

Article

Papillary-Muscle-Derived Radiomic Features for Hypertrophic Cardiomyopathy versus Hypertensive Heart Disease Classification

Qiming Liu [†] , Qifan Lu [†] , Yezi Chai , Zhengyu Tao, Qizhen Wu, Meng Jiang * and Jun Pu *

Department of Cardiology, RenJi Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai 200120, China; undomiel.lqm@gmail.com or 090503liu@sjtu.edu.cn (Q.L.); czzx_2012_505@126.com (Q.L.); cyz960707@126.com (Y.C.)

* Correspondence: jiangmeng_renji@163.com (M.J.); junpu_1970@163.com (J.P.)

† These authors contributed equally to this work.

Abstract: Purpose: This study aimed to assess the value of radiomic features derived from the myocardium (MYO) and papillary muscle (PM) for left ventricular hypertrophy (LVH) detection and hypertrophic cardiomyopathy (HCM) versus hypertensive heart disease (HHD) differentiation. **Methods:** There were 345 subjects who underwent cardiovascular magnetic resonance (CMR) examinations that were analyzed. After quality control and manual segmentation, the 3D radiomic features were extracted from the MYO and PM. The data were randomly split into training (70%) and testing (30%) datasets. Feature selection was performed on the training dataset. Five machine learning models were evaluated using the MYO, PM, and MYO+PM features in the detection and differentiation tasks. The optimal differentiation model was further evaluated using CMR parameters and combined features. **Results:** Six features were selected for the MYO, PM, and MYO+PM groups. The support vector machine models performed best in both the detection and differentiation tasks. For LVH detection, the highest area under the curve (AUC) was 0.966 in the MYO group. For HCM vs. HHD differentiation, the best AUC was 0.935 in the MYO+PM group. Comparing the radiomics models to the CMR parameter models for the differentiation tasks, the radiomics models achieved significantly improved the performance ($p = 0.002$). **Conclusions:** The radiomics model with the MYO+PM features showed similar performance to the models developed from the MYO features in the detection task, but outperformed the models developed from the MYO or PM features in the differentiation task. In addition, the radiomic models performed better than the CMR parameters' models.

Keywords: radiomics; papillary muscle; hypertrophic cardiomyopathy; hypertensive heart disease; cardiac magnetic resonance



Citation: Liu, Q.; Lu, Q.; Chai, Y.; Tao, Z.; Wu, Q.; Jiang, M.; Pu, J. Papillary-Muscle-Derived Radiomic Features for Hypertrophic Cardiomyopathy versus Hypertensive Heart Disease Classification. *Diagnostics* **2023**, *13*, 1544. <https://doi.org/10.3390/diagnostics13091544>

Academic Editors: Renato Cuocolo, Lorenzo Ugga and Gaia Spadarella

Received: 30 March 2023

Revised: 23 April 2023

Accepted: 24 April 2023

Published: 25 April 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Among the various causes of left ventricular hypertrophy (LVH), hypