



Population-based personalization of geometric models of myocardial infarction

Kannara Mom, Patrick Clarysse, Nicolas Duchateau

► To cite this version:

Kannara Mom, Patrick Clarysse, Nicolas Duchateau. Population-based personalization of geometric models of myocardial infarction. Functional Imaging and Modelling of the Heart, 2021, Stanford, United States. In press. hal-03203372

HAL Id: hal-03203372

<https://hal.science/hal-03203372>

Submitted on 20 Apr 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Population-based personalization of geometric models of myocardial infarction

Kannara Mom¹, Patrick Clarysse¹, and Nicolas Duchateau¹

Univ Lyon, Université Claude Bernard Lyon 1, INSA-Lyon, CNRS, Inserm,
CREATIS UMR 5220, U1294, F-69621, LYON, France
`nicolas.duchateau@creatis.insa-lyon.fr`

Abstract. We propose a strategy to perform population-based personalization of a model, to overcome the limits of case-based personalization for generating virtual populations from models that include randomness. We formulate the problem as matching the synthetic and real populations by minimizing the Kullback-Leibler divergence between their distributions. As an analytical formulation of the models is complex or even impossible, the personalization is addressed by a gradient-free method: the CMA-ES algorithm, whose relevance was demonstrated for the case-based personalization of complex biomechanical cardiac models. The algorithm iteratively adapts the covariance matrix which in our problem encodes the distribution of the synthetic data.

We demonstrate the feasibility of this approach on two simple geometrical models of myocardial infarction, in 2D, to better focus on the relevance of the personalization process. Our strategy is able to reproduce the distribution of 2D myocardial infarcts from the segmented late Gadolinium images of 123 subjects with acute myocardial infarction.

Keywords: Model personalization ; myocardial infarction ; late Gadolinium enhancement ; cardiac magnetic resonance.

1 Introduction

The interest of developing realistic simulations of myocardial infarction is undeniable for understanding and validation purposes [1]. As for many biophysical models, personalization is a necessary milestone for the realism of such simulations. The standard personalization approach consists in adjusting the model to the data of a given individual. In order to generate large virtual populations that can for example feed machine learning algorithms, the personalization process should be extended beyond the fit to individuals' data. In this sense, two approaches may be adopted:

- *Case-based* personalization, which consists in finding the optimal model parameters for each individual, from which a range of relevant values can be determined a-posteriori. A synthetic population can therefore be generated by randomly sampling within this range and running the corresponding simulations. This approach is suited for fully deterministic models, but cannot

cover models with randomness inside (e.g. generating a lesion at a random location or of random shape).

- *Population-based* personalization (the approach we propose), which can overcome the latter issue by finding the optimal model parameters that best match a virtual population to the real population under study. This strategy is somehow comparable to the matching of distributions pursued in variational auto-encoders [2], except that in our case the virtual population is generated from the personalized model parameters, not from a related latent space that encodes the data.

In this context, we propose a strategy to perform population-based personalization of a model, which we illustrate on simple geometrical models of myocardial infarction.

Tissue-level geometrical models of the lesions have been proposed based on a regional prior about a given coronary territory [3–5] or even up to mimicking the wavefront phenomenon [6] for the lesion propagation around an existing coronary segmentation [7]. Here we rely on two very simple models whose output is rather straightforward to visualize and assess, in 2D, so that we can better focus on the relevance of the personalization process: one that iteratively models an infarct as the union of spheres of random size [4], and one that uses an ellipsoid centered on the endocardium to represent the infarct [8].

Our primary objective is to demonstrate the feasibility of the population-based personalization. In addition, our secondary objective is to state on the relevance of these geometrical models of myocardial infarction to mimic a real population of patients with acute infarcts.

2 Methods

2.1 Data and pre-processing

The population under study consisted of 123 subjects for which late Gadolinium enhancement images were available and segmented, distributed into 45, 17, and 61 cases for which the LAD, LCX, or RCA coronary arteries were responsible of the infarct. These data came from the Minimalist Immediate Mechanical Intervention (MIMI) study [9], which consists of acute ST-Elevation Myocardial Infarction (STEMI) patients who underwent either immediate or delayed stenting. Cardiac magnetic resonance was performed 5 days (interquartile range 4-6) after inclusion with Avanto 1.5T systems (Siemens, Erlangen, Germany). The infarct patterns were derived from the late Gadolinium enhancement images, performed 10 minutes after bolus injection, with an inversion time around 240-280 ms. The myocardial contours were manually segmented offline by consensus reading of three experienced observers using commercial software (CVI42 v.5.1.0 Circle Cardiovascular Imaging, Calgary, CA). The LV ranged over 17 ± 2 slices. The infarct zone was determined semi-automatically using the full-width half-maximum (FWHM) method. All contours were controlled and corrected manually if needed.

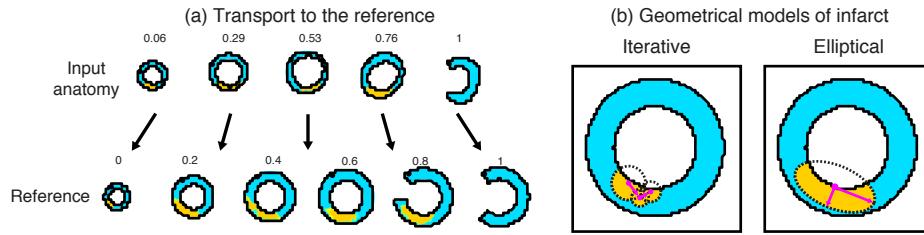


Fig. 1: (a) Transport of the data from an individual to the reference anatomy, illustrated on a subset of slices. The black dot points out the LV-RV junction. (b) Schematic view of the two simple geometrical models of infarct used in this paper.

Radial, circumferential, and long-axis coordinates ranging from 0 to 1 were automatically defined on each stack of slices, after manually identifying the LV-RV junction on each slice and the myocardial borders around the LV outflow tract, when relevant. For this automatic parameterization of the LV geometry, we used images upsampled by a factor of 4 to prevent artifacts that may occur with few pixels covering the myocardium, in particular along the radial direction.

The lesion segmentations were transported to a reference anatomy (defined as a semi-ellipsoid with maximal endocardial and epicardial radii of respectively 30 and 50 pixels, represented as a stack of 21 slices of 80x80 pixels each), using linear interpolation tailored for data defined on a scattered grid (Fig.1a).

Finally, we rotated all infarct patterns along the circumference such that the centers of mass of the infarcts for each coronary territory are aligned to the LAD one. This allows better focusing on the infarct shape and extent, and lowering the effects of the infarct localization without being restricted to a given coronary territory. Although this is arguable, as wall characteristics and infarct patterns may differ across coronary territories, it allowed us to demonstrate the feasibility of the personalization process on large enough populations.

2.2 Geometric models of myocardial infarction

We focused on two geometrical models that generate synthetic lesions of varying shape, extent, and location. Each model is parameterized by a starting point on the endocardial surface, which conditions the infarct location, and other parameters that encode the shape and extent of the lesion (Fig.1b):

- The first model approximates a random infarct shape as the union of several spheres of random sizes (intersected with the myocardium) [4]. The first sphere is centered on a given endocardial point, and the model generates at each iteration a new sphere centered on a random point of the previous sphere. This model requires setting the total number of iterations and the maximal radius of the spheres (2 parameters).
- The second model generates an ellipsoid centered at a given endocardial point, and sets the infarcted region as the intersection between the ellipsoid and the myocardium [8]. It requires setting the short- and long-axes of the ellipsoid, which respectively lie on the radial and circumferential/longitudinal

directions of the endocardium. Randomness is introduced in this model by sampling the short- and long-axis values using a uniform distribution (4 parameters: for each axis, the distribution center and its extension).

To better initialize the models and limit the number of parameters to optimize, the starting point was sampled within a small zone around the center of mass of the infarcts obtained from the real subjects, for both models.

In 3D, the synthetic lesions are generated on a template mesh that corresponds to the reference anatomy used to align the real image data (Sec.2.1). Synthetic slices are then estimated by resampling the mesh data on an image grid. In this paper, to better focus on the personalization process within a reasonable computing time, we focused on a 2D version of the geometrical models at the mid-level of the left ventricle.

2.3 Personalization of the models

The personalization relied on the CMA-ES algorithm (Covariance Matrix Adaptation – Evolution Strategy) [10], a method that consists in adapting across iterations (referred to as generations) the synthetic data distribution parameterized by its covariance matrix. It uses a stochastic search that retains at each iteration a subset of the best cases (in the sense of the energy to minimize) sampled from a multi-normal distribution, and updates the covariance matrix of the distribution with such samples.

CMA-ES is a gradient-free method whose relevance was demonstrated for the personalization of complex models as in cardiac electromechanical simulations [11]. Besides, this generic personalization strategy allows to remain non-specific to a given model.

Let's denote $\{\mathbf{x}_1, \dots, \mathbf{x}_p\}$ and $\{\mathbf{y}_1, \dots, \mathbf{y}_q\}$ the sets of synthetic and real infarct images, represented by the distributions P and Q , respectively.

The *case-based personalization* means finding for each $i \in [1, q]$ the optimal parameters $\theta_i \in \mathbb{R}^{m_d}$ such that $\mathbf{x}_i = f_d(\theta_i) \approx \mathbf{y}_i$, where m_d is the number of parameters of a given (deterministic) model f_d (e.g. the number of iterations and the maximal sphere radius for the iterative model).

In contrast, the *population-based personalization* consists in finding the optimal parameters $\theta \in \mathbb{R}^{m_r}$ that lead to comparable synthetic and real distributions P and Q , where $\mathbf{x}_i = f_r(\theta)$ and f_r is the (random) model. From the point of view of the energy to minimize during the personalization process, this can be formulated as minimizing the Kullback-Leibler divergence between the distributions P and Q , approximated from the known samples $\{\mathbf{x}_i\}_{i=1}^p$ and $\{\mathbf{y}_i\}_{i=1}^q$ using a Gaussian kernel of bandwidth σ . The personalization is stopped when one of the three following conditions are met: the energy to minimize is below a given threshold ϵ , or N_{max} iterations have been performed, or the average and covariance of the synthetic population no longer substantially evolve.

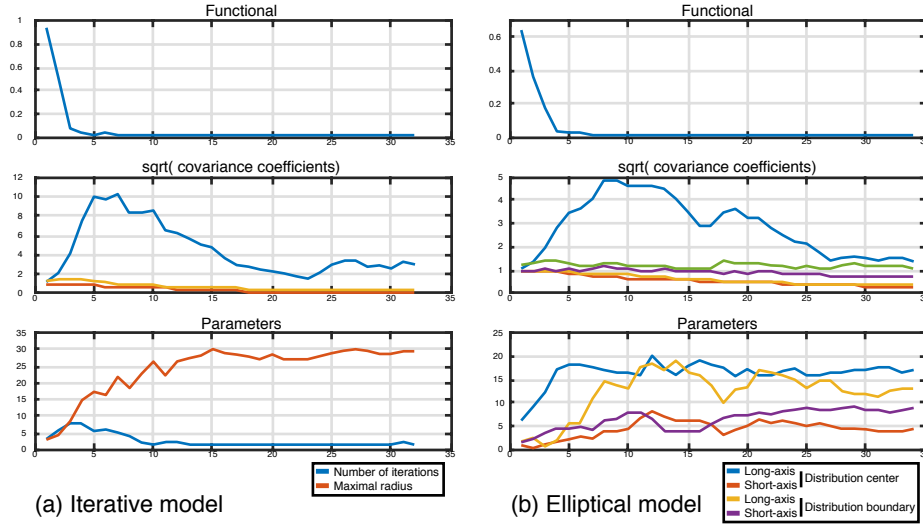


Fig. 2: Minimized energy (top), covariance matrix coefficients (middle), and model parameters (bottom) across iterations of the CMA-ES algorithm, for the iterative (a) and elliptical (b) models.

3 Experiments and results

3.1 CMA-ES parameters and convergence

For both models, the CMA-ES hyperparameters were empirically set to: 100 samples generated at each new generation of the algorithm, 15 samples retained to estimate the covariance matrix, maximum number of iterations $N_{max} = 500$, and the functional threshold $\epsilon = 0.004$ such that the covariance coefficients and the model parameters have enough iterations to stabilize. The kernel bandwidth σ used for the Kullback-Leibler divergence was set to the average distance between each sample and its 10 nearest neighbors.

The initial values were the identity for the covariance matrix, and 2 for each parameter of the model. However, as recommended by the CMA-ES authors, the long-axis of the elliptical model was rescaled by a factor 3 so that the sensitivity to the long- and short-axis parameters are (a-priori) comparable. This means a wider range of values for the infarct spread along the circumferential direction, although this does not explicitly constrain infarcts to spread more along this direction (to better correspond to what is observed on the real data).

Figure 2 reports the evolution of the minimized energy, the covariance matrix coefficients, and the model parameters across the generations of the CMA-ES algorithm. It shows that although the energy is rapidly minimized, the evolution of the covariance matrix coefficients needs to be monitored to ensure a stable solution: the highest coefficient exhibits a first increase necessary to get the energy minimum within the reach of the distribution P , and then decreases and stabilizes while the model parameters also converge, indicating that the personalization has been achieved and is stable. Of note, due to the randomness

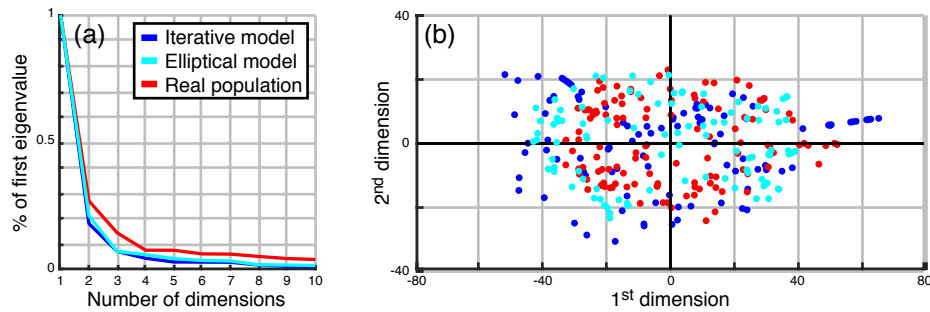


Fig. 3: Relative decrease of the eigenvalues (a) and distribution in the latent space (b) using non-linear dimensionality reduction with Isomap on the real (123 subjects) and synthetic populations generated using the iterative and elliptical models (100 samples for each model).

in the models, we do not expect a perfect stabilization of the model parameters and covariance coefficients fully converging to 0, unless an infinite set of synthetic samples is generated.

3.2 Distribution of the synthetic and real populations

We applied non-linear dimensionality reduction (Isomap) on the infarct patterns from the synthetic and real populations to better examine their distributions and the main variations they encode.

Figure 3 illustrates the relative decrease of the eigenvalues and the distribution of samples in the latent space. It shows that after personalization, the synthetic populations generated from the two models and the real population have similar distribution, although the synthetic populations are slightly more compact (faster decrease of eigenvalues), which is expected given that they originate from rather simple models.

Figure 4 complements these observations by representing the main variations encoded in these latent spaces, obtained by reconstructing the infarct patterns associated to $\{-2, -1, 0, 1, 2\}$ standard deviations along the first two dimensions. Reconstruction was achieved by multi-scale kernel regression [12], with equally weighted regularization and data terms. The regression used for the reconstruction of samples from the latent space may provide intermediate values between the healthy myocardium (blue) and the infarct (yellow): we kept them on purpose to assess the soundness of the reconstructed patterns, and only thresholded the values above/below these limits.

