



HAL
open science

Automatic myocardial ischemic lesion detection on magnetic resonance perfusion weighted imaging prior perfusion quantification: A pre-modeling strategy

Clément Daviller, Thomas Grenier, Hélène Ratiney, Michaël Sdika, Pierre Croisille, Magalie Viallon

► To cite this version:

Clément Daviller, Thomas Grenier, Hélène Ratiney, Michaël Sdika, Pierre Croisille, et al.. Automatic myocardial ischemic lesion detection on magnetic resonance perfusion weighted imaging prior perfusion quantification: A pre-modeling strategy. *Computers in Biology and Medicine*, 2019, 110, pp.108-119. 10.1016/j.compbiomed.2019.05.001 . hal-02135018

HAL Id: hal-02135018

<https://hal.science/hal-02135018v1>

Submitted on 26 Oct 2020

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.

1 **Automatic Myocardial Ischemic Lesion Detection on**
2 **Magnetic Resonance Perfusion Weighted Imaging prior**
3 **Perfusion Quantification: a pre-modeling strategy.**

4 Clément Daviller¹, Thomas Grenier¹, Helene Ratiney¹, Michaël Sdika¹, Pierre Croisille^{2,3} and
5 Magalie Viallon^{2,3}

6 1. Univ. Lyon, INSA-Lyon, Université Claude Bernard Lyon 1, UJM-Saint Etienne, CNRS, Inserm,
7 CREATIS UMR 5520, U1206, F-69621, VILLEURBANNE, France

8 2. Univ Lyon, UJM-Saint-Etienne, INSA, CNRS UMR 5520, INSERM U1206, CREATIS, F-42023,
9 SAINT-ETIENNE, France

10 3. Radiology Dept. CHU de Saint Etienne, Univ Lyon, UJM-Saint-Etienne, F-42023, SAINT-ETIENNE,
11 France

12
13
14 **Corresponding author : Clément Daviller**

15 Mail : clement.daviller@creatis.insa-lyon.fr

16 Phone : +33 (0)4 72 44 80 15 fax: +33 (0)4 72 44 81 99

17 Address: 3, rue Victor Grignard 69616 Villeurbanne Cedex France

18

19 Abstract

20 Even if cardiovascular magnetic resonance (CMR) perfusion imaging has proven its relevance
21 for visual detection of ischemia, myocardial blood flow (MBF) quantification at the voxel
22 observation scale remains challenging. Integration of an automated segmentation step, prior
23 to perfusion index estimation, might be a significant reconstruction component that could allow
24 sustainable assumptions and constraint enlargement prior to advanced modeling. Current
25 clustering techniques, such as bullseye representation or manual delineation, are not
26 designed to discriminate voxels belonging to the lesion from healthy areas. Hence, the
27 resulting average time–intensity curve, which is assumed to represent the dynamic contrast
28 enhancement inside of a lesion, might be contaminated by voxels with perfectly healthy
29 microcirculation.

30 This study introduces a hierarchical lesion segmentation approach based on time–intensity
31 curve features that considers the spatial particularities of CMR myocardial perfusion. A first k-
32 means clustering approach enables this method to perform coarse clustering, which is refined
33 by a novel spatiotemporal region-growing (STRG) segmentation, thus ensuring spatial and
34 time–intensity curve homogeneity.

35 Over a cohort of 30 patients, myocardial blood flow (MBF) measured in voxels of lesion regions
36 detected with STRG was significantly lower than in regions drawn manually (mean difference
37 = 0.14, 95% CI [0.07, 0.2]) and defined with the bullseye template (mean difference = 0.25,
38 95% CI [0.17, 0.36]). Over the 90 analyzed slices, the median Dice scores calculated against
39 the ground truth ranged between 0.62 and 0.67, the inclusion coefficients ranged between
40 0.62 and 0.76 and the centroid distances ranged between 0.97–3.88 mm. Therefore, though
41 these metrics highlight spatial differences, they could not be used as an index to evaluate the
42 accuracy and performance of the method, which can only be attested by the variability of the
43 MBF clinical index.

44

45

46

47

48 Introduction

49 MR contrast-enhanced myocardial perfusion imaging (ceMPI) can supply insight into
50 microcirculation in myocardial tissue[1]. In coronary artery disease (CAD), this method offers
51 crucial information on the impact of vessel lesions with flow reduction. ceMPI also allows
52 myocardial blood flow assessment, topics that are of major interest in CAD lesion detection
53 and clinical decisions. Numerous studies have further demonstrated the ability to quantify
54 myocardial perfusion from MR perfusion-weighted (PW) image series [2,3].

55 However, although quantitative perfusion indexes such as myocardial blood flow (MBF) have
56 proven their usefulness for diagnosis [4,5], they remain rare in clinical use. Among the
57 obstacles that must be overcome are (i) the complexity of the protocol from PW image
58 acquisition to quantification, (ii) the wide variability between results produced by the
59 approaches, (iii) the large amount of data that must be processed, and (iv) the numerous
60 assumptions that must be met to reduce the number of solutions because deconvolution is an
61 ill-posed inverse problem.

62 Quantitative measurement approaches are usually divided into two categories: model
63 independent and parametric approaches. Model-independent approaches assume the
64 myocardial circulation system is linear and shift invariant (LSI) [2]. These approaches are often
65 preferred but are noise sensitive because the introduction of small input errors might result in
66 large measurement bias [6]. Large measurement biases particularly occur in regions with
67 perfusion defects where the contrast-to-noise ratio (CNR) is poor due to low signal
68 enhancement of the time-intensity curves. In contrast, model-dependent approaches rely on
69 parametric determination of the residue function $R(t)$ and are constructed with various levels
70 of complexity [7]. Despite their theoretical accuracy, numerous assumptions must be satisfied
71 prior to measurement that are often not sustainable and cannot be made at the voxel level in
72 a reasonable amount of time.

73 For these reasons, observation at a higher scale than the voxel is often preferred. Among the
74 alternative methods are manual segmentation of lesions or the use of the American Heart
75 Association (AHA) bullseye template, in which the myocardium regions are assigned to 17
76 predetermined anatomical segments [8]. Manual segmentation requires that the boundary
77 lesion delineation be left to the clinician, which is time-consuming and subjective due to
78 operator dependence. If the AHA bullseye template method is rapid and sufficient in the supply
79 of a semiquantitative coarse overview of the lesion extent in clinical summary reports, it is
80 actually not an acceptable template for quantification. Indeed, the AHA theoretical anatomical
81 boundaries rarely match the individual lesion shapes. This approximation can lead to a nearly

82 systematic over- or underestimation bias of the perfusion index, measurements depending on
83 the perfusion level of the contaminating tissue.

84 Recently, promising segmentation methods [9,10] have been proposed that locally cluster the
85 voxels based on quantitative perfusion index similarity. However, even if these techniques
86 supply segmentation that considers the shape of the suffering territory, the segmentation is
87 conducted after the quantification.

88 In this work, we propose a new segmentation method based on spatiotemporal region growing
89 which clusters the neighboring myocardium voxels with similar tissue characteristics. This
90 approach is based on temporal signal behavior that is wisely conditioned by lesion-specific
91 features. Therefore, a first contribution is the extraction of the classically used lesion-specific
92 features in a more robust way.

93 The performance of our unsupervised boundary lesion delineation was evaluated by
94 i) comparing the MBF values calculated in the obtained classified regions against these
95 extracted from expert segmentation, over a cohort of 30 subjects referred for known or
96 suspected CAD for a perfusion CMR with pharmacological stress in clinical settings and
97 ii) comparing the ROIs obtained by our unsupervised technique and ROIs defined by expert
98 segmentation by calculating the Dice score inclusion coefficient and the centroid distance.

99

100 **Materials and methods**

101 **Study population and MR Acquisition**

102 We recruited thirty patients referred for known or suspected CAD with perfusion CMR. The
103 study was performed with the approval of the Institutional Review Board (IRBN 052019/
104 CHUSTE), and written informed consent was obtained from all subjects. We excluded patients
105 if they had metallic implants or implanted cardiac devices incompatible with CMR, a glomerular
106 filtration rate ≤ 30 ml/min, a high degree of atrioventricular (AV) blocks, severe chronic
107 obstructive pulmonary disease, or claustrophobia. Patients were asked to abstain from
108 caffeine-containing products for ≥ 12 hours prior to CMR examination.

109 Perfusion imaging was performed using a dual-acquisition approach as described by
110 Gatehouse et al. [11], on a 3 T MAGNETOM PRISMA scanner (Siemens Healthineers,
111 Erlangen, Germany) with an 18-element surface coil. A standard protocol was used with
112 cardiac localization, steady-state free-precession cine images acquired to cover the heart from
113 base to apex. Contrast-enhanced inversion-recovery images were acquired 10 min after

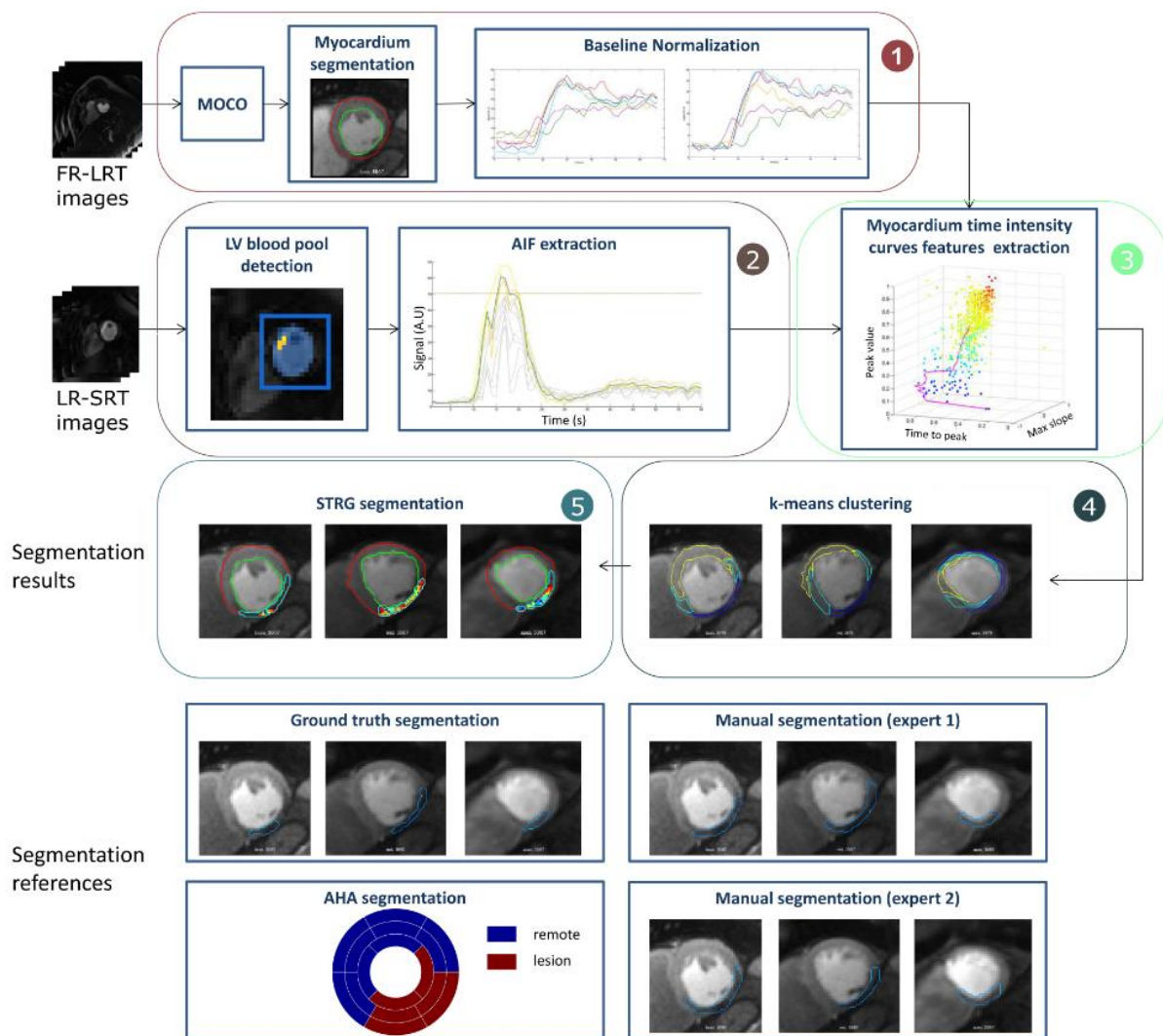
114 injection. Vasodilation was obtained with a 0.4 ml bolus injection of Regadenoson (Rapiscan,
115 GEMS) 1 minute before imaging. Perfusion acquisition was subsequently performed at 3 to 5
116 short-axis locations at every heart beat with a bolus injection (6 ml/s) of gadoterate meglumine
117 (0.2 mmol/kg) (Dotarem, France). We used the first 10 interbeat RR intervals for precontrast
118 baseline signal measurement, and the data were acquired in free-breathing mode 60 to 110
119 heartbeats (~60-70 s). Two types of images were acquired at every heart beat: i) a saturation
120 recovery (SR)-prepared sequence with low resolution and a short-recovery-time (LR-SRT)
121 image to avoid saturation of the LV-blood pool signal, and ii) SR-prepared with fine resolution
122 and long-recovery-time (FR-LRT) images for 3 to 5 slices that could fit into the RR interval.
123 The LR-SRT image series was used in arterial input function AIF estimation, and the FR-LRT
124 series were used in perfusion analysis. The sequence acquired the first 3 proton density (PD)-
125 weighted scans planned for signal spatial normalization (flip angle 5°) before imaging of the
126 T1-weighted frames. T1 weighting was obtained with a nonselective SR pulse train followed
127 by a saturation recovery time, defined as the duration between the end of the saturation pulse
128 and the beginning of k-space acquisition. The acquisition kernel was a 2D single-shot turbo-
129 Flash sequence. The FR-LRT main acquisition parameters were assigned as follows: spatial
130 resolution = $1.98 \times 1.98 \text{ mm}^2$, long saturation-recovery time = 43 ms, flip angle $\alpha = 10^\circ$,
131 TR = 2ms, echo time TE = 0.95 ms, parallel acquisition (TPAT) mode using generalized
132 autocalibrating partially parallel acquisitions (GRAPPA) with acceleration factor = 3, and linear
133 k-space reordering. The LR-SRT images main MR parameters were:
134 spatial resolution = $5.94 \times 5.94 \text{ mm}^2$, flip angle $\alpha = 8^\circ$, repetition time TR = 1.3 ms, echo time
135 TE = 0.74 ms, short saturation recovery time = 5 ms, centric k-space reordering. For both
136 acquisitions, slice thickness = 8 mm, FOV = $380 \times 380 \text{ mm}^2$.

137

138 **Data analysis**

139 This section details the proposed data processing, which is summarized in figure 1. In the
140 following, we denote an image series as the time-intensity curves $c_i(t)$ of each voxel v_i .

141



143

144 **Figure 1:** Global overview of the proposed pipeline for the myocardium segmentation process.
 145 Inputs of the pipeline are the low-resolution short-recovery-time (LR-SRT) and fine resolution
 146 long-recovery-time (FR-LRT) image series. The steps described in the Materials and methods
 147 are numbered here by the bullet points from 1 to 5. Both series types are processed
 148 independently (step 1 and 2) to extract myocardium time intensity curve features (step 3).
 149 These features are exploited to process the segmentation approach composed of a first
 150 coarse k-means clustering (step 4), allowing myocardium lesion location. It is followed by
 151 spatio temporal region-growing (STRG) segmentation refinement giving a precise contour of
 152 the lesion (step 5). The results are finally compared to reference segmentations performed by
 153 two experts. The ground-truth segmentation is the result of their consensus of what they
 154 considered as definite ischemic lesion.

155

156 **Step 1: Preprocessing of MOCO, normalization and manual myocardial segmentation**

157 First, all MR images (LR-SRT and FR-LRT) were normalized and motion corrected by the non-
158 rigid Siemens MOCO algorithm [12], as reported in Figure 1. For each slice level, LV
159 myocardium segmentation was manually performed by drawing the endo- and epicardial
160 contours on the temporal-maximal-intensity-projection (tMIP) image [10,13]. This image was
161 enhanced with an emphasis filter to sharpen the myocardium borders. This image clearly
162 delineated the endo- and epicardial contours, especially by highlighting the left and right
163 ventricle cavities, as well as the LV lateral portion of the myocardial wall. We applied the same
164 myocardium segmentation template for all analyses.

165 **Step 2: Estimating AIF to improve feature extraction**

166 Most of the approaches that address myocardial lesion segmentation [14] or analysis are
167 based on the features extracted from FR-LRT: peak value, time to peak, maximum upslope,
168 area under the curve, contrast agent bolus arrival delay, etc. These features might not be
169 robustly extracted from the image series because artifacts, noise, movement correction errors
170 or normalization bias can occur in a clinical acquisition context. Therefore, we attempted to
171 improve the feature extraction using assumptions based on AIF knowledge.

172 We first estimated the AIF from the LR-SRT images to avoid signal distortions. On these
173 images, the blood pool was detected first. From this detection, the curve with the maximum
174 peak intensity value $C_{\max}(t)$ served as a reference to aggregate other voxels, with $c_i(t)$ reaching
175 80% of this maximum value(see appendix A1). Finally, all of the curves were averaged to
176 compute the mean AIF time-intensity curve $S_a(t)$

177 Subsequently, we extracted the following features from $S_a(t)$:

- 178 • peak value P_a , as the maximum value reached by $S_a(t)$
- 179 • peak date t_{Pa} , as the time required for S_a to reach P_a
- 180 • foot t_{Fa} , as the time required for S_a to reach 5% of P_a before reaching P_a
- 181 • AIF first pass end date t_{Ea} , as the minimum signal intensity value reached by S_a after
182 t_{Pa} and before the recirculation peak.

183 **Step 3: Robust tissue features extraction for improved segmentation**

184 As described in step 2, robust segmentation relies on features that are wisely extracted from
185 the myocardium tissue time-intensity curves $S_m(t)$. Indeed, the robustness and accuracy of the
186 segmentation method described in this article rely on these features. A high contrast-to-noise
187 ratio (CNR) of $S_a(t)$ makes AIF features extraction easy. However, myocardium time-intensity

188 curves $S_m(t)$ with lower CNR and more exposure to artifacts require more complex processing
 189 to yield accurate feature extraction. In the following section, we propose a new lesion specific
 190 feature extraction method that is more robust to artifacts. We combine the prior knowledge of
 191 the $S_a(t)$ features with two assumptions. First, the studied system is causal. As a consequence,
 192 myocardium signal enhancement cannot start before t_{Fa} . Second, we consider that the $S_m(t)$
 193 peak value cannot occur after the end of the AIF first pass. These latter values are set as
 194 hierarchical constraints for search of $S_m(t)$ features of voxels present in the myocardium region
 195 (manually delineated) extracted from the FR-LRT images in manner similar to AIF estimation:

- 196 • The peak value P_m in the tissue is calculated as the maximum value of $S_m(t)$ with t
 197 ranging from the AIF foot and the end of the AIF first pass (equation 1)

$$P_m = \max_{t \in [t_{Pa}; t_{Ea}]} (S_m(t)) \quad (1)$$

198

- 199 • The peak date t_{Pm} is the time required for $S_m(t)$ to reach P_m
- 200 • The area under the curve AUC is calculated as the sum of the $S_m(t)$ values
 201 (equation 2)

$$AUC = \sum_{t=0}^{t_{Pm}} |S_m(t)| \quad (2)$$

- 202 • The $S_m(t)$ foot t_{Fm} is calculated in the same way as t_{Fa}
- 203 • The bolus arrival delay Δt_{Fm} is calculated as the time difference between t_{Fa} and t_{Fm}
 204 (equation 3)

$$\Delta t_{Fm} = t_{Fm} - t_{Fa} \quad (3)$$

- 205 • The maximum upslope is calculated as the maximum intensity difference in a time
 206 interval Δt between two consecutive acquisitions in the time range from t_{Fm} to t_{Pm}
 207 (equation 4)

$$\delta_{max} = \max_{t \in [t_{Fm}; t_{Pm}]} (S_m(t + \Delta t) - S_m(t)) \quad (4)$$

