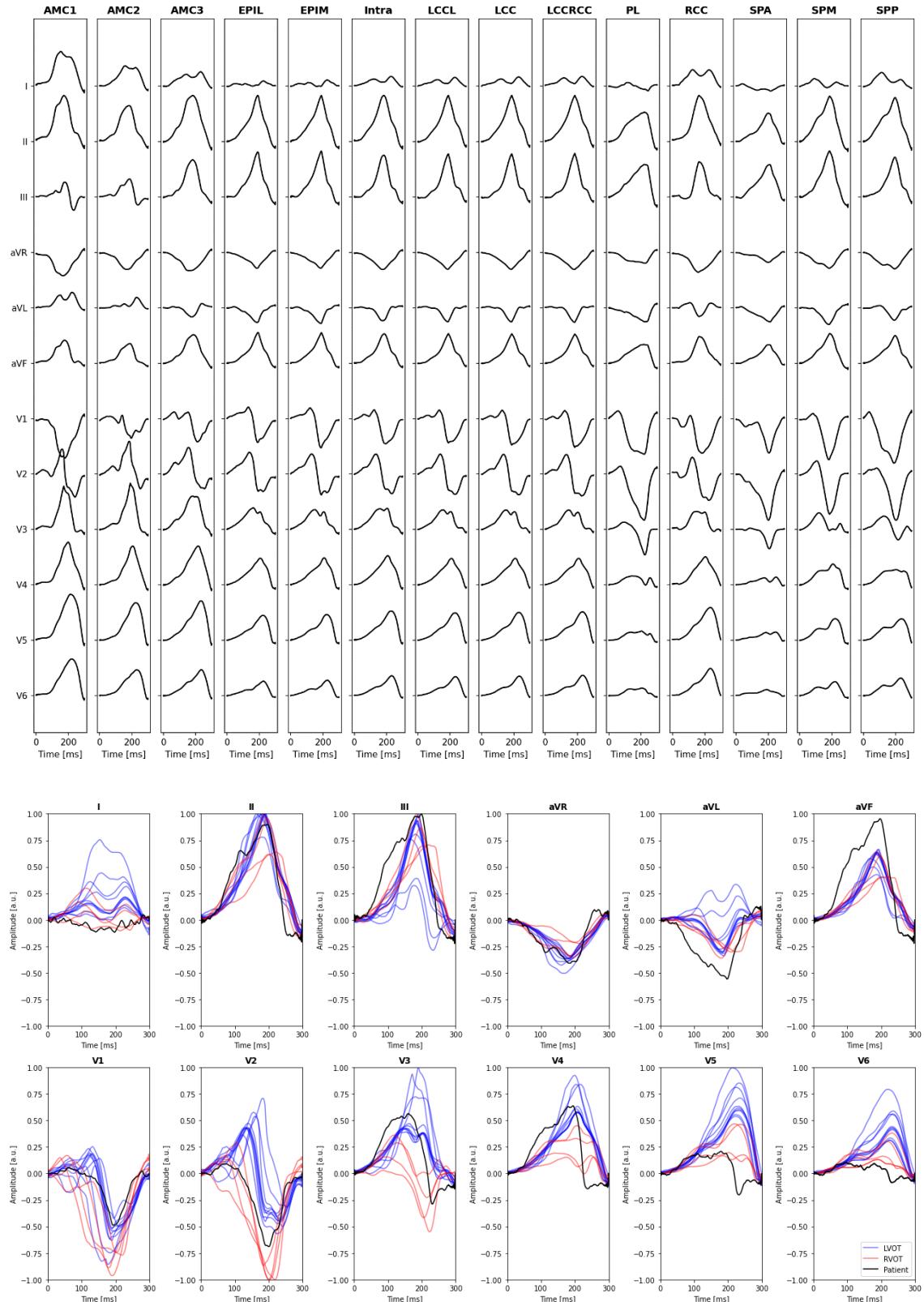
## Patient 5. Real SOO: RCC



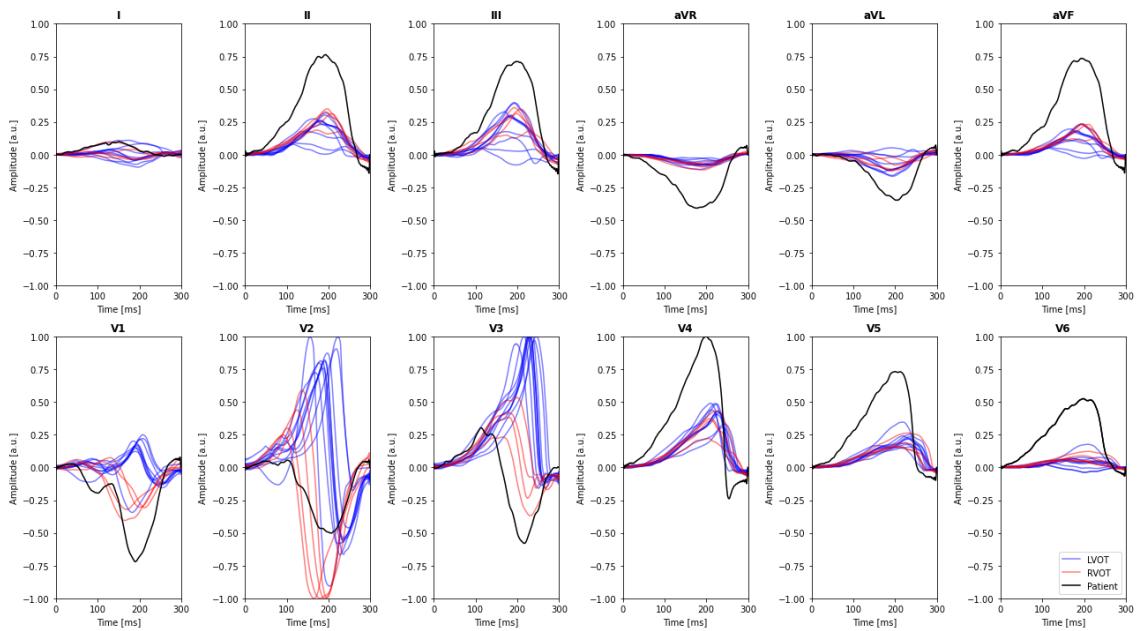