The synthetic and real populations exhibit comparable infarct shape, extent, and location along these modes of variation. Similar results are visible along the next dimensions (not displayed here to fit the page limit), although variations become more subtle as these distributions are rather compact. Slightly unplausible patterns may be observed at extreme values (e.g. +2 standard deviations for the elliptical model), due to the limits of the reconstruction method and the lower density of samples in this region of the latent space. The intrinsic limits

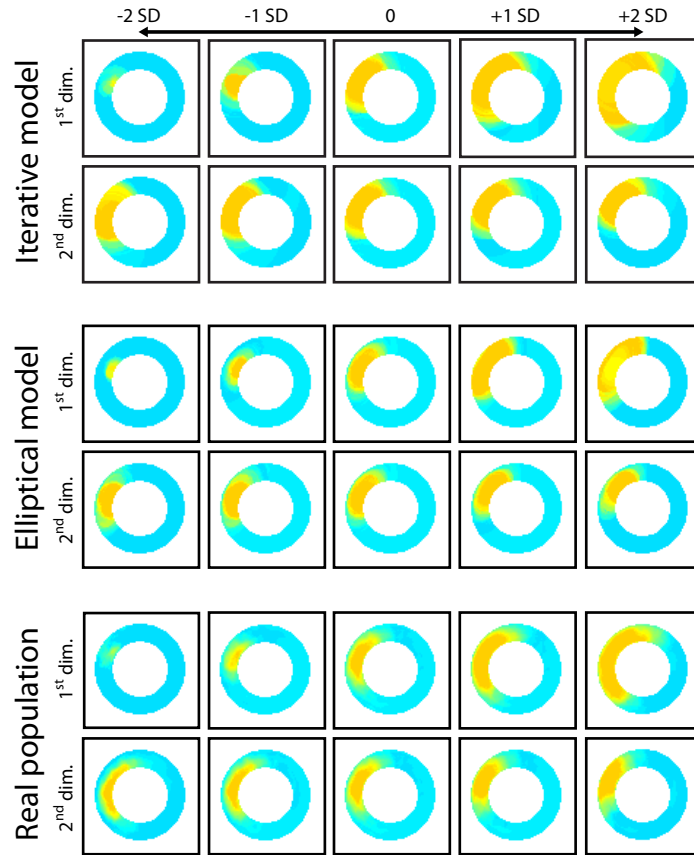


Fig. 4: First modes of variation (obtained through non-linear dimensionality reduction with Isomap on the infarct patterns) for the real and synthetic populations generated using the iterative and elliptical models. Intermediate values are observed between the healthy myocardium (blue) and the infarct (yellow) due to the regression used to reconstruct samples from the latent space.

of the models are also slightly visible: the iterative model tends to provide fully transmural lesions, while the elliptical model provides simpler infarct shapes.

4 Discussion

We proposed a method for the population-based personalization of simulations, illustrated in 2D on two simple geometrical models of myocardial infarction with respect to a real population of acute infarcts. We demonstrated the feasibility of this personalization strategy, which is therefore promising to generate virtual populations from models that intrinsically contain randomness.

The iterative model provides limited control on the infarct shape and its propagation from the endocardium. This intrinsic randomness, combined with

the stochastic search performed by the CMA-ES algorithm, are rather challenging to reach stable personalizations, although the examination of the optimization values (energy, covariance coefficients, and optimized parameters) confirms the convergence of the personalization. Besides, comparable optimal values of the model parameters were obtained by using a large covariance matrix as initialization. Adding relevant constraints on this model may improve realism and the robustness and efficiency of the personalization process.

Examining 2D data allowed focusing on the personalization process in a reasonable computing time (around 1 minute for the personalization of each model). The Kullback-Leibler divergence seems viable to compare the synthetic and real populations during personalization, although comparisons are directly performed on the infarct patterns and not on a relevant latent space as in variational auto-encoders [2]. We obtained comparable results with the Maximum Mean Discrepancy [13], although it tended to provide less stable results.

