Journal of Cardiovascular Computed Tomography

Identification of subclinical cardiac amyloidosis in aortic stenosis patients undergoing transaortic valve replacement using radiomic analysis of computed tomography myocardial texture
--Manuscript Draft--

Corresponding Author:	Gianluca Pontone, MD, PhD Centro Cardiologico Monzino IRCCS Milan, ITALY
First Author:	Francesca Lo Iacono
Order of Authors:	Francesca Lo Iacono
	Riccardo Maragna, MD
	Marco Gugliemo, MD, FEACVI, FSCCT
	Mattia Chiesa, PhD
	Laura Fusini, MD
	Andrea Annoni, MD
	Mario Babbaro, MD
	Andrea Baggiano, MD
	Maria Ludovica Carerj, MD
	Francesco Cilia, MD
	Alberico Del Torto, MD
	Alberto Formenti, MD
	Maria Elisabetta Mancini, MD
	Francesca Marchetti, MD
	Manuela Muratori, MD
	Saima Mushtaq, MD
	Marco Penso
	Sergio Pirola, MD
	Alessandra Volpe, MD
	Andrea Igoren Guaricci, MD
	Marianna Fontana, MD, PhD
	Gloria Tamborini, MD
	Thomas Treibel, MD
	James Moon, MD
	Valentina D.A Corino, PhD
	Gianluca Pontone, MD, PhD, FESC, FSCCT

Abstract:	Background: Cardiac amyloidosis (CA) is an increasingly diagnosed pathology sharing several phenotypical features with aortic stenosis (AS). As diagnosing the two diseases has important prognostic and therapeutic implications, this study aims to identify a set

of stable and discriminative radiomic features derived from cardiac computed tomography (CCT) to differentiate them.

Methods: Forty-two patients were included in the study. For each patient, 107 radiomics features were extracted and evaluated by means of geometrical transformations (translations) to the region of interests (ROIs), and ICC (intra class correlation coefficient) computation. A stratified 7-fold cross validation (k=7) was performed to split data into learning, validation and test set. Three features selection methods (Wilcoxon signed rank-based method and/or LASSO regression) and five machine learning classifiers Results: Ninety radiomic features satisfied the robustness criteria and 10 were kept after feature selection. The best results were obtained using the logistic regression classifier, combined with Wilcoxon signed rank and LASSO regression, obtaining an accuracy of 95% ± 7% and sensitivity and specificity both equal to 95% ± 12% in the test set.

Conclusions: In this study, radiomics has shown promising results in distinguishing left ventricle hypertrophy caused by CA from AS and might be used as non-invasive tool able to support clinical decision making.

Suggested Reviewers:	Marco Francone marco.francone@hunimed.eu
	Antonio Esposito antonio.esposito@hsr.it
	Carlo De Cecco carlo.dececco@emory.edu
Opposed Reviewers:	

Dear Editor,

please find enclosed our manuscript entitled "Identification of subclinical cardiac amyloidosis in aortic stenosis patients undergoing transaortic valve replacement using radiomic analysis of computed tomography myocardial texture" to be considered for publication in International Journal of Cardiology by Riccardo Maragna, Marco Guglielmo, Mattia Chiesa, Laura Fusini, Andrea Annoni, Mario Babbaro, Andrea Baggiano, Maria Ludovica Carerj, Francesco Cilia, Alberico Del Torto, Alberto Formenti, Maria Elisabetta Mancini, Francesca Marchetti, Manuela Muratori, Saima Mushtaq, Marco Penso, Sergio Pirola, Luigi Tassetti, Alessandra Volpe, Andrea Igoren Guaricci, Marianna Fontana, Gloria Tamborini, Thomas Treibel, James Moon, Valentina D. A. Corino, Gianluca Pontone, and myself.

The manuscript is original, with no portion under simultaneous consideration for publication elsewhere or previously published. The study conforms to the Declaration of Helsinki, and all participants gave informed consent of the study that was approved by the Institutional Ethical Committee.

Thank you for your attention.

Best regards
Francesca Lo Iacono

Identification of subclinical cardiac amyloidosis in aortic stenosis patients undergoing transaortic valve replacement using radiomic analysis of computed tomography myocardial texture

Short title: Radiomics for cardiac amyloidosis identification

Francesca Lo Iacono[1], Riccardo Maragna[2], Marco Guglielmo[3], Mattia Chiesa[1,2], Laura Fusini[1,2], Andrea Annoni[2], Mario Babbaro[2], Andrea Baggiano[2,4], Maria Ludovica Carerj[2,5], Francesco Cilia[2], Alberico Del Torto[2], Alberto Formenti[2], Maria Elisabetta Mancini[2], Francesca Marchetti[2], Manuela Muratori[2], Saima Mushtaq[2], Marco Penso[1,2], Sergio Pirola[2], Luigi Tassetti[2], Alessandra Volpe[2], Andrea Igoren Guaricci[6], Marianna Fontana[7], Gloria Tamborini[1], Thomas Treibel[8,9], James Moon[8,9], *Valentina D. A. Corino[1,2], *Gianluca Pontone[2]

[1] Department of Electronics, Information and Bioengineering, Politecnico di Milano, Italy

[2] Centro Cardiologico Monzino IRCCS, Milan, Italy

[3] University Medical Center Utrecht, Netherland

[4] Department of Clinical Sciences and Community Health, University of Milan, Milan, Italy.