## Patient 6. Real SOO: RV



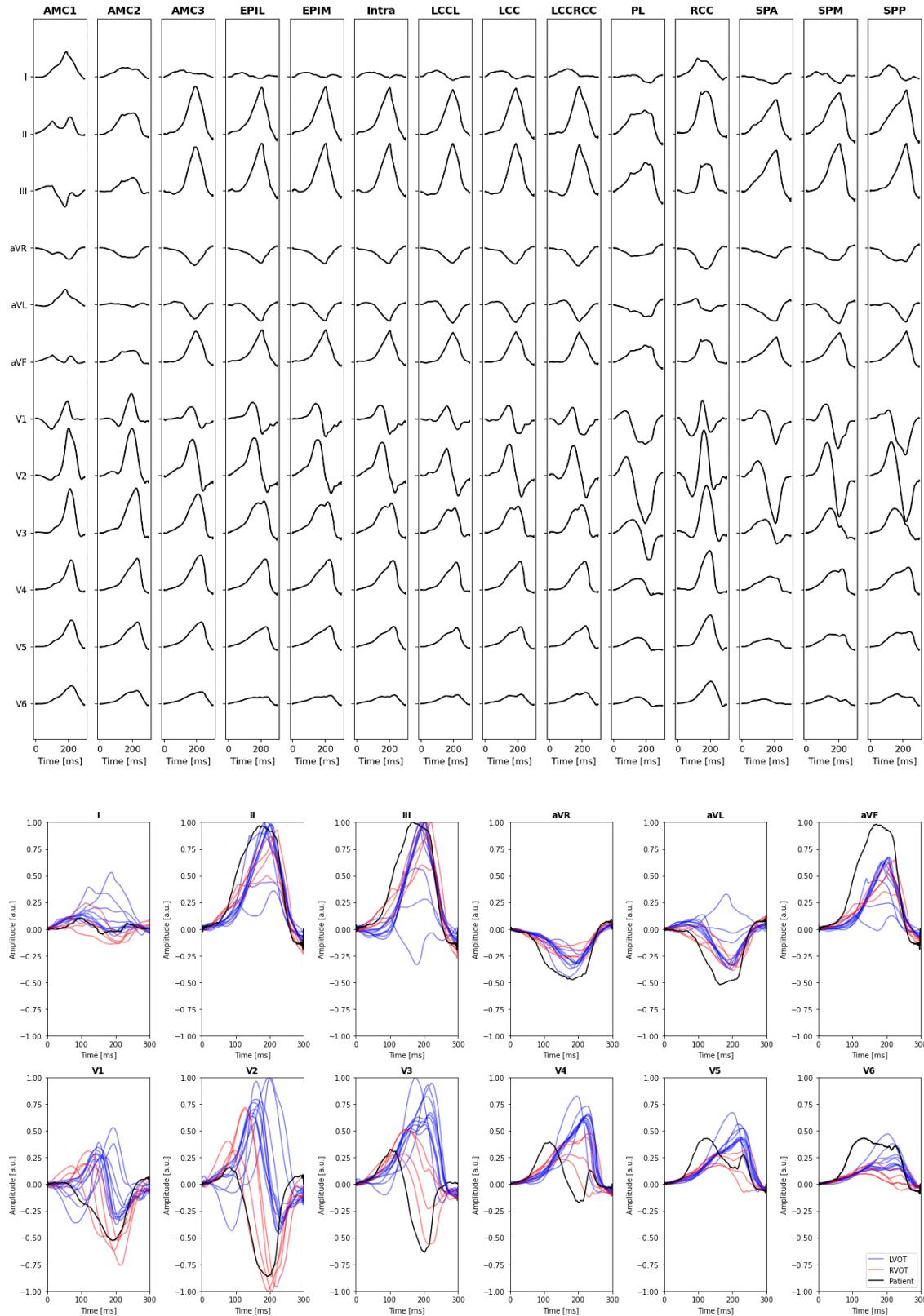
## Patient 7. Real SOO: LCC



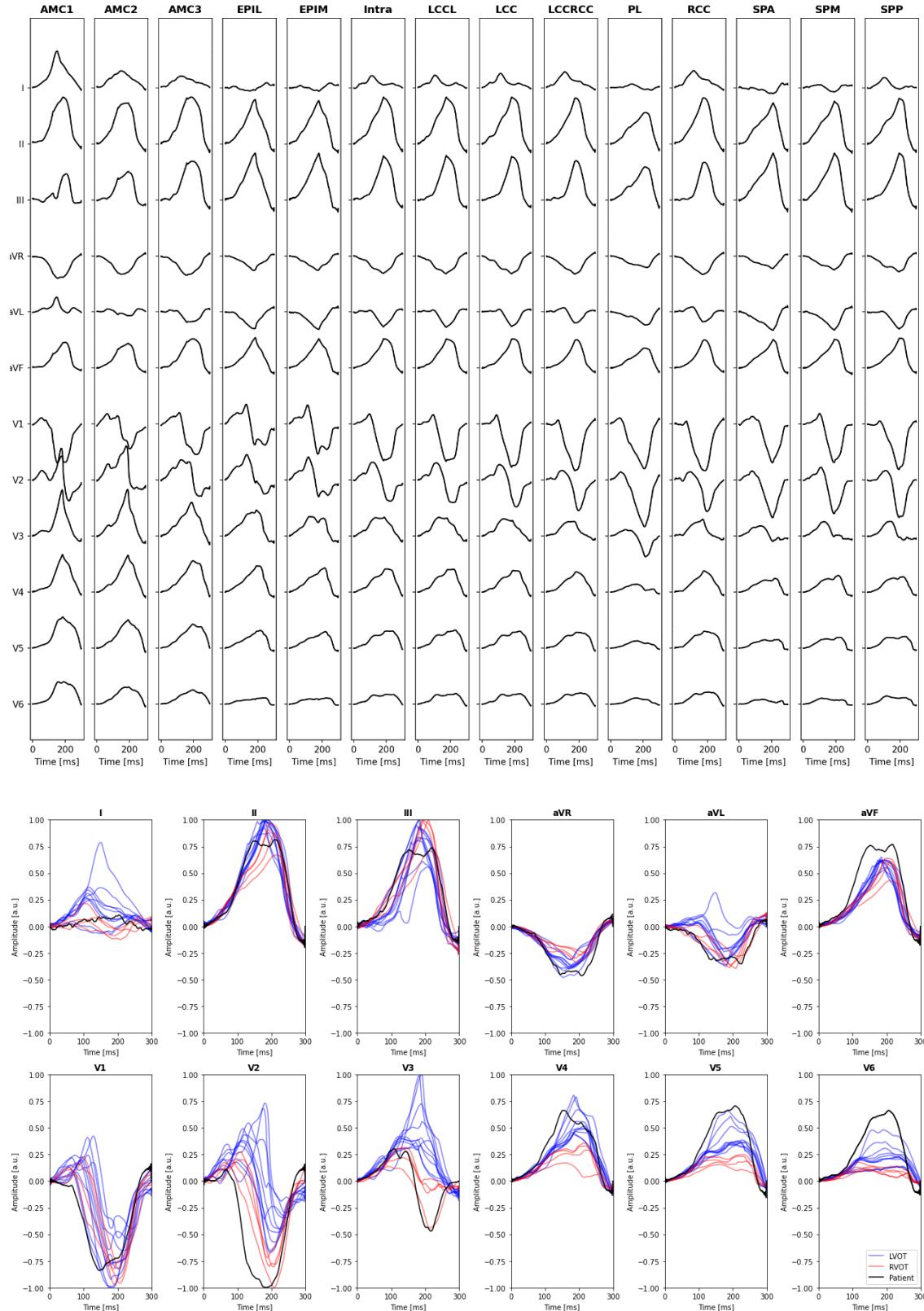
## Patient 8. Real SOO: RCC



## Patient 9. Real SOO: RV



## Patient 10. Real SOO: RV



## Patient 11. Real SOO: LCC