In 2D, the elliptical model seemed to better encode the infarct transmuralities but the generated shapes are rather simple. However, generalizing this work to 3D (planned for future work) will help to better state on the realism of such models to generate relevant synthetic populations. In any case, our primary focus here was to assess the feasibility of population-based personalization, and not on developing more complex infarct models. Nonetheless, further work should consider reducing the gap between synthetic and real data, using heterogeneous (instead of binary) infarct patterns, or at least a third region that could represent the border zone, and infarcts that may originate from the occlusion of several arteries. Provided relevant real data are available, our work is also generalizable beyond the acute infarct setting, for example to chronic infarcts, or to consider remodeling that may occur over time.

Acknowledgements. The authors acknowledge the support from the French ANR (LABEX PRIMES of Univ. Lyon [ANR-11-LABX-0063] within the program “Investissements d’Avenir” [ANR-11-IDEX-0007], the JCJC project “MIC-MAC” [ANR-19-CE45-0005]), and the Fédération Française de Cardiologie (“MI-MIX” project, Allocation René Foudon). They are also grateful to P Croisille and M Viallon (CREATIS, CHU Saint Etienne) for providing the imaging data for the MIMI population, and M Di Folco (CREATIS) for the preliminary exploration of the data alignment and simulation tools.

References

1. AJ Connolly and MJ Bishop. Computational representations of myocardial infarct scars and implications for arrhythmogenesis. *Clin Med Insights Cardiol*, 10:27–40, 2016.
2. DP Kingma and M Welling. Auto-encoding variational Bayes. *arXiv*, 2014.
3. RC Kerckhoffs, AD McCulloch, JH Omens, et al. Effects of biventricular pacing and scar size in a computational model of the failing heart with left bundle branch block. *Med Image Anal*, 13:362–9, 2009.
4. N Duchateau, M De Craene, P Allain, et al. Infarct localization from myocardial deformation: prediction and uncertainty quantification by regression from a low-dimensional space. *IEEE Trans Med Imaging*, 35:2340–52, 2016.

5. CN Leong, E Lim, A Andriyana, et al. The role of infarct transmural extent in infarct extension: A computational study. *Int J Numer Method Biomed Eng*, 33:e02794, 2017.
6. KA Reimer, JE Lowe, MM Rasmussen, et al. The wavefront phenomenon of ischemic cell death. 1. myocardial infarct size vs duration of coronary occlusion in dogs. *Circulation*, 56:786–94, 1977.
7. A Pashaei, C Hoogendoorn, R Sebastián, et al. Effect of scar development on fast electrophysiological models of the human heart: In-silico study on atlas-based virtual populations. *Proc. FIMH, LNCS*, 6666:427–36, 2011.
8. GK Rumindo, N Duchateau, P Croisille, et al. Strain-based parameters for infarct localization: evaluation via a learning algorithm on a synthetic database of pathological hearts. *FIMH, LNCS*, 10263:106–14, 2017.
9. L Belle, P Motreff, L Mangin, et al. Comparison of immediate with delayed stenting using the Minimalist Immediate Mechanical Intervention approach in acute ST-segment-elevation myocardial infarction: the MIMI Study. *Circ Cardiovasc Interv*, 9:e003388, 2016.
10. N Hansen and A Ostermeier. Adapting arbitrary normal mutation distributions in evolution strategies: The covariance matrix adaptation. *Proc. ICEC*, pages 312–7, 1996.
11. R Molléro, X Pennec, Delingette H, et al. Multifidelity-CMA: a multifidelity approach for efficient personalisation of 3D cardiac electromechanical models. *Biomech Model Mechanobiol*, 17:285–300, 2018.
12. N Duchateau, M De Craene, M Sitges, et al. Adaptation of multiscale function extension to inexact matching. *Proc. SEE GSI, LNCS*, 8085:578–86, 2013.
13. A Gretton, KM Borgwardt, MJ Rasch, et al. A kernel two-sample test. *J Mach Learn Res*, 13:723–73, 2012.