[5] Diagnostic and Interventional Radiology Unit, Department of Biomedical Sciences and Morphological and Functional Imaging, "G. Martino" University Hospital Messina, Messina, Italy

[6] Institute of Cardiovascular Disease, Department of Interdisciplinary Medicine, University Hospital "Policlinico" of Bari, Bari, Italy

[7] National Amyloidosis Centre, University College London, Royal Free Campus, London, United Kingdom

[8] Institute of Cardiovascular Science, University College London, London, United Kingdom

[9] Barts Heart Centre, St Bartholomew's Hospital, London, United Kingdom

*: equal senior contribution

Address for correspondence: Gianluca Pontone, MD, PhD, FESC, FSCCT
Via C. Parea 4, 20138 Milan, Italy
Ph: +39-02-58002574
Fax: +39-02-58002231
E-mail: gianluca.pontone@ccfm.it

Conflict of interest

TOC Summary

Cardiac amyloidosis (CA) and aortic stenosis (AS) present phenotypical common features and differentiating them represents a challenge to properly assess patient prognosis and therapy. In this study, a set of robust radiomic features is derived from cardiac computed tomography to build a machine learning model to distinguish CA from AS patients. An overall diagnostic accuracy of 95% is reached. These promising results show that radiomics might be used as non-invasive tool able to support clinical decision making.

Abstract

Background: Cardiac amyloidosis (CA) is an increasingly diagnosed pathology sharing several phenotypical features with aortic stenosis (AS). As diagnosing the two diseases has important prognostic and therapeutic implications, this study aims to identify a set of stable and discriminative radiomic features derived from cardiac computed tomography (CCT) to differentiate them.

Methods: Forty-two patients were included in the study. For each patient, 107 radiomics features were extracted and evaluated by means of geometrical transformations (translations) to the region of interests (ROIs), and ICC (intra class correlation coefficient) computation. A stratified 7-fold cross validation (k=7) was performed to split data into learning, validation and test set. Three features selection methods (Wilcoxon signed rank-based method and/or LASSO regression) and five machine learning classifiers (k-nearest neighbors, support vector classifier, decision tree, logistic regression and gradient boosting) were tested.

Results: Ninety radiomic features satisfied the robustness criteria and 10 were kept after feature selection. The best results were obtained using the logistic regression classifier, combined with Wilcoxon signed rank and LASSO regression, obtaining an accuracy of 95% $\pm$ 7% and sensitivity and specificity both equal to 95% $\pm$ 12% in the test set.

Conclusions: In this study, radiomics has shown promising results in distinguishing left ventricle hypertrophy caused by CA from AS and might be used as non-invasive tool able to support clinical decision making.

1

Abbreviations:

AL: amyloid light chain

AS: aortic stenoses

ATTR: amyloid transthyretin

CA: cardiac amyloidosis

CCT: cardiac computed tomography

EACVI: European association cardiovascular imaging

ECG: electrocardiogram

ECV: extracellular volume

IBSI: Image Biomarker Standardization Initiative

ICC: intraclass correlation coefficient

LASSO: Least Absolute Shrinkage and Selection Operator

LVEDV: left ventricle end-diastolic volume

LVEF: left ventricle ejection fraction

LVESV: left ventricle end-systolic volume

ML: machine learning

NGTDM: neighbouring gray tone difference matrix

wtATTR: wild type amyloid transthyretin

vATTR: variable amyloid transthyretin

ROI: region of interest

TAVI: transaortic valve implantation

TTE: transthoracic echocardiography

KNN: k-nearest neighbor

SVC: support vector classifier

DT: decision tree

LR: logistic regression

GB: gradient boosting

ROC: receiver operating characteristic

1. Introduction

Cardiac amyloidosis (CA) is an increasingly diagnosed condition caused by conformational changes in amyloidogenic proteins leading to amyloid fibril deposition in the heart [1]. The two predominant amyloid proteins found in the heart are immunoglobulin light chain (AL) and transthyretin (ATTR) [2]. ATTR-CA occurs in two most common forms: wild-type, or senile, transthyretin (wtATTR), more prevalent in elderly people, and hereditary or variable TTR (vATTR), genetic autosomal dominant disease [3]. Clinical presentation associated to cardiac amyloidosis shows an increased biventricular wall thickness, myocardial stiffening and restrictive physiology of the left and right ventricles [2] caused by amyloid fibril deposition in the extracellular myocardial space.

Similar myocardial remodeling processes affect the heart of patients with aortic stenosis (AS). AS is the most common cause of valvular heart disease [4]. It is characterized by a progressive aortic valve narrowing leading to an increase in afterload and wall stress, compensated with a gradual hypertrophy of left ventricle [5].

Therefore, as CA shares several common phenotypical features with AS and considering the high prevalence of subclinical CA among AS undergoing TAVI, the differential diagnosis of these two entities has important prognostic and therapeutic implications.

In this framework, radiomics can be a non-invasive tool useful to perform differential diagnosis starting from medical images such as cardiac computed tomography (CCT) usually used for interventional planning of AS patients undergoing TAVI. Radiomics is an emerging research field aimed to improve diagnosis, characterization, and prognosis using quantitative features extracted from medical images. Radiomics is widely employed in oncology for tumors characterization [6,7], treatment response [8,9,10] and overall survival analysis [11,12]. Recently, it has been proposed in the cardiovascular field to improve diagnostic accuracy, patients cardiac risk prediction and stratification [13,14,15,16,17,18,19,20,21,22].

As radiomics generates hundreds of features, a crucial step of radiomic workflow is selecting features according to their relevance with respect to the clinical question of interest. Assessing features robustness is a necessary preliminary step in the process of feature selection [23]. The most applied techniques to evaluate features reliability are test-retest and multiple delineations of the ROI [15,17,18], methods affected by a time-consuming intrinsic limit: test-retest requires multiple scan acquisitions, whereas multiple delineations require several ROI segmentations. To overcome these problems, in the current study, for the first time in CCT, features robustness was assessed performing geometrical transformations of the ROIs [23]. To mimic multiple manual delineations, small ROI transformations are applied to simulate errors due to manual delineation thus assessing feature stability. In addition, large ROI transformations are performed to assess feature discrimination capacity. The underlying hypothesis is that robust features need to be stable, i.e., similar for small transformations, and discriminative, i.e., different, for large transformations. Studies on feature robustness already exist in oncological field [23,24,25], whereas a very limited research is available in cardiovascular radiomics [26,15].

The aim of this study is to identify a set of stable and discriminative radiomic features derived from CCT to differentiate left ventricle hypertrophy due to CA versus AS.