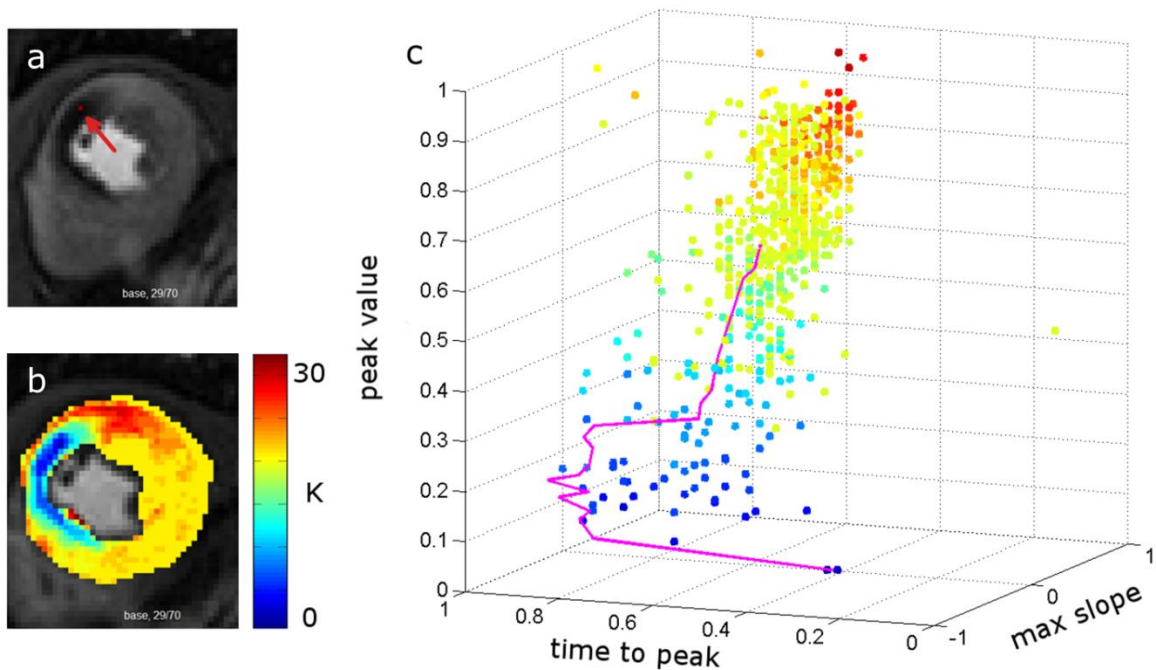
- 208 • The time to peak Δt_{max} is the time difference between t_{Fm} and t_{Pm} (equation 5)

$$\Delta t_{max} = t_{Pm} - t_{Fm} \quad (5)$$

209

210 This hierarchical approach avoids feature computation errors due to noise and variability of
 211 the time-intensity curve shapes and supplies robust material for the following lesion detection
 212 step. In the following segmentation approach, the 4 features P_m , δ_{max} , AUC and Δt_{max} are used.

213 Figure 2 presents a 3-dimensional space representation of the features extracted from the
 214 time-intensity curves in the PW images acquired on a patient suffering from a severe stenosis.
 215 The three dimensions are the usual perfusion semi quantitative parameters: maximum slope
 216 δ_{\max} , peak value P_m and time to peak Δt_{\max} . This plot gives a simple visualization of the
 217 differences between the time-intensity curve features from the affected regions and from
 218 remote tissue.



219

220 **Figure 2:** Extraction of myocardium voxels' time-intensity curve features (peak value, time to
 221 peak and maximum slope). The extraction starts from the voxel seed (red voxel in a.) with the
 222 lowest area under the curve and considered the center of the lesion. The seed is iteratively
 223 expanded to its neighbors, according to equation 9. At each iteration, the tolerance threshold
 224 K is incremented until covering the entire myocardium. The map on (b) represents the
 225 minimum value of K required to include a voxel in the region. (c) Time-intensity curve features
 226 of each voxel are displayed in the 3D feature space. Color map of the points is identical to (b).
 227 The magenta curve in (c) represents the region-growing scheme of the average time-intensity
 228 curve feature evolution of the growing region as a function of the iteration.

229 **Step 4: Segmentation based on K-means clustering**

230 From the features P_m , δ_{max} , AUC and Δt_{max} extracted from the voxels in the myocardial region
231 and for each slice, a first coarse clustering pass was applied using the k-means [15] algorithm.
232 The choice of k-means was motivated by the fact that only one parameter is needed by this
233 simple and well used method for segmentation problems. The principle of this unsupervised
234 approach is briefly reviewed.

235 The aim is to cluster the myocardium voxels v_i into a defined number of clusters k . Each voxel
236 v_i is associated with a vector of feature values φ_i . k-means clustering consists of minimizing
237 the vector distance of voxels within a cluster by iteratively moving μ_j (equation 6):

$$\min \sum_{j=1}^k \sum_{\varphi_i \in C_j} \|\varphi_i - \mu_j\|^2 \quad (6)$$

238

239 where j is the cluster index, i is the voxel vector index number, φ_i and μ_j are respectively the
240 feature vectors $[P_m, \delta_{max}, AUC, \Delta t_{max}]$ of voxel i and cluster centroid j , and C_j is the set of voxels
241 associated with cluster j (a voxel φ_i is associated with its cluster C_j according to the Euclidean
242 distance).

243 This minimization is performed by k-means based on the algorithm proposed by Arthur and
244 Vassilvitskii [16], which improves the running time and the overall quality of the final solution.

245 Based on physiological considerations, the number of clusters k was set to 3, assuming that
246 a gradient exists in the voxel features from the healthy tissue areas to the affected tissue
247 areas. The third cluster gathers voxels with maximal time-intensity curve features, and the
248 second gathers voxels with gradual time-intensity curve enhancement. The first cluster is
249 composed of voxels with the poorest enhancement features.

250 Voxels v_i belonging to the first cluster C_1 (i.e., the voxels cluster with the lower average time-
251 intensity curve sum) are considered abnormal perfusion areas and are reported in a 2-
252 dimensional mask R . This mask R is finally set as an input of a connected component analysis
253 to identify possible multiple lesion regions labeled R_n in the myocardium, where all 8-connectivity
254 connected voxels v_i reported in R are labeled with an identical value. This rule can be
255 described by equation 7:

$$R = C_1 = \bigcup_n R_n \quad (7)$$

256

257 **Step 5: Myocardial lesion segmentation refinement using spatiotemporal region**
 258 **growing (STRG)**

259 This step details our major methodological contribution: the STRG algorithm is applied on
 260 each labeled region R_n outputted by the k-means clustering described in step 4. The principle
 261 of this approach is to expand a seed defined as the voxel with the lowest time-intensity curve
 262 sum, as defined by equation 8:

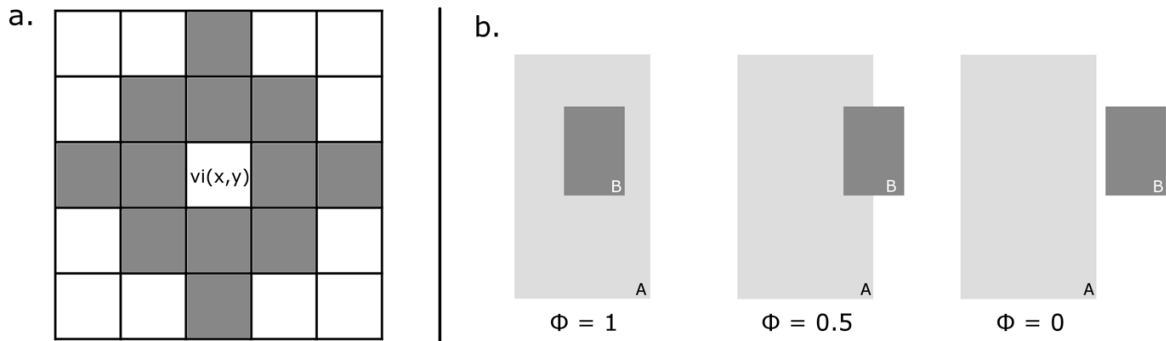
$$v_{seed,n} = \arg \min_{v \in R_n} \sum_{t=0}^T C_v(t) \quad (8)$$

263 where n is the index of the labeled region R_n , and T is the last acquisition of the time series.

264 The region R_n is expanded on its 12 connected neighbors (Figure 3-a) belonging to R_n , and
 265 matching the condition set by equation 9:

$$\|C_i(t) - C_{seed,R_n}(t)\|_{\infty} \leq K, \text{ with } i \text{ such as } v_i \in N_{12}(R_n) \quad (9)$$

266 We note that each v_i is the time-intensity curve $C_i(t)$ and that $\|\cdot\|_{\infty}$ represents the infinity norm,
 267 meaning that we retain the maximum difference between $C_i(t)$ and $C_{seed,R_n}(t)$. In other words,
 268 at each acquisition the intensity difference between the current analyzed voxel v_i and the seed
 269 must be lower than a tolerance threshold K .



270

271 **Figure 3:** (a) 12-connected neighboring pattern and (b) Principle of inclusion coefficient of set
 272 B in A with 3 examples. (Left) B is totally included in A ($\Phi = 1$), (center) exactly half of set B is
 273 out of A ($\Phi = 0.5$), and (right) B is completely out of A ($\Phi = 0$).

274

275 STRG is repeated by incrementing the value of K, starting from K = 1. At each iteration, the
 276 average time-intensity curve $S_{avg}^K(t)$ of the region R_i^K found by STRG and its feature vector
 277 $F_{avg}^K = [P_m, \delta_{max}, AUC, \Delta t_{max}]$ are calculated. This algorithm is iterated until it reaches the
 278 feature consistency condition as defined in the following:

Lesion segmentation refinement:

Input: R_i input (coarse) segmentation from k-means

λ : consistency coefficient

Outputs: K^* : the optimal tolerance threshold,

M_λ : the refined segmentation of R_i

beginning

$K \leftarrow 0$

do

$K \leftarrow K + 1$

$R_i^K \leftarrow \text{STRG}(R_i, K)$

$S_{avg}^K(t) \leftarrow \text{mean of } C_i(t) \text{ of } v_i \in R_i^K$

$F_{avg}^K \leftarrow \text{Compute Features of } S_{avg}^K(t)$

while $\left(\frac{\|F_{avg}^K\| - \|F_{seed}\|}{\|F_{seed}\|} \leq \lambda \right)$ **AND** $(R_i^K < R_n)$

$M_\lambda \leftarrow R_i^K$

$K^* \leftarrow K$

End

279

280 This finer segmentation M_λ supplies spatial consistency and guarantees time-intensity curve
 281 similarity in the region outputted by STRG. Indeed, the neighboring voxels included in the
 282 lesion mask are associated with a signal for which the degree of characteristic variation is
 283 controlled by the λ criterion.

284 This algorithm was applied by varying the feature consistency coefficient λ from 0.1 to 0.7 with
 285 an incremental value of 0.1. The results were compared with the medical expert segmentation
 286 ground-truth masks G considered as a reference by calculating the Dice scores D [17] to
 287 determine the optimal value for λ^* . λ^* was defined as the lower λ value that gave the higher
 288 Dice score (equation 10) average over the basal, mid-cavity and apical slices.

$$\lambda^* = \arg \max_{\lambda} (Dice(M_\lambda, G)) \quad (10)$$

289 where M_λ is the segmentation mask processed by automated segmentation. The Dice score
 290 is an efficient index that compares the similarity of two segmentations: the proposed
 291 segmentation mask M and the ground truth mask G . If the masks perfectly overlap, the Dice
 292 score is equal to 1. If masks share no pixels in common, the score is 0. The Dice score is
 293 computed as follow (equation 11):

$$Dice(M_\lambda, G) = \frac{2 \cdot (M_\lambda \cap G)}{M_\lambda + G} \quad (11)$$

294 where $|\cdot|$ means the cardinality of the set and \cap is the intersection of the two masks. An
 295 inclusion coefficient was also calculated from the results computed with STRG and configured
 296 with $\lambda = \lambda^*$. This criterion is based on medical expert segmentation used as a reference and
 297 reflects the automated segmentation mask inclusion in the reference. The inclusion coefficient
 298 is established by equation 12:

$$\Phi = \frac{M_{\lambda^*} - M_{\lambda^*} \setminus G}{M_{\lambda^*}} \quad (12)$$

299

300 This coefficient calculates the ratio between the surface of M_{λ^*} overlapping with G over the
 301 surface of M_{λ^*} . Hence, if M_{λ^*} is completely included in G , then the resulting inclusion criterion
 302 is $\Phi = 1$. The score decreases as the surface of M_{λ^*} deviates from mask G until it reaches
 303 $\Phi = 0$, when both masks have no common voxels. The principle of this index is represented
 304 in Figure 3-b.

305 We finally assessed the lesion detection accuracy from the centroid distance. This metric
 306 evaluates the distance between the lesion centroids detected by STRG and the ground truth
 307 segmentations. The centroid C of a lesion region was computed from equation 13:

$$C = \left(\frac{M_{10}}{M_{00}}, \frac{M_{01}}{M_{00}} \right) \quad (13)$$

308

309 where M_{00} , M_{10} , and M_{01} , represent respectively the area, the horizontal moment, and the
 310 vertical moment, calculated with equation 14:

$$M_{ij} = \sum_{x=1}^W \sum_{y=1}^H x^i y^j R(x, y) \quad (14)$$

311

312 The distance D between two centroids C_1 and C_2 was then calculated as the Euclidian distance
 313 (equation 15).

$$D = \sqrt{((x_{c2} - x_{c1}) - (y_{c2} - y_{c1}))} \quad (15)$$

314

315

316 ***Manual segmentation***

317 Ischemic lesion segmentation was conducted blindly, randomly and independently on a
318 dedicated postprocessing workstation by two level-3 CMR-experienced observers who were
319 unaware of the clinical and angiographic data. The observers were asked to delineate what
320 they would consider the definite ischemic lesion and the remote normal region. They were
321 only allowed to use the original DICOM images for analysis, and lesion segmentation was
322 performed on the frame that was considered the most representative.

323 ***Ground-truth segmentation***

324 Both experts were asked to reach a consensus on the definite ischemic lesion with all available
325 clinical and angiographic data made available to them. They were also allowed to use an
326 interactive spatiotemporal clustering mapping tool to refine their decision based on all
327 spatiotemporal available data. The algorithm of this semi-automated segmentation tool is quite
328 similar to STRG because it allows the user to select a voxel as a seed that is iteratively
329 expanded on its neighbors by matching the condition set by equation (9). In our case, the
330 value of the tolerance threshold was manually set by the user. This final ground-truth
331 segmentation (also known as G) mask set was defined as the reference for subsequent
332 comparisons.

333

334 **Statistical analysis**

335 Data were screened for normality using the Shapiro-Wilk test, and were accordingly reported
336 as the mean \pm standard deviation (SD) or median and 95% confidence interval [95% CI]. Using
337 two-way mixed ANOVA modeling, the MBF values were compared among segmentation
338 methods (within-subject factor), with assessment of the main effect and potential interaction
339 of the slice level or type of disease (CAD or microvascular disease) (between-subject factors).
340 Box's conservation correction factor was applied to account for sphericity violation. Posthoc
341 pairwise comparisons were performed when applicable using Bonferroni adjustments.

342 Statistical analysis was performed using Medcalc 18.2.1 (Ostend, Belgium) and Stata SE 15.1
343 (Statacorp, College Station, TX, USA). For all analyses, statistical significance was accepted
344 at $p < 0.05$.

345

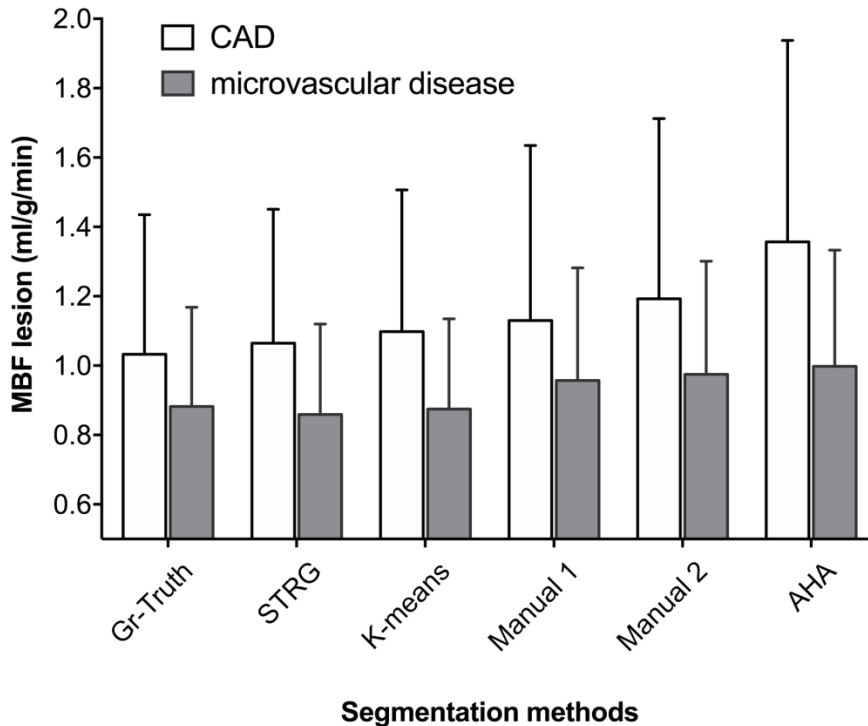
346 **Results**

347 Among the 30 included patients, 22 demonstrated focal hypoperfused lesions that were
348 related to macrovascular CAD disease (1-vessel = 9, 2-vessels = 6, 3-vessels = 7), and 8
349 patients displayed a microvascular diffuse disease (dilated cardiomyopathy = 2, diabetes = 3,
350 hypertrophic cardiomyopathy = 3).

351 ***Comparison of MBF values obtained across segmentation approaches***

352 The MBF values obtained after lesion segmentation with the various approaches are shown
353 in Figure 4. Overall, although no main effect of slice location ($p = 0.27$) or disease type
354 ($p = 0.19$) was observed, a significant interaction was noted between disease type and method
355 ($F(5; 410) = 3.55$, $p = 0.004$). A statistically significant difference in mean MBF values was
356 found across methods ($F(5; 415) = 31.24$, $p < 0.001$). Pairwise comparisons showed that
357 although only a minimal and nonsignificant difference was found between the ground-truth
358 reference and STRG (mean difference = 0.02, 95% CI[-0.07, 0.04], $p > 0.9$) or k-means (mean
359 difference = 0.046, 95% CI [-0.1,0.01], $p = 0.24$), significantly higher mean MBF was observed
360 for both manual segmentation (mean difference = 0.14, 95% CI [0.07, 0.2]) and AHA (mean
361 difference = 0.25, 95% CI [0.17, 0.36]).

362



363

364 **Figure 4:** Average MBF measurements calculated at voxel observation scales in lesions
 365 regions detected by ground-truth segmentation defined by a consensus between two medical
 366 experts (Gr-Truth), STRG and prior coarse segmentation (STRG and K-means, respectively),
 367 manual segmentations performed by the two experts (Manual 1 and Manual 2), and AHA
 368 bullseye segmentation (AHA).

369

370 In the lesion regions detected by STRG, we calculated the mean and standard deviation of
 371 the blood flows of each MBF measured directly at the voxel scale, $\overline{MBF_{vox}}$ and $SD(MBF_{vox})$,
 372 respectively. We compared these with the blood-flow MBF_{ROI} measured on the ROI average-
 373 time curve. The median values of the differences observed in the base, middle and apex slices
 374 were 20%, 8% and 16% respectively. The standard deviation over the time-intensity curve
 375 baseline showed reductions of 58%, 48% and 43% when calculated on the ROI average time
 376 curve, hence improving significantly the average time-intensity curve CNR.

377 Overall, although no main effect of slice location ($p = 0.27$) or disease type ($p = 0.19$) was
 378 observed, a significant interaction was noted between disease-type and method
 379 ($F(5; 410) = 3.55, p = 0.004$). A statistically significant difference in mean MBF values was
 380 found across methods ($F(5; 415) = 31.24, p < 0.001$). Pairwise comparisons showed that
 381 although only a minimal and nonsignificant difference between the ground-truth reference and
 382 STRG (mean difference = 0.02, 95% CI[-0.07, 0.04], $p > 0.9$) or k-means (mean difference =

0.046, 95% CI [-0.1,0.01], $p = 0.24$) occurred, significantly higher mean MBF was observed for both manual segmentation (mean difference = 0.14, 95% CI [0.07, 0.2]) and AHA (mean difference = 0.25, 95% CI [0.17, 0.36]). Such results suggest that STRG is more efficient over the 3 slice levels (base, middle and apex) than the other tested approaches.

When considering the MBF difference between the lesion and the remote region (delta-MBF) in CAD subjects, significant overall differences were found across methods ($F(4; 4236) = 17.43, p < 0.001$). No difference in delta-MBF was observed between ground-truth and STRG approaches (mean difference = -0.03, 95% CI[-0.103, 0.040], $p > 0.9$). The delta-MBF was significantly underestimated by manual (m1 $p = 0.004$, m2 $p = 0.02$) and AHA approaches ($p < 0.001$).

Comparison of classifiers using Dice score inclusion coefficient and centroid distance metrics

The median Dice scores calculated for pairs of ground-truth and automatically computed masks M_λ for each value of λ and by slice are listed in Table 1. The first column shows the Dice score reached with the simple k-means preclustering approach masks R. Based on the Dice scores, overall higher patient mask matches were obtained for a coherence coefficient value of $\lambda = 0.5$. Median scores were respectively equal to $D_{base} = 0.62$, $D_{middle} = 0.63$ and $D_{apex} = 0.66$ for base, middle and apical slices. These scores were higher than those obtained with k-means presegmentation ($D_{base} = 0.62$, $D_{middle} = 0.59$ and $D_{apex} = 0.56$, respectively) showing that similarity with the ground-truth mask was better after STRG refinement.

	k-means	$\lambda = 0.1$	$\lambda = 0.2$	$\lambda = 0.3$	$\lambda = 0.4$	$\lambda = 0.5$	$\lambda = 0.6$	$\lambda = 0.7$
Base	0.62	0.60	0.58	0.62	0.64	0.62	0.65	0.64
Middle	0.59	0.60	0.61	0.63	0.64	0.63	0.63	0.63
Apex	0.56	0.60	0.60	0.60	0.61	0.66	0.63	0.63
Median	0.59	0.60	0.60	0.62	0.63	0.64	0.64	0.63

403

Table 1: Median Dice scores among all patients for the three slice levels calculated between expert segmentation masks and k-means clustering (first column), or STRG segmentation masks with λ coefficient values ranging from 0.1 to 0.7. The best Dice score (green column) was obtained for $\lambda = 0.5$.