2. Methods

2.1. Study population and baseline characteristics

Twenty-one patients with CA (Mean age and interquartile range: 74 [67-76]; frequency of male: 67%) were randomly extracted by our database.

ATTR patients underwent bone scintigraphy using 3,3-diphosphono-1,2-propanodicarboxylicacid (DPD), whereas AL patients also underwent SAP scintigraphy. For ATTR, cardiac amyloidosis was defined by presence of ATTR amyloid in a myocardial biopsy (Congo red and immunohistochemical staining) or positive DPD scintigraphy. All

5

ATTR patients also underwent sequencing of exons 2, 3, and 4 of the TTR gene. For AL, systemic AL amyloidosis was proven with biopsies from non-cardiac tissues.

As for CA patients, we randomly extracted a population of 21 AS patients (Mean age and interquartile range: 82 [78-84]; Male: 53%) from a cohort of consecutive patients with severe AS referred to CCT for TAVI planning. The concomitant presence of CA was excluded by bone scintigraphy and/or cardiac magnetic resonance.

All patients underwent comprehensive evaluation with transthoracic echocardiography using commercially available equipment (iE33 or Epiq, Philips Medical System, or Vivid-9, GE Healthcare) measuring LV end-diastolic (LVEDV) and end-systolic (LVESV) volumes indexed for body surface area and LV ejection fraction (LVEF).

The institutional Ethical Committee approved the study, and all the patients signed the informed consent.

2.2. CCT scan protocol

CCT examinations were performed using 256-slices (Revolution CT; GE Healthcare, Milwaukee, WI) or 320-slices wide volume coverage CT scanner (Aquilion ONE Vision™; Canon Medical Systems Corp., Tokyo, Japan). No premedication with beta-blockers or nitrates was added before CT acquisition. Revolution CT scans were acquired using the following parameters: peak tube voltage, 100 kV; Detector collimation: 160 mm using 256 rows by 0.625 mm on Z axis. Detector geometry: 256 rows by 832 detection elements per row. High contrast spatial resolution: 0.23 mm. Slice thickness: 0.625; gantry rotation time, 280 ms; prospective triggering; and iterative reconstruction algorithm (ASIR-V; GE Healthcare). Patients received a fixed dose of 50 ml bolus of contrast medium (400 mg of iodine per milliliter, Iomeprol; Bracco, Milan, Italy) despite the BMI via an antecubital vein at an infusion rate of 5 ml s−1 followed by 50 ml of saline solution at 5 ml·s−1. Aquilion ONE Vision scans were acquired using the following parameters: peak tube voltage, 120 kV; tube

current–time product, 160 mAs; section collimation, 320 detector rows, 1.2-mm section thickness; gantry rotation time, 275 msec. The iodinated contrast material used in this case was iohexol (300 mg of iodine per milliliter Omnipaque 300; Nycomed Amersham, Oslo, Norway) at a standard dose of 1mL/kg and injection rate of 3ml/sec without a saline chaser. A body mass index (BMI)-adapted protocol was used for the tube current. The bolus tracking technique was used to synchronize the arrival of contrast material at the aortic root with the start of acquisition.

2.3. Images segmentation, preprocessing and radiomics features extraction

For each patient, the left ventricular wall was manually segmented by expert cardiac imagers with level III European Association of Cardiovascular Imaging (EACVI) [27].

Image preprocessing was performed to reduce the imaging-related variability: a 3D Gaussian filter with a 3x3x3 voxel kernel and σ=0.5 was used to denoise the images. Then, voxel size resampling to an isotropic resolution of 2 mm (as in [28]) was performed with B-spline interpolation.

The extraction of radiomic features was performed using Pyradiomics 3.0 [29]. A total of 107 features, belonging to different categories, were extracted: shape and size (14 features), first order statistics (18 features) and textural variables (75 features) [29]. A fixed-bin width histogram discretization (0.5 Hounsfield units per bin) was used prior to features extraction. Since features were extracted with Pyradiomics, they were compatible with the Image Biomarker Standardization Initiative (IBSI) [30].

Radiomic features underwent a series of features selection steps:

a) Selection of stable and discriminant features

Features stability and discrimination capacity were assessed using geometrical transformations (translations) of the ROIs as in [23]. The entire workflow is implemented in

MATLAB 2017a (Mathworks, Natick, MA, USA) and applied to first order and textural features.

Various entity translations were applied to the ROIs, along both the x (medial-lateral) and y (antero-posterior) directions. The minimal entity translation was ±0.5% the length of the bounding box surrounding the ROI in the direction of interest **(Figure 1A)**, and the maximal entity translation was ±30% **(Figure 1B)**. Eleven different translations ([0.5, 1, 2, 4, 6, 8, 10, 15, 20, 25, 30] %) were defined and, for each of them, 4 shifts of the ROI were applied, one positive and one negative along both the x and y axis direction. Radiomic features were computed on each translated ROI and compared to the ones obtained with the original one segmented by expert cardiac radiologists. For each translation, the percentage of volume overlap was computed to assess whether most of the ROI was considered when transforming the ROI.

b) Selection of non-redundant features

The second step consisted of a correlation-based feature selection, performed to ensure a set of features with low internal redundancy.

c) Selection of more relevant features

Three feature selection methods were tested [31]. The first method included only features statistically different between CA and AS (p-value < 0.05). The second method selected features using LASSO method. The third method combined the first two methods: only significant features were considered when applying LASSO algorithm.

2.4. Machine Learning Classification

Five machine learning classification models were implemented in Scikit-Learn Python library to identify the radiomic features that better distinguish patients with CA from patients affected by AS.

A stratified 7-fold cross validation was performed to split data into train and test set preserving the percentage of samples for each class.

In all ML models, training set was further split into learning and validation set (70% and 30% of the data, respectively) preserving the same class proportions observed in the original dataset.

Before classification, data belonging to the three sets were z-scored. The mean and standard deviation were computed on the training set and re-applied on the validation and test set. Five ML models were considered: KNN, SVC, DT, LR, GB. In each train-test split, classifiers were trained on the learning set and applied to the validation set. Sensitivity, specificity, diagnostic accuracy, and area under the curve (AUC) of five ML models were calculated.

The entire workflow is shown in **Figure 2**.

2.5. Statistical Analysis

Statistical analysis was performed using MATLAB 2017a (Mathworks, Natick, MA, USA). Descriptive statistics was used to characterize the study population and Student's independent t-test, Wilcoxon Rank sum test tests, Chi-square, or Fisher's exact test were used as appropriate to compare the distribution of continuous and categorical variables among AS vs CA groups.

Radiomic features robustness was measured by the intraclass correlation coefficient (ICC). For each feature calculated from a single ROI, 4 ICCs with their mean (ICC_{mean}) were calculated. A total of 11 ICC_{mean} were computed. Minimal and maximal entity translation ICCs were employed to assess features stability and discrimination capacity. Following general guidelines [32], two ICC threshold values were identified to select robust features: ICC = 0.75, which indicates good agreement between data, and ICC = 0.5, which reflects poor similarity. Thus, features having an ICC_{mean} higher than 0.75 for minimal entity translation

and an ICC_{mean} lower than 0.5 for maximal entity translation, were considered stable and discriminative.

To select the non-redundant features, Spearman correlation coefficient ρ was computed for each pair of features and in case a pair had |ρ| > 0.85, only the feature with the lower mean Spearman coefficient with all the others (n-2) features was selected, 'n' being the total number of features.

3. Results

A total of 42 patients were enrolled. Among the 21 patients with CA, 11 were affected by AL-CA and 10 by ATTR-CA. In twenty-one out of 42 patients with AS (AS group), CA was excluded by cardiac magnetic resonance and by cardiac tracer uptake on bone scintigraphy with Perugini grade

< 2 when CMR was equivocal.

Table 1 shows patients baseline characteristics.

Regarding to feature selection, first step identified 76 radiomic features as robust (18 first-order statistics and 58 textural features) as shown in **Figure 3 Panel A**. These features have an ICC_{mean} higher than 0.75 at the minimal translation and an ICC_{mean} lower than 0.5 at the maximal entity translation. Features not responding to the two previous criteria are shown in **Figure 3 Panel B**. After this first feature selection step, 90 features were considered, i.e., the 76 robust ones and the 14 shape-based features. **Table 2** reports the list of all radiomic features with the ICC values for minimal and maximal translations. **Figure 4** shows the percentages of overlap between the original and the translated ROI as a function of the different entity translation. The median overlap for the minimal translation entity was 96.7 % while the median overlap for the maximal translation was 8.9 %. At each train-test split, an

average of 24 radiomic features, ranging from 22 to 26, were further selected as non-redundant with the second feature selection step.

Finally, 10±1 features were selected using Wilcoxon Signed Rank test, 7±1 using the LASSO method and 7±1 using the combination of the two methods. **(Table 3)**.

Figure 5 shows the performance of the five ML classifiers in the validation set using the three different feature selection methods. Comparing the ML models, a general trend showing higher performance in LR can be observed. The LR classifier, applied with Wilcoxon signed rank-based method and its combination with the LASSO regression, shows the best performances (increasing mean accuracy of more than 0.10), resulting in statistically significant differences with DT and GB models. Thus, LR classifier, with these two feature selection methods, was selected to be used in the test set in each split.

Figure 6 shows the ROC curve, with AUC 0f 0.94, for the test set considering the LR classifier used with Wilcoxon signed rank and LASSO regression combination as feature selection method.

4. Discussion

The main finding of our study is that a CT dataset acquired for the usual planning of TAVI in AS patients can be used to extract radiomic features to diagnose subclinical CA. More specifically 90 radiomic features satisfied the robustness criteria and 10 were kept after feature selection. The best results were obtained using the LR classifier, combined with Wilcoxon signed rank and LASSO regression, obtaining an accuracy of our model of 95%±7% in the test set (86% ± 4% in the validation set).

Endomyocardial biopsy is identified as the gold standard of CA diagnosis. However, it is invasive, and some risks are associated with this technique. Nowadays, it is widely recognized that bone scintigraphy represents a reliable diagnostic tool for CA, in particular for the ATTR variant, avoiding endomyocardial biopsy. Other imaging techniques have been

investigated. Echocardiography is valuable and widely accessible tool for investigating heart failure [33] but most echocardiographic parameters do not provide CA diagnosis in the early stages of disease and do not to allow to distinguish CA from other restrictive or hypertrophic cardiomyopathy [34]. Cardiac magnetic resonance (CMR) imaging offers a greater diagnostic power in CA, but at the same time it is expensive, contraindicated in a substantial proportion of patients and could provide false-positive and false-negative results [33].

Recently, several studies have also developed artificial intelligence algorithms aimed to diagnose CA.

Goto et al. [35] developed two deep learning algorithms using electrocardiogram (ECG) and echocardiography data coming from respectively 3 and 5 academic medical centers. The ECG-based model reached a mean C-statistic of 0.86 in differentiating 587 CA patients from 8612 controls while the echocardiography-based model achieved better performances with a mean C-statistics of 0.95 in distinguishing 609 CA patients from 303 controls.

Also, CMR imaging was employed in different studies [36]. Zhou et al. classified 139 patients (79 CA positive vs 60 controls) by employing CMR radiomics based machine learning algorithm with a mean accuracy of 80%, and Martini et al. [37] developed a deep learning model to diagnose CA in 206 patients achieving an accuracy of 88%.

As compared with the previous experience, several points of strength could be considered in our study. First, our results based on CCT radiomics obtained an accuracy of 95%, significantly higher as compared previous studies. Second, to the best of our knowledge, this is the first study to employing CCT for amyloidosis identification.