Tables 2 and 3 report the Dice scores obtained on the same population of patients but divided into two categories: those suffering from focal myocardial lesions (22 patients) and those

410 suffering from diffuse ischemic diseases (8 patients). The best Dice score for focal lesions was
 411 obtained with a slightly higher λ value than for the diffuse lesions. With respect to the relatively
 412 low number of patients included in the study and especially when subdivided into subsets,
 413 median calculation was chosen instead of average.

414

	k- means	$\lambda = 0.1$	$\lambda = 0.2$	$\lambda = 0.3$	$\lambda = 0.4$	$\lambda = 0.5$	$\lambda = 0.6$	$\lambda = 0.7$
Base	0.62	0.60	0.58	0.66	0.62	0.59	0.64	0.63
Mid	0.59	0.58	0.61	0.63	0.64	0.63	0.67	0.63
Apex	0.59	0.65	0.63	0.66	0.66	0.67	0.67	0.67
Median	0.60	0.61	0.61	0.65	0.64	0.63	0.66	0.64

415

416 **Table 2:** Median Dice score calculated identically to Table 1, by selecting only the patient sub-
 417 population with a focal 1- or 2-vessel disease. The best Dice score (green column) was
 418 obtained for $\lambda = 0.6$.

419

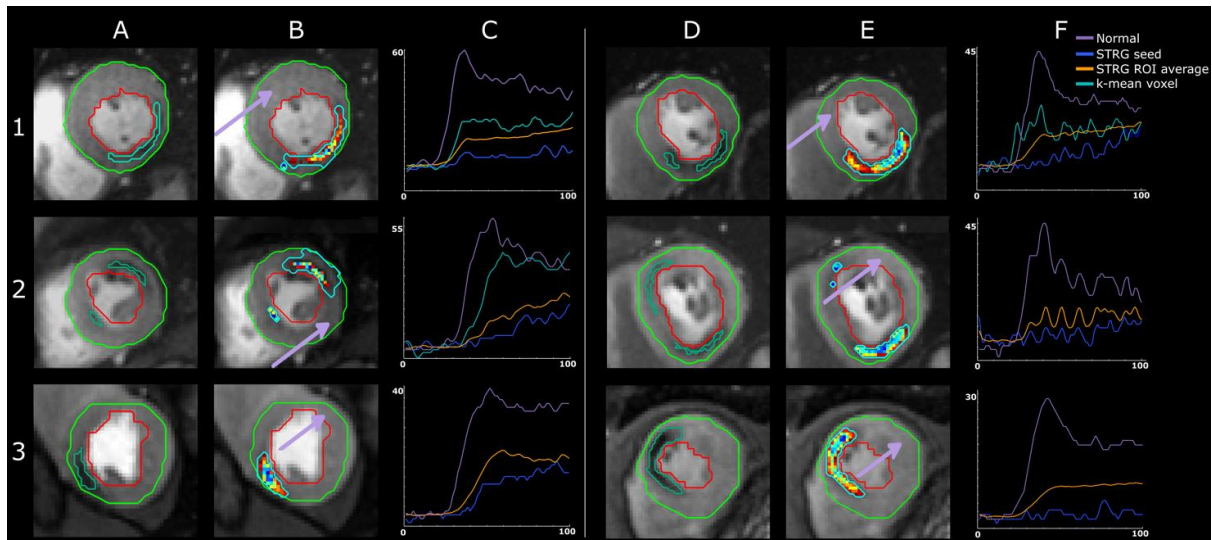
	k- means	$\lambda = 0.1$	$\lambda = 0.2$	$\lambda = 0.3$	$\lambda = 0.4$	$\lambda = 0.5$	$\lambda = 0.6$	$\lambda = 0.7$
Base	0.61	0.61	0.56	0.67	0.67	0.67	0.67	0.67
Mid	0.60	0.65	0.60	0.63	0.63	0.62	0.58	0.62
Apex	0.56	0.60	0.60	0.60	0.61	0.66	0.63	0.63
Median	0.59	0.62	0.59	0.63	0.64	0.65	0.63	0.64

420

421 **Table 3:** Median Dice score calculated identically to Table 1, by selecting only the patient sub-
 422 population with diffuse micro-vascular lesions. The best Dice score (green column) was
 423 obtained for $\lambda = 0.5$ which was the same as for the complete patient population.

424 Figure 5 presents the segmentation results from image series acquired on 6 patients suffering
 425 from various ischemic lesions. The plots display the average time-intensity curves of the
 426 largest regions identified by the final segmentation. In 3 cases (B1, B2 and F1) out of 6, the
 427 masks found by the STRG algorithm were more restrictive than the masks found by k-means
 428 presegmentation. In all cases and as expected, the average time-intensity curve of the region
 429 outputted by STRG $C_{STRG}(t)$ presented a greater peak value enhancement and was smoother
 430 than that of $C_{seed}(t)$. When the optimal STRG region M_{λ^*} was more restrictive than the k-means

431 region R, the behavior of the voxels present in R and outside of M_{λ^*} showed important
 432 differences (higher peak value, maximum slope and AUC). This result shows the importance
 433 of the refinement supplied by the STRG approach to reach an average signal that is as close
 434 as possible to the lesion center without contamination by voxels showing intermediate or low
 435 perfusion defects. Oscillations in the time-intensity curves displayed in plot F1 and especially
 436 F2 reveal motion correction mismatches. Despite these artifacts, segmentation was accurately
 437 detected in both cases.



438
 439 **Figure 5:** Segmentation results on mid-ventricular slices acquired in 6 different subjects. Endo
 440 and epicardium borders are defined by the red and green contours, respectively. Images on
 441 columns A and D expose the manual lesion segmentation. k-means and STRG segmentations
 442 are displayed in columns B and E, respectively, with light-blue boundaries and with colored
 443 masks. The color map of the STRG mask indicates the threshold value K for which the voxel
 444 has been included in the region. The plots in columns C and F show time-intensity curves of:
 445 (purple) normal voxels; (light blue) voxels belonging to k-means segmentation and out of the
 446 STRG mask (if any); (orange) the STRG mask average; and (dark blue) the STRG seed voxel.
 447 Purple arrows indicate the spatial position of the voxel from which was extracted normal signal.
 448

449 Table 4 indicates the median and interquartile range (IQR) of inclusion coefficients Φ of STRG
 450 segmentation masks M_{λ^*} in the ground-truth masks. These values were calculated over all
 451 patient cohorts and over the dense and diffuse patient subsets on basal, middle and apical
 452 slices. In the complete patient set, the median (IQR) inclusion coefficients Φ were equal to
 453 $\Phi_{\text{base}} = 0.68$ (0.30), $\Phi_{\text{middle}} = 0.62$ (0.27) and $\Phi_{\text{apex}} = 0.75$ (0.30). In comparison, the median
 454 (IQR) inclusion coefficients of the k-means segmentation masks R in the ground truth mask
 455 were $\Phi_{\text{base}} = 0.56$ (0.28), $\Phi_{\text{middle}} = 0.56$ (0.22) and $\Phi_{\text{apex}} = 0.55$ (0.34). This result indicates that
 456 in all cases, inclusion was significantly greater with STRG segmentation than with k-means
 457 presegmentation. Table 4 also indicates that the inclusion coefficients are greater on diffuse
 458 lesions than on focal ones.

	Base		Middle		Apex	
	Median	IQR	Median	IQR	Median	IQR
All patients ($\lambda = 0.5$)	0.68	0.30	0.62	0.27	0.75	0.30
focal lesions ($\lambda = 0.6$)	0.63	0.30	0.62	0.28	0.74	0.34
Diffuse lesions ($\lambda = 0.5$)	0.73	0.23	0.69	0.21	0.76	0.19

459 **Table 4:** Median and interquartile range inclusion coefficients of automated STRG
 460 segmentation against expert segmentation, calculated over the entire patient population, focal
 461 lesion subset, and diffuse lesion subset. The coefficient was calculated over basal, middle and
 462 apical slices.

463

		<i>Focal 1- or 2-vessel disease</i>		<i>Diffuse micro-vascular disease</i>	
		<i>k-means</i>	<i>STRG</i> ($\lambda = 0.6$)	<i>k-means</i>	<i>STRG</i> ($\lambda = 0.5$)
Dice Score	Base	0.62(0.20)	0.64(0.19)	0.61(0.20)	0.67(0.16)
	Middle	0.59(0.16)	0.67(0.14)	0.60(0.26)	0.62(0.17)
	Apex	0.59(0.21)	0.67(0.21)	0.56(0.18)	0.66(0.17)
Inclusion Coefficient	Base	0.53(0.21)	0.63(0.30)	0.67(0.48)	0.73(0.23)
	Middle	0.56(0.25)	0.62(0.28)	0.57(0.18)	0.69(0.21)
	Apex	0.58(0.37)	0.74(0.34)	0.55(0.15)	0.76(0.19)
Centroid Distance (mm)	Base	3.22(2.80)	2.54(1.38)	2.31(3.45)	1.64 (1.44)
	Middle	3.15(2.24)	2.33(2.81)	1.84(1.12)	1.73(0.68)
	Apex	1.67(1.42)	0.97(0.99)	3.88(5.51)	2.93(1.52)

464 **Table 5** : Median (IQR) Dice scores, inclusion coefficients and centroid distances (in mm),
465 calculated over the two patient sub populations for basal, middle and apical slices.

466 Table 5 summarizes the medians and interquartile ranges of Dice scores, inclusion coefficients
467 and centroid distances between ground truth and regions detected by k-means and STRG for
468 the two categories of populations. In all cases, the scores were improved by STRG refinement
469 compared to the respective k-means. While the Dice score gave information on the lesion
470 coverage, the centroid distance informed on the location accuracy of the lesion core.
471 Reduction of this distance with the use of STRG, suggests a closer lesion core location to
472 ground truth as defined by experts.