This has several clinical implications. Indeed, several studies showed that CA is frequent (11.8%) [38] in patients with severe AS referred for TAVI and the challenge, in this context, is to differentiate a wooden horse (lone AS) from a Trojan horse (AS with CA) [39]. Nitsche et al. [38] developed in this setting the RAISE [remodelling, age, injury, system, and electrical] score, in order to predict the presence of CA in patients with severe AS. This multiparameter

score assessed age >85 years (1 point), history of carpal tunnel syndrome (3 points), presence of right bundle branch block (2 points), Sokolow/Lyon index <1.9 mV (1 point), high sensitivity troponin level >20 ng/ml (1 point), and E/A ratio >1.4 (1 point), with a sensitivity and specificity of 84% and 94%, respectively. Alternatively to RAISE score, Oda et al. [40] propose the measurement of extracellular volume (ECV) by using CCT dataset. However, this approach requires triple scan acquisition (unenhanced scan arterial phase acquisition and late scan) and higher volume of contrast agent. On the contrary our approach is easily performed by using the single arterial phase CCT dataset acquired during the usual diagnostic work-up of these patients.

Moreover, within radiomic workflow, assessing features robustness is an important preliminary step to improve generalizability of radiomic models. In most of the studies, radiomic features stability has been tested using multiple delineation method on a subgroup of patients [15,17,18], evaluating inter-observer agreement amongst different operators using the ICC index. These methods are very time-consuming. This is the first study performing features robustness assessment using image perturbations in cardiac CT based on small and large entity translation as in [23].

Considering radiomics features, five appeared in each split of each features selection method (10th percentile, signal maximum, small area emphasis, shape elongation and maximum 2D diameter column). Five additional features (energy, major axis length, flatness, sphericity, and small dependence emphasis) were included in all the splits using Wilcoxon signed rank-based method and excluded by LASSO regression in most.

Having high values of sensitivity represents a promising result from the clinical point of view as recognizing CA, whose prognosis results worse compared to AS, is the main interest of this study.

Several study limitations should be considered. First, the sample size employed in the study is limited thus these results are only preliminary and should be confirmed by a larger dataset. In addition, the considered patients were randomly extracted from a cohort of AS and CA patients referred for CCT. Finally, CA and AS were explored separately without including patients affected by both pathologies, this latter group of patients might be included in a further study.

5. Conclusions

In conclusion in the current study a set of stable and robust radiomics features, able to discriminate CA and AS patients, has been found. Features robustness assessment has been performed using small and large transformations of the ROIs. This method overcomes time-consuming intrinsic limits which affect the most applied techniques evaluating feature reliability. This study shows promising results in differentiating CA from AS that might be applied to clinical routine available images.

References

1. Baker KR, Rice L. The Amyloidoses: Clinical Features, Diagnosis and Treatment. *Methodist DeBakey Cardiovascular Journal*. 2012;8(3):3. doi:10.14797/mdcj-8-3-3

2. Bonelli A, Paris S, Nardi M, et al. Aortic Valve Stenosis and Cardiac Amyloidosis: A Misleading Association. *J Clin Med*. 2021;10(18):4234. doi:10.3390/jcm10184234

3. GHERASIM, L. Amyloid Cardiopathy and Aortic Stenosis. *Maedica (Bucur)*. 2021;16(3):473-481. doi:10.26574/maedica.2020.16.3.473

4. Elmariah S. Patterns of Left Ventricular Remodeling in Aortic Stenosis: Therapeutic Implications. *Curr Treat Options Cardio Med*. 2015;17(7):31. doi:10.1007/s11936-015-0391-0

5. Shah ASV, Chin CWL, Vassiliou V, et al. Left Ventricular Hypertrophy With Strain and Aortic Stenosis. *Circulation*. 2014;130(18):1607-1616. doi:10.1161/CIRCULATIONAHA.114.011085

6. La Greca Saint-Esteven A, Vuong D, Tschanz F, et al. Systematic Review on the Association of Radiomics with Tumor Biological Endpoints. *Cancers*. 2021;13(12):3015. doi:10.3390/cancers13123015

7. Corino VDA, Montin E, Messina A, et al. Radiomic analysis of soft tissues sarcomas can distinguish intermediate from high-grade lesions. *Journal of Magnetic Resonance Imaging*. 2018;47(3):829-840. doi:10.1002/jmri.25791

8. Kothari G. Role of radiomics in predicting immunotherapy response. *Journal of Medical Imaging and Radiation Oncology*. 2022;66(4):575-591. doi:10.1111/1754-9485.13426

9. Relevance of apparent diffusion coefficient features for a radiomics-based prediction of response to induction chemotherapy in sinonasal cancer - Bologna - 2022 - NMR in Biomedicine - Wiley Online Library. Accessed June 17, 2022. https://analyticalsciencejournals.onlinelibrary.wiley.com/doi/10.1002/nbm.4265

10. Zhang B, Ouyang F, Gu D, et al. Advanced nasopharyngeal carcinoma: pre-treatment prediction of progression based on multi-parametric MRI radiomics. *Oncotarget*. 2017;8(42):72457-72465. doi:10.18632/oncotarget.19799

11. Bologna M, Corino V, Calareso G, et al. Baseline MRI-Radiomics Can Predict Overall Survival in Non-Endemic EBV-Related Nasopharyngeal Carcinoma Patients. *Cancers (Basel)*. 2020;12(10):2958. doi:10.3390/cancers12102958

12. Zhao L, Gong J, Xi Y, et al. MRI-based radiomics nomogram may predict the response to induction chemotherapy and survival in locally advanced nasopharyngeal carcinoma. *Eur Radiol*. 2020;30(1):537-546. doi:10.1007/s00330-019-06211-x

13. Ponsiglione A, Stanzione A, Cuocolo R, et al. Cardiac CT and MRI radiomics: systematic review of the literature and radiomics quality score assessment. *Eur Radiol*. 2022;32(4):2629-2638. doi:10.1007/s00330-021-08375-x

14. Shang J, Guo Y, Ma Y, Hou Y. Cardiac computed tomography radiomics: a narrative review of current status and future directions. *Quant Imaging Med Surg.* 2022;12(6):3436-3453. doi:10.21037/qims-21-1022

15. Hu W, Wu X, Dong D, et al. Novel radiomics features from CCTA images for the functional evaluation of significant ischaemic lesions based on the coronary fractional flow reserve score. *Int J Cardiovasc Imaging.* 2020;36(10):2039-2050. doi:10.1007/s10554-020-01896-4

16. Lin A, Kolossváry M, Yuvaraj J, et al. Myocardial Infarction Associates With a Distinct Pericoronary Adipose Tissue Radiomic Phenotype. *JACC: Cardiovascular Imaging.* 2020;13(11):2371-2383. doi:10.1016/j.jcmg.2020.06.033

17. Oikonomou EK, Williams MC, Kotanidis CP, et al. A novel machine learning-derived radiotranscriptomic signature of perivascular fat improves cardiac risk prediction using coronary CT angiography. *European Heart Journal.* 2019;40(43):3529-3543. doi:10.1093/eurheartj/ehz592

18. Shang J, Ma S, Guo Y, et al. Prediction of acute coronary syndrome within 3 years using radiomics signature of pericoronary adipose tissue based on coronary computed tomography angiography. *Eur Radiol.* 2022;32(2):1256-1266. doi:10.1007/s00330-021-08109-z

19. Nam K, Suh YJ, Han K, Park SJ, Kim YJ, Choi BW. Value of Computed Tomography Radiomic Features for Differentiation of Periprosthetic Mass in Patients With Suspected Prosthetic Valve Obstruction. *Circ: Cardiovascular Imaging.* 2019;12(11):e009496. doi:10.1161/CIRCIMAGING.119.009496

20. Raisi-Estabragh Z, Jaggi A, Gkontra P, et al. Cardiac Magnetic Resonance Radiomics Reveal Differential Impact of Sex, Age, and Vascular Risk Factors on Cardiac Structure and Myocardial Tissue. *Front Cardiovasc Med.* 2021;8:763361. doi:10.3389/fcvm.2021.763361

21. Izquierdo C, Casas G, Martin-Isla C, et al. Radiomics-Based Classification of Left Ventricular Non-compaction, Hypertrophic Cardiomyopathy, and Dilated Cardiomyopathy in Cardiovascular Magnetic Resonance. *Front Cardiovasc Med.* 2021;8:764312. doi:10.3389/fcvm.2021.764312

22. Rauseo E, Izquierdo Morcillo C, Raisi-Estabragh Z, et al. New Imaging Signatures of Cardiac Alterations in Ischaemic Heart Disease and Cerebrovascular Disease Using CMR Radiomics. *Front Cardiovasc Med.* 2021;8:716577. doi:10.3389/fcvm.2021.716577

23. Bologna M, Corino VDA, Montin E, et al. Assessment of Stability and Discrimination Capacity of Radiomic Features on Apparent Diffusion Coefficient Images. *J Digit Imaging.* 2018;31(6):879-894. doi:10.1007/s10278-018-0092-9

24. Gevaert O, Mitchell LA, Achrol AS, et al. Glioblastoma multiforme: exploratory radiogenomic analysis by using quantitative image features. *Radiology.* 2014;273(1):168-174. doi:10.1148/radiol.14131731

25. Aerts HJWL, Velazquez ER, Leijenaar RTH, et al. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. *Nat Commun.* 2014;5:4006. doi:10.1038/ncomms5006

26. Raisi-Estabragh Z, Gkontra P, Jaggi A, et al. Repeatability of Cardiac Magnetic Resonance Radiomics: A Multi-Centre Multi-Vendor Test-Retest Study. *Frontiers in Cardiovascular Medicine.* 2020;7. Accessed June 15, 2022. https://www.frontiersin.org/article/10.3389/fcvm.2020.586236

27. Pontone G, Moharem-Elgamal S, Maurovich-Horvat P, et al. Training in cardiac computed tomography: EACVI certification process. *European Heart Journal - Cardiovascular Imaging.* 2018;19(2):123-126. doi:10.1093/ehjci/jex310

28. Leijenaar RT, Bogowicz M, Jochems A, et al. Development and validation of a radiomic signature to predict HPV (p16) status from standard CT imaging: a multicenter study. *The British Journal of Radiology.* 2018;91(1086):20170498.

29. van Griethuysen JJM, Fedorov A, Parmar C, et al. Computational Radiomics System to Decode the Radiographic Phenotype. *Cancer Res.* 2017;77(21):e104-e107. doi:10.1158/0008-5472.CAN-17-0339

30. Zwanenburg A, Vallières M, Abdalah MA, et al. The Image Biomarker Standardization Initiative: Standardized Quantitative Radiomics for High-Throughput Image-based Phenotyping. *Radiology.* 2020;295(2):328-338. doi:10.1148/radiol.2020191145

31. James G, Witten D, Hastie T, Tibshirani R, eds. *An Introduction to Statistical Learning: With Applications in R.* Springer; 2013.

32. Koo TK, Li MY. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *Journal of Chiropractic Medicine.* 2016;15(2):155-163. doi:10.1016/j.jcm.2016.02.012

33. Mattana F, Muraglia L, Girardi F, et al. Clinical application of cardiac scintigraphy with bone tracers: controversies and pitfalls in cardiac amyloidosis. *VP.* Published online 2022. doi:10.20517/2574-1209.2021.87

34. Jung MH, Chang S, Han EJ, Youn JC. Multimodal Imaging and Biomarkers in Cardiac Amyloidosis. *Diagnostics.* 2022;12(3):627. doi:10.3390/diagnostics12030627

35. Goto S, Mahara K, Beussink-Nelson L, et al. Artificial intelligence-enabled fully automated detection of cardiac amyloidosis using electrocardiograms and echocardiograms. *Nat Commun.* 2021;12(1):2726. doi:10.1038/s41467-021-22877-8

36. Zhou XY, Tang CX, Guo YK, et al. Diagnosis of Cardiac Amyloidosis Using a Radiomics Approach Applied to Late Gadolinium-Enhanced Cardiac Magnetic Resonance Images: A Retrospective, Multicohort, Diagnostic Study. *Front Cardiovasc Med.* 2022;9:818957. doi:10.3389/fcvm.2022.818957

37. Martini N, Aimo A, Barison A, et al. Deep learning to diagnose cardiac amyloidosis from cardiovascular magnetic resonance. *Journal of Cardiovascular Magnetic Resonance.* 2020;22(1):84. doi:10.1186/s12968-020-00690-4

38. Prevalence and Outcomes of Concomitant Aortic Stenosis and Cardiac Amyloidosis | Elsevier Enhanced Reader. doi:10.1016/j.jacc.2020.11.006

39. Pibarot P, Lancellotti P, Narula J. Concomitant Cardiac Amyloidosis in Severe Aortic Stenosis: The Trojan Horse?∗. *Journal of the American College of Cardiology*. 2021;77(2):140-143. doi:10.1016/j.jacc.2020.11.007

40. Oda S, Kidoh M, Takashio S, et al. Quantification of Myocardial Extracellular Volume With Planning Computed Tomography for Transcatheter Aortic Valve Replacement to Identify Occult Cardiac Amyloidosis in Patients With Severe Aortic Stenosis. *Circ: Cardiovascular Imaging*. 2020;13(5):e010358. doi:10.1161/CIRCIMAGING.119.010358

Figure legends

Figure 1: Example of translations applied to the ROIs. (a) Minimal entity translation: dashed red line represents a translation of 0.5% in the *x* direction. (b) Maximal entity translation: dashed red line represents a translation of 30% in the *x* direction. In both figures, continuous black lines represent the contours of the original ROIs. ROI: region of interest.

Figure 2: Machine learning workflow. (a) Visual representation of the learning, validation and test split using 7-fold cross-validation. (b) Visual representation of machine learning workflow.

Figure 3: ICCs as a function of the translation entity. (a) Stable and discriminative features: ICCs at the minimal entity translation (0.5%) is higher than the threshold value 0.75 (ICC$_{min}$) and ICCs calculated at the maximal entity translation (30%) are lower than the threshold value 0.5 (ICC$_{max}$). superimposed in red, ICC mean and standard deviation. (b) Non-valid features: black lines represent ICC trend for features whose ICC at the maximal entity translation is higher than ICC$_{max}$; grey lines represent ICC trend for features whose ICC at the minimal entity translation is lower than ICC$_{min}$.

Figure 4: Boxplots representing the overlap percentage between translated ROIs and original ROIs, at each translation value.

Figure 5: Accuracy, sensitivity, and specificity for the validation set. (a) Features selection based on the Wilcoxon signed rank-based method. (b) Features selection based on the LASSO regression method. (c) Features selection based on the combination of Wilcoxon signed rank-based method and LASSO regression. KNN: K-nearest neighbor, SVC: support vector classifier, DT: decision tree, LR: logistic regression, GB: gradient boosting.

Figure 6: Receiver Operating Characteristics (ROC) curve computed from the test set

(a)

(b)

(a)

Iteration 1
Iteration 2
Iteration 3
Iteration 4
Iteration 5
Iteration 6
Iteration 7

Dataset = 42 patients

Learning + Validation = 36 patients
(70%)   (30%)

Test = 6 patients

Learning

Validation

Test

Figure

(b)

7 Fold Cross Validation

- Accuracy
- Sensitivity
- Specificity

Model Training

Model Evaluation

Best model choice

Prediction

Learning Set

70%

Split

30%

Validation Set

Features Selection

Training Set

Split

Test Set

Dataset

(a)

(b)

(a)

(b)

(c)

Table 1. Baseline characteristics of study population

	All (n=42)	AS (n=21)	ATTR-CA (n= 21)	P value
Age, years	77 (72-82)	82 (78-84)	74 (67-76)	p<0.01
Female	17 (40%)	10 (48%)	7 (33%)	ns
Body Surface Area, m²	1.9 (1.7-2)	1.8(1.7-1.9)	1.9 (1,7-2)	ns
Body Mass index, kg/m²	26 (23-31)	27 (24-30)	26 (22-32)	ns
LVESVi, ml/m²	19 (15-24)	21 (16-24)	18 (15-24)	ns
LVEDVi, ml/m²	46 (36-59)	57 (45-65)	36 (29-54)	p<0.01
LVEF, %	58 (45-66)	65 (58-66)	48 (43-58)	p<0.01

Values are expressed as absolute number and percentage or median and IQR

LVEDVi: left ventricle end-diastolic volume index; LVEF: left ventricle ejection fraction; LVESVi: left ventricle end-systolic volume index; ns: not significant
LVEF: left ventricle ejection fraction

Table 2. ICC values for the minimal (0.5% translation) and maximal (30% translation) perturbation. In bold, the robust features, i.e., with ICC (0.5% translation) > 0.75 and ICC (30% translation) < 50%. ICC: intraclass correlation.

Feature name	ICC (0.5% translation)	ICC (30% translation)
firstorder 10Percentile	0.9760	0.0404
firstorder 90Percentile	0.9975	0.3870
firstorder Energy	0.9909	0.2208
firstorder Entropy	0.9801	0.0786
firstorder InterquartileRange	0.9882	0.0367
firstorder Kurtosis	0.9768	0.0035
firstorder Maximum	0.8334	0.1303
firstorder MeanAbsoluteDeviation	0.9579	0.0388
firstorder Mean	0.9925	0.0627
firstorder Median	0.9986	0.0000
firstorder Minimum	0.9594	0.1515
firstorder Range	0.9580	0.1365
firstorder RobustMeanAbsoluteDeviation	0.9856	0.0365
firstorder RootMeanSquared	0.9924	0.0586
firstorder Skewness	0.9303	0.0252
firstorder TotalEnergy	0.9909	0.2208
firstorder Uniformity	0.9926	0.0649
firstorder Variance	0.9399	0.0625
glcm Autocorrelation	0.9952	0.2149
glcm ClusterProminence	0.9699	0.0475
glcm ClusterShade	0.9367	0.0522
glcm ClusterTendency	0.8723	0.0608
glcm Contrast	0.9042	0.0658
glcm Correlation	0.9720	0.0040
glcm DifferenceAverage	0.9853	0.1422
glcm DifferenceEntropy	0.9847	0.1761
glcm DifferenceVariance	0.7925	0.0605
glcm Id	0.9967	0.3672
glcm Idm	0.9970	0.3929
glcm Idmn	0.7337	0.0570
glcm Idn	0.7941	0.0929
glcm Imc1	0.9839	0.0225
glcm Imc2	0.9737	0.0012
glcm InverseVariance	0.9925	0.2651

glcm JointAverage	0.9597	0.1835
glcm JointEnergy	0.9970	0.1143
glcm JointEntropy	0.9882	0.1216
glcm MCC	0.8092	0.0037
glcm MaximumProbability	0.9988	0.1952
glcm SumAverage	0.9597	0.1835
glcm SumEntropy	0.9783	0.0652
glcm SumSquares	0.8739	0.0606
glrlm GrayLevelNonUniformity	0.9993	0.2126
glrlm GrayLevelNonUniformityNormalized	0.9891	0.0573
glrlm GrayLevelVariance	0.9272	0.0455
glrlm HighGrayLevelRunEmphasis	0.9953	0.2158
glrlm LongRunEmphasis	0.9991	0.5656
glrlm LongRunHighGrayLevelEmphasis	0.9928	0.1976
glrlm LongRunLowGrayLevelEmphasis	0.7614	0.0901
glrlm LowGrayLevelRunEmphasis	0.7410	0.0915
glrlm RunEntropy	0.9580	0.0462
glrlm RunLengthNonUniformity	0.9998	0.9243
glrlm RunLengthNonUniformityNormalized	0.9984	0.5213
glrlm RunPercentage	0.9988	0.5588
glrlm RunVariance	0.9994	0.4987
glrlm ShortRunEmphasis	0.9984	0.5337
glrlm ShortRunHighGrayLevelEmphasis	0.9957	0.2211
glrlm ShortRunLowGrayLevelEmphasis	0.7373	0.0925
glszm GrayLevelNonUniformity	0.9967	0.3324
glszm GrayLevelNonUniformityNormalized	0.9579	0.0298
glszm GrayLevelVariance	0.7339	0.0467
glszm HighGrayLevelZoneEmphasis	0.9957	0.2193
glszm LargeAreaEmphasis	0.9947	0.1799
glszm LargeAreaHighGrayLevelEmphasis	0.9922	0.0702
glszm LargeAreaLowGrayLevelEmphasis	0.7743	0.0180
glszm LowGrayLevelZoneEmphasis	0.7074	0.1064
glszm SizeZoneNonUniformity	0.9926	0.4355
glszm SizeZoneNonUniformityNormalized	0.9033	0.1671
glszm SmallAreaEmphasis	0.9038	0.1650
glszm SmallAreaHighGrayLevelEmphasis	0.9952	0.2220
glszm SmallAreaLowGrayLevelEmphasis	0.6920	0.1015
glszm ZoneEntropy	0.9445	0.0403
glszm ZonePercentage	0.9888	0.3438
glszm ZoneVariance	0.9947	0.1785
ngtdm Busyness	0.7901	0.0175
ngtdm Coarseness	0.9996	0.5760
ngtdm Complexity	0.7857	0.0960

ngtdm Contrast	0.9202	0.0331
ngtdm Strength	0.8659	0.1475
gldm DependenceEntropy	0.9819	0.0480
gldm DependenceNonUniformity	0.9999	0.8447
gldm DependenceNonUniformityNormalized	0.9993	0.5797
gldm DependenceVariance	0.9995	0.4636
gldm GrayLevelNonUniformity	0.9994	0.2354
gldm GrayLevelVariance	0.9394	0.0622
gldm HighGrayLevelEmphasis	0.9952	0.2152
gldm LargeDependenceEmphasis	0.9993	0.5956
gldm LargeDependenceHighGrayLevelEmphasis	0.9860	0.1717
gldm LargeDependenceLowGrayLevelEmphasis	0.7972	0.1084
gldm LowGrayLevelEmphasis	0.7419	0.0864
gldm SmallDependenceEmphasis	0.9921	0.3534
gldm SmallDependenceHighGrayLevelEmphasis	0.9973	0.2672
gldm SmallDependenceLowGrayLevelEmphasis	0.7102	0.1202

ICC: intraclass correlation coefficient

GLCM: gray level co-occurrence matrix

GLRLM: gray level run length matrix

GLSZM: the gray level size zone matrix

NGTDM: neighbouring gray tone difference matrix

GLDM: gray level dependence matrix

Table 3. Features selected in each of the 7 splits as final input to the ML classifiers using the Wilcoxon test, the LASSO or their combination. In bold, features which have been selected in each split by all the features selection methods employed. LASSO: least absolute shrinkage and selection operator.

Feature name	Wilcoxon	LASSO	Wilcoxon + LASSO
Firstorder 10Percentile	7/7	7/7	7/7
Firstorder Maximum	7/7	7/7	7/7
Firstorder Energy	7/7	3/7	5/7
Glszm SmallAreaEmphasis	7/7	7/7	7/7
Shape MajorAxisLength	7/7	2/7	1/7
Shape Flatness	7/7	4/7	2/7
Shape Elongation	7/7	7/7	7/7
Shape Maximum2DDiameterColumn	7/7	7/7	7/7
Shape Sphericity	7/7	1/7	6/7
Gldm SmallDependenceEmphasis	7/7	0/7	3/7
Shape SurfaceVolumeRatio	2/7	0/7	0/7
Glcm MCC	0/7	1/7	0/7
Ngtdm Contrast	0/7	1/7	0/7
Gldm DependenceEntropy	0/7	4/7	0/7
Glszm LargeAreaHighGrayLevelEmphasis	0/7	1/7	0/7

GLSZM: the gray level size zone matrix

GLDM: gray level dependence matrix

GLCM: gray level co-occurrence matrix

NGTDM: neighbouring gray tone difference matrix

GLDM: gray level dependence matrix