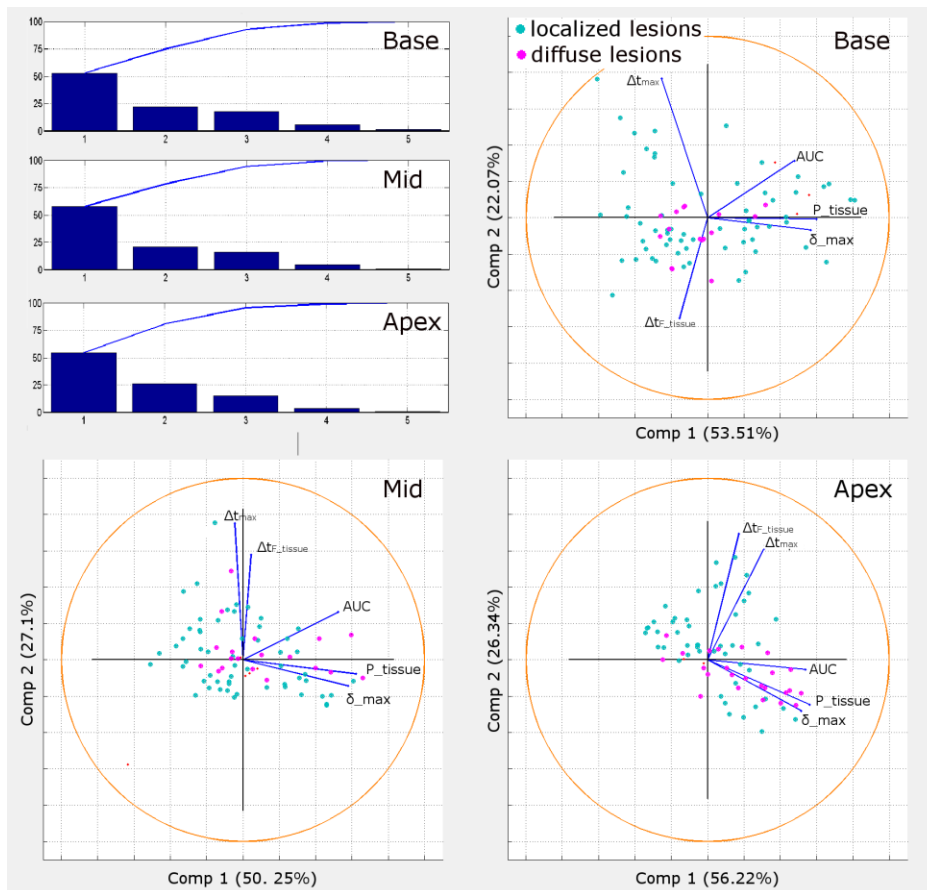
473 We calculated the median and IQR of the K^* values used in segmentation over the all-patient
474 set and by disease type as defined above (CAD or diffuse diseases) with their respective λ^*
475 values. These values are reported in Table 6 by slice (base, middle and apex). Minor
476 differences were observed between the optimal thresholds used in the context of either CAD
477 of diffuse disease lesion. Nevertheless, in all cases the coefficient IQR calculations indicated
478 a high variability because they ranged from 10.5 to 12. This observation highlights an
479 important amplitude of the used thresholds over all datasets that was independent of the type
480 of lesion and shows that the use of an adaptive threshold was necessary for robust and
481 accurate segmentation.

	Base		Mid		Apex	
	Median	IQR	Median	IQR	Median	IQR
All patients ($\lambda = 0.5$)	8.50	11.5	7	10.5	10.5	12
Focal lesions ($\lambda = 0.6$)	8	13	7	13	12	11
Diffuse lesions ($\lambda = 0.5$)	10.5	9	11	10	6	14

482 **Table 6:** Median and interquartile range applied threshold K^* for STRG segmentation over
483 (line 1) the complete patient population with $\lambda = 0.5$, (line 2) the patient population with a
484 focal ischemic lesion and with $\lambda = 0.6$, and (line 3) the patient population with diffuse
485 ischemic lesion $\lambda = 0.5$

486 **Principle component analysis**

487 Principle component analysis (PCA) was applied by using the time-intensity curve features
488 extracted from the lesion regions obtained by automated segmentations. PCA was performed
489 to determine selected salient features of the perfusion curves within the different types of
490 segmented lesions. The Pareto diagram in Figure 6 indicates that more than 90% of the
491 information stored in the time-intensity curve features was carried by the first 3 components.
492 For all slices, the first component was essentially based on the peak value and maximum
493 slope, which were highly correlated. The second component was mainly defined by the delay
494 and time-to-peak features. The orthogonality of these features to the maximum slope and peak
495 value suggests independence between the two groups of features. Delay and time to peak
496 were highly correlated in the middle-and apex plot and negatively correlated in the base plot.



497

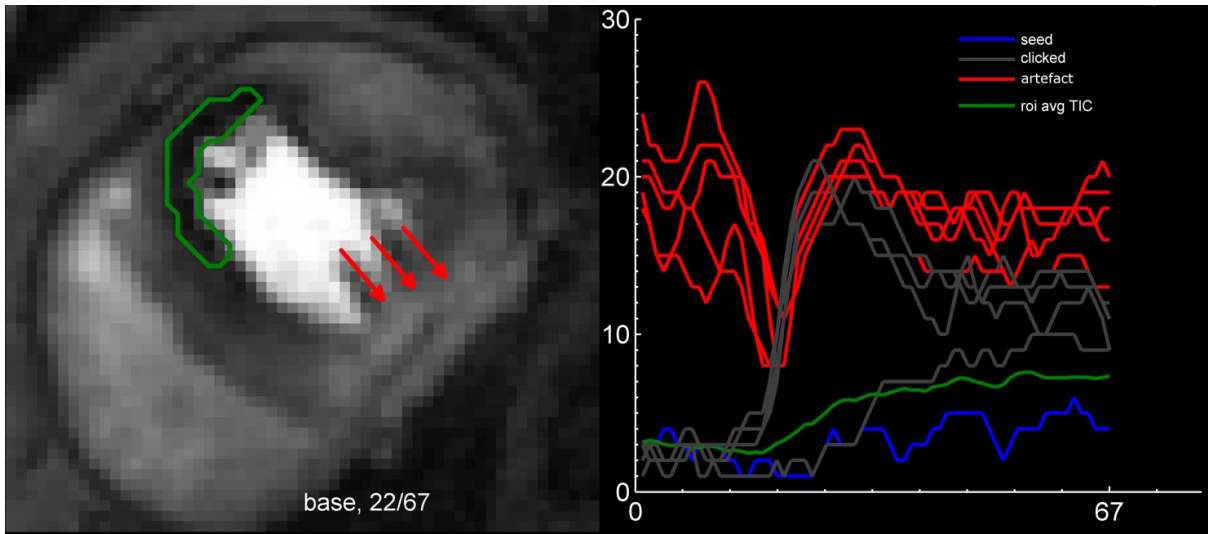
498 **Figure 6:** (top-left) Pareto diagram and principle component analysis plots of average time-
 499 intensity curves features in regions automatically segmented by the proposed approach. For
 500 ease of visualization, PCA diagrams were plotted over 2 dimensions that already contained
 501 more than 75% of the information. Blue points represent focal lesions and magenta
 502 diffuse ones.

503 The diffuse lesion region signal features appeared to be significantly influenced by the time to
 504 peak value and maximum slope feature, especially when examining the apex PCA plot. Focal
 505 lesion features were spread over the diagram plot with no significant influence from any
 506 feature.

507 No significant difference between diffuse and focal lesions could be observed. However, it is
 508 clear that the diffuse lesion population is more concentrated around the peak value and
 509 maximum slope feature vector. This outcome was actually expected because enhancement
 510 is more effective in this type of pathology than in focal lesions where the peak value was
 511 extremely weak. Hence, the variability could not be determined.

512 Figure 7 indicates a lesion detected on a patient slice with a focal lesion and aliasing artifact.
 513 The presence of this artifact had no impact on the final segmentation due to the use of 3

514 clusters at the k-means segmentation step. Indeed, a cluster was “sacrificed” for the artifact,
515 but the lesion cluster was sufficiently robust to allow accurate STRG segmentation.



516

517 **Figure 7:** (left) STRG segmentation result (green) on an image series with aliasing artifact
518 occurring on an inferior segment shown by red arrows; (right) Plots of time-intensity curves of
519 voxels and regions associated with artifact (red), lesion region detected by STRG (green),
520 lesion region's seed (blue), voxels out of lesion region (gray).

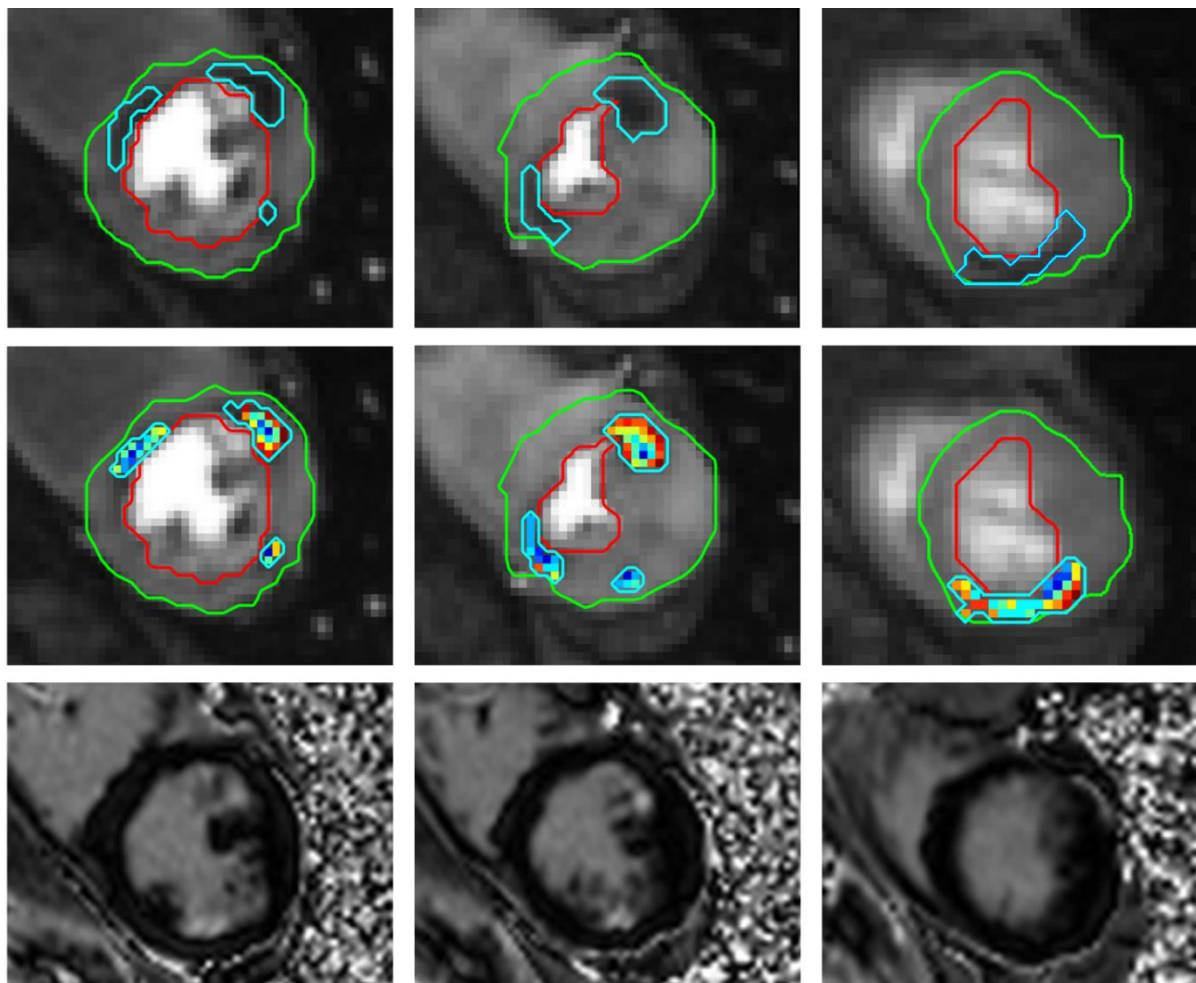
521

522 **Discussion**

523 This proposed segmentation approach offers a robust, easy, reliable and objective automatic
524 method to detect most representative pixels from ischemic coronary lesions using PW images
525 based on the time-intensity curve shapes and features of voxels belonging to the myocardium.
526 Average MBF values measured in targeted lesion regions are indeed lower and closer to the
527 ones obtained by the expert ground-truth classification and have lower standard deviation in
528 the determined region.

529 Our hierarchical approach proposes a first rough segmentation based on time-intensity curve
530 feature similarity and refinement by a method that sets spatial and temporal constraints
531 regulated by the time-intensity curve features to exploit the totality of the information given by
532 the PW images, especially in the temporal dimension. Exploitation of the spatial constraint
533 only after the second step (STRG) enable multiple lesion detection when it occurs (Figure
534 5[B2-E2], and Figure 8) since STRG is performed as many times as the number of connected
535 regions detected by k-means clustering.

536



537

538 **Figure 8:** (top line) Ground-truth segmentation and (middle line) segmentation results on PW-
539 image series at peak enhancement acquired sequentially without delay after the RR wave on
540 a patient with multiple lesions on anterior, septal, and inferolateral segments. Green and red
541 lines indicate the endo and epicardial bounds. k-means (cyan curves) and STRG (color
542 masks) segmentation results totally matched on septal and inferolateral segments. STRG
543 segmentation was slightly more restrictive than k-means on anterior segmentation. (Bottom
544 line) late gadolinium enhancement images acquired on same patient and slices at end-diastole
545 mismatched with perfusion images, indicating pure ischemic lesion without necrosis.

546 This method exquisitely considers the high interpatient variability of the signal collected in the
547 images to propose an adaptive method for accurate segmentation. STRG segmentation relies
548 on an automatically set and appropriate tolerance threshold K , based on the average time-
549 intensity curve coherence. This feature gives the method a wide level of flexibility while making
550 it robust to temporal frame coregistration artifacts (Figure 5 – plots F1 and F2) and to aliasing
551 artifacts, as shown in Figure 7.

552 We observed that measurements of $\overline{MBF_{\text{vox}}}$, calculated as an average of MBF values from
553 a map region, showed differences from measurements of MBF_{ROI} , as well as difference from
554 blood flow measurement of the average time curve in the same region. This result shows that
555 in a certain number of cases, there is value in clustering before quantification, even with a
556 simple quantification approach, demonstrating the utility of the segmentation approach.

557 As shown above by the PCA, the focal ischemic lesions' signal was less homogenous than in
558 the case of diffuse regions, leading to the need for a slightly higher λ value. However, this
559 higher tolerance is compatible with accurate segmentation considering ground-truth
560 comparison because the feature difference between the lesion and healthy tissue is important.
561 In this study, the algorithm was calibrated to obtain the best match between automated and
562 manual segmentations based on the Dice score. Nevertheless, the use of a more restrictive
563 λ -value could guarantee a higher similarity between time-intensity curves and is of interest for
564 analysis of the lesion center because a significant enhancement difference might occur
565 between the time-intensity curves of the seed and final lesion region, as shown in Figure 5,
566 plot F3.

567 By showing that lesion feature variability was important, especially for focal lesions, PCA also
568 confirmed that an accurate and robust segmentation cannot only consider one feature, but a
569 combination of features, done in this study.

570 Nevertheless, this technique presents one limitation. This technique relies on the exploitation
571 of relative differences between time-intensity curve features regardless of any analysis of the

572 meaning of these quantities. As a consequence, this method cannot be considered solely for
573 triage between healthy and CAD patients.

574 In all contexts, the Dice score reached a modest value although the results met visual
575 agreement. We note that the Dice score, which is a common and frequently used comparison
576 method, is intrinsically not the best approach for evaluating such results. Indeed, if physicians
577 are trained to detect the culprit sectors with high sensitivity, they never precisely size the
578 ischemic lesions and only the voxels with minimal doubt are conserved for the ground-truth
579 definition. Moreover, although the Dice score is highly dependent on the lesion size, the spatial
580 surface of the smallest lesions can be less than 10 voxels. In this particular case, small
581 variations between detection and ground-truth might lead to important reduction of the score.
582 Therefore, our proposed method could perhaps be considered as more invariant and reliable
583 than subjective manual segmentation.

584 The Dice score was complemented by inclusion coefficient Φ , and centroid distance. The first
585 showed the segmentation accuracy of STRG with a median ranging between $\Phi = 0.62$ and
586 $\Phi = 0.76$. Hence, a major portion of the STRG segmentation mask M_{λ^*} was similar to the
587 ground-truth mask, but a portion of this latter group was somehow rejected by STRG because
588 the time-intensity curve behavior of this portion of the surface did not match the seed signal
589 under the conditions laid down by the consistency coefficient λ^* . This result suggests that the
590 STRG segmentation was more restrictive than that performed by the clinician in charge of
591 reference segmentation. This restrictive segmentation is nevertheless crucial because the
592 behavior of the average region time-intensity curve must reflect the microcirculation of the
593 segmented region particularly, in the case of the perfusion indexes measurement context in
594 which the quantification approaches are often ill-conditioned problems and rely on
595 assumptions required to isolate a solution. The centroid distance showed the improvement of
596 lesion center location accuracy when STRG was applied compared to k-means. In other
597 words, STRG also modifies the k-means presegmentation to “refocus” the lesion core because
598 STRG can very accurately detect the location of the seed that is the voxel showing the lowest
599 indicator uptake. The lesion region is then built around this seed by aggregating adjacent
600 voxels matching the initial curve behavior. It is therefore not surprising that STRG improves
601 centroid distance score.

602 Even if these metrics allowed us to highlight spatial differences, only the MBF variability should
603 be considered to evaluate the approach accuracy. The reason is mainly the difficulty of
604 defining a robust and precise delineation of the lesion because the temporal dimension
605 acquired image series makes it difficult to make an incontestable lesion region definition.
606 Indeed, only one frame of the series – often the peak contrast frame - is used, due to the

607 representation limitation, when the perfusion defect information is spread over the entire
608 series.

609 We also emphasize that this segmentation technique used AIF calculated from LR-SRT
610 images to improve the robustness of the myocardium TIC features in the FR-LRT images. The
611 LR-SRT images are usually exploited, as in this work, for perfusion quantification purposes
612 because they allow avoidance of the signal distortions in the LV blood pool due to important
613 contrast agent concentration [11]. However, the CNR from the LV blood pool signal of the FR-
614 LRT images could be sufficient to perform accurate myocardium $S_t(t)$ feature extraction.
615 Hence, we anticipate no difficulty in adapting the pipeline approach with PW images acquired
616 using other techniques, as long as they supply a sufficient CNR.

617 This segmentation technique proved its ability to yield robust and accurate lesion
618 segmentation results with the use of only semiquantitative perfusion parameters. This
619 approach also guarantees time-intensity curve homogeneity by constraining the variation of
620 features by a single coefficient. Moreover, this approach could be combined with prior
621 segmentation via a myocardium segmentation technique. Further work should be performed
622 to tackle this issue by using convolutional neural networks.

623 The use of complex perfusion models for perfusion indexes quantification is often limited
624 because it requires assumptions that are difficult to make, due to a lack of knowledge of the
625 physiological mechanisms in the lesion, and in particular, at the voxel observation scale level.
626 The average time-intensity curve regions where the homogeneity of microcirculation behavior
627 is guaranteed could offer possibilities for assumptions by analysis of its shape where
628 implementation would be far too long if this analysis would have to be done voxel by voxel.
629 Associated with LSI deconvolution methods, this approach can supply regions with attenuated
630 breathing and heart movement artifacts, and improve CNR in lesion locations where it is
631 particularly low. Hence, the accuracy and robustness of measurements could be improved.

632 From the perspective of improvements, this study did not exploit the spatial inter-slice lesion
633 detection coherence to enhance the algorithm's robustness. However, this information, which
634 is widely used by clinicians, could be a good indicator to eliminate casual false positive
635 detection. Moreover, according to recent advances in segmentation based on machine
636 learning approaches, the manually performed myocardium segmentation on FR-LRT images
637 should soon be performed automatically with acceptable robustness [18–21], even in
638 ischemic patients.

639 In conclusion, the approach proposed in this work aimed to aggregate the voxels most likely
640 belonging to ischemic lesions. This classification was compared with the ground truth selection

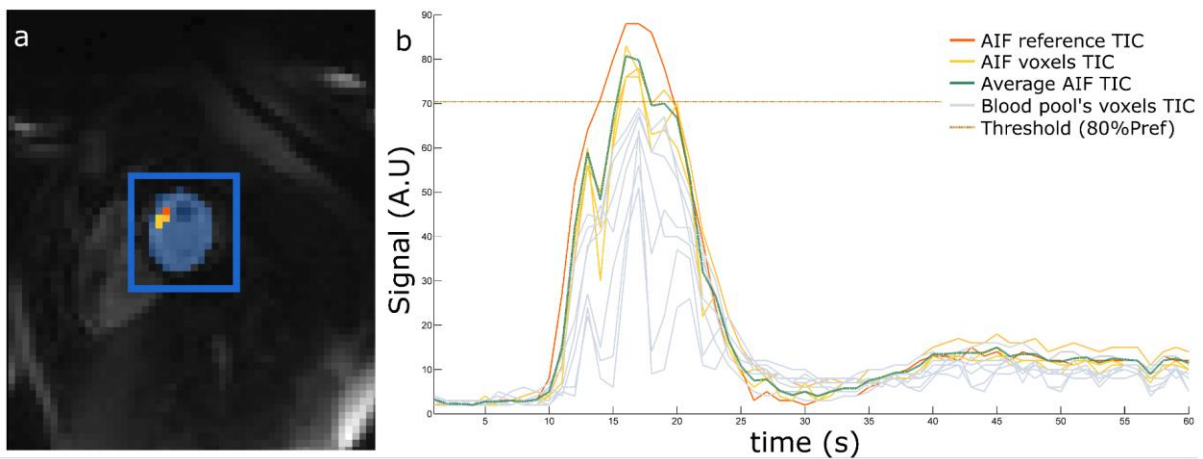
641 performed by two experts. Our study demonstrated that this approach automatically achieved
642 the closest ground truth clustering of voxels belonging to the lesion, provided that at least one
643 lesion exists, which is the only limitation of the developed solution. To our knowledge, there
644 are no large clinical studies that permit us to clearly understand how MBF values are spread
645 between different patient groups clustered by their pathologies. We could legitimately consider
646 that a large-scale descriptive statistic could be used to characterize normal values and thus
647 introduce additional characteristics such as normal value limits, perhaps concluding that MBF
648 values are normal in the absence of lesions.

649 The output of our algorithm can be directly applied for inline calculation of the MBF in the
650 determined culprit region of interest to supply i) automatic and robust detection of the culprit
651 sector(s) that intrinsically excludes susceptibility artifacts and ii) the most representative
652 stress MBF values in the suspected regions. Indeed, stress MBF are as the most clinically
653 relevant parameter for clinical decision-making prior to revascularization [22].

654

655 **Appendix**

656 Figure A-1: Arterial input function extraction process. The AIF is extracted from (a) the LR-
657 SRT image series by segmentation (light blue surface) of the LV blood pool cavity (b) Time-
658 intensity curves of voxels owning to this region are then analyzed to identify the one (light
659 yellow) reaching the highest peak value. This value is set as the peak reference P_{ref} , and
660 (yellow) voxels v_a whose peak values reaching more than 80% of P_{ref} are considered as
661 relevant for the AIF estimation. The (green) AIF time-intensity curve $S_a(t)$ is then calculated
662 as the average time-intensity curves $c_a(t)$ of voxels v_a .



663

664

665 **Acknowledgments:**

666 This work was supported by the LABEX PRIMES (ANR-11-LABX-0063) of Université de
667 Lyon, within the program "Investissements d'Avenir" (ANR-11-IDEX-0007) operated by the
668 French National Research Agency (ANR).

669 This work was supported by the RHU MARVELOUS (ANR-16-RHUS-0009) of l'Université
670 Claude Bernard Lyon 1 (UCBL), within the program "Investissements d'Avenir" operated by
671 the French National Research Agency (ANR).

672 This work was performed on a platform / by a laboratory member of France Life Imaging
673 network (grant ANR-11-INBS-0006)

674

675

- 677 [1] K.L. Gould, N.P. Johnson, T.M. Bateman, R.S. Beanlands, F.M. Bengel, R. Bober, P.G.
678 Camici, M.D. Cerqueira, B.J.W. Chow, M.F. Di Carli, S. Dorbala, H. Gewirtz, R.J.
679 Gropler, P.A. Kaufmann, P. Knaapen, J. Knuuti, M.E. Merhige, K.P. Rentrop, T.D.
680 Ruddy, H.R. Schelbert, T.H. Schindler, M. Schwaiger, S. Sdringola, J. Vitarello, K.A.
681 Williams, D. Gordon, V. Dilsizian, J. Narula, Anatomic Versus Physiologic Assessment of
682 Coronary Artery Disease, *J. Am. Coll. Cardiol.* 62 (2013) 1639–1653.
683 doi:10.1016/j.jacc.2013.07.076.
- 684 [2] D.C. Lee, N.P. Johnson, Quantification of Absolute Myocardial Blood Flow by Magnetic
685 Resonance Perfusion Imaging, *JACC-Cardiovasc. Imag.* 2 (2009) 761–770.
686 doi:10.1016/j.jcmg.2009.04.003.
- 687 [3] M. Jerosch-Herold, Quantification of myocardial perfusion by cardiovascular magnetic
688 resonance, *J. Cardiovasc. Magn. Reson.* 12 (2010) 57. doi:10.1186/1532-429X-12-57.
- 689 [4] M. Li, T. Zhou, L. Yang, Z. Peng, J. Ding, G. Sun, Diagnostic accuracy of myocardial
690 magnetic resonance perfusion to diagnose ischemic stenosis with fractional flow reserve
691 as reference: systematic review and meta-analysis, *JACC Cardiovasc Imaging.* 7 (2014)
692 1098–1105. doi:10.1016/j.jcmg.2014.07.011.
- 693 [5] J.P. Greenwood, N. Maredia, J.F. Younger, J.M. Brown, J. Nixon, C.C. Everett, P.
694 Bijsterveld, J.P. Ridgway, A. Radjenovic, C.J. Dickinson, S.G. Ball, S. Plein,
695 Cardiovascular magnetic resonance and single-photon emission computed tomography
696 for diagnosis of coronary heart disease (CE-MARC): a prospective trial, *Lancet.* 379
697 (2012) 453–460. doi:10.1016/S0140-6736(11)61335-4.
- 698 [6] K. Kudo, S. Christensen, M. Sasaki, L. Østergaard, H. Shirato, K. Ogasawara, M.
699 Wintermark, S. Warach, Stroke Imaging Repository (STIR) Investigators, Accuracy and
700 reliability assessment of CT and MR perfusion analysis software using a digital phantom,
701 *Radiology.* 267 (2013) 201–211. doi:10.1148/radiol.12112618.
- 702 [7] S.P. Sourbron, D.L. Buckley, Tracer kinetic modelling in MRI: estimating perfusion and
703 capillary permeability, *Phys. Med. Biol.* 57 (2012) R1–R33. doi:10.1088/0031-
704 9155/57/2/R1.
- 705 [8] M.D. Cerqueira, N.J. Weissman, V. Dilsizian, A.K. Jacobs, S. Kaul, W.K. Laskey, D.J.
706 Pennell, J.A. Rumberger, T. Ryan, M.S. Verani, American Heart Association Writing
707 Group on Myocardial Segmentation and Registration for Cardiac Imaging, Standardized
708 myocardial segmentation and nomenclature for tomographic imaging of the heart. A
709 statement for healthcare professionals from the Cardiac Imaging Committee of the
710 Council on Clinical Cardiology of the American Heart Association, *Int J Cardiovasc*
711 *Imaging.* 18 (2002) 539–542.
- 712 [9] T.F. Ismail, L.-Y. Hsu, A.M. Greve, C. Goncalves, A. Jabbour, A. Gulati, B. Hewins, N.
713 Mistry, R. Wage, M. Roughton, P.F. Ferreira, P. Gatehouse, D. Firmin, R. O’Hanlon, D.J.
714 Pennell, S.K. Prasad, A.E. Arai, Coronary microvascular ischemia in hypertrophic
715 cardiomyopathy - a pixel-wise quantitative cardiovascular magnetic resonance perfusion
716 study, *J. Cardiovasc. Magn. Reson.* 16 (2014) 49. doi:10.1186/s12968-014-0049-1.
- 717 [10] T. Chitiboi, A. Hennemuth, L. Tautz, P. Stolzmann, O.F. Donati, L. Linsen, H.K. Hahn,
718 Automatic Detection of Myocardial Perfusion Defects using Object-based Myocardium
719 Segmentation, in: 2013 Computing in Cardiology Conference (Cinc), Ieee, New York,
720 2013: pp. 639–642.
- 721 [11] P.D. Gatehouse, A.G. Elkington, N.A. Ablitt, G.-Z. Yang, D.J. Pennell, D.N. Firmin,
722 Accurate assessment of the arterial input function during high-dose myocardial perfusion
723 cardiovascular magnetic resonance, *J Magn Reson Imaging.* 20 (2004) 39–45.
724 doi:10.1002/jmri.20054.
- 725 [12] H. Xue, J. Guehring, L. Srinivasan, S. Zuehlsdorff, K. Saddi, C. Chefdhotel, J.V. Hajnal,
726 D. Rueckert, Evaluation of Rigid and Non-rigid Motion Compensation of Cardiac
727 Perfusion MRI, in: D. Metaxas, L. Axel, G. Fichtinger, G. Szekely (Eds.), *Medical Image*

728 Computing and Computer-Assisted Intervention - Miccai 2008, Pt li, Proceedings,
729 Springer-Verlag Berlin, Berlin, 2008: pp. 35–43.

730 [13]R.W.M. Underberg, F.J. Lagerwaard, B.J. Slotman, J.P. Cuijpers, S. Senan, Use of
731 maximum intensity projections (MIP) for target volume generation in 4DCT scans for
732 lung cancer, *Int. J. Radiat. Oncol. Biol. Phys.* 63 (2005) 253–260.
733 doi:10.1016/j.ijrobp.2005.05.045.

734 [14]M.S. Hansen, H. Olafsdottir, K. Sjostrand, S.G. Erbou, M.B. Stegmann, H.B.W. Larsson,
735 R. Larsen, Ischemic segment detection using the support vector domain description, in:
736 J.P.W. Pluim, J.M. Reinhardt (Eds.), *Medical Imaging 2007: Image Processing, Pts 1-3*,
737 Spie-Int Soc Optical Engineering, Bellingham, 2007: p. UNSP 65120F.

738 [15]S. Lloyd, Least-Squares Quantization in Pcm, *IEEE Trans. Inf. Theory.* 28 (1982) 129–
739 137. doi:10.1109/TIT.1982.1056489.

740 [16]D. Arthur, S. Vassilvitskii, *k-means plus plus : The Advantages of Careful Seeding*, Siam,
741 Philadelphia, 2007.

742 [17]Lee Raymond Dice, Measures of the Amount of Ecologic Association Between Species,
743 *Ecology.* 26 (1945) 297–302.

744 [18]M.R. Avendi, A. Kheradvar, H. Jafarkhani, A combined deep-learning and deformable-
745 model approach to fully automatic segmentation of the left ventricle in cardiac MRI, *Med.*
746 *Image Anal.* 30 (2016) 108–119. doi:10.1016/j.media.2016.01.005.

747 [19]X. Alba, K. Lekadir, M. Pereanez, P. Medrano-Gracia, A.A. Young, A.F. Frangi,
748 Automatic initialization and quality control of large-scale cardiac MRI segmentations,
749 *Med. Image Anal.* 43 (2018) 129–141. doi:10.1016/j.media.2017.10.001.

750 [20]A.H. Curiale, F.D. Colavecchia, P. Kaluza, R.A. Isoardi, G. Mato, *Automatic Myocardial*
751 *Segmentation by Using A Deep Learning Network in Cardiac MRI*, Ieee, New York,
752 2017.

753 [21]P.V. Tran, *A Fully Convolutional Neural Network for Cardiac Segmentation in Short-Axis*
754 *MRI*, CoRR. abs/1604.00494 (2016).

755 [22]J. Schwitter, C.M. Wacker, N. Wilke, N. Al-Saadi, E. Sauer, K. Huettle, S.O. Schönberg,
756 A. Luchner, O. Strohm, H. Ahlstrom, T. Dill, N. Hoebel, T. Simor, MR-IMPACT II:
757 Magnetic Resonance Imaging for Myocardial Perfusion Assessment in Coronary artery
758 disease Trial: perfusion-cardiac magnetic resonance vs. single-photon emission
759 computed tomography for the detection of coronary artery disease: a comparative
760 multicentre, multivendor trial, *Eur Heart J.* 34 (2013) 775–781.
761 doi:10.1093/eurheartj/ehs022.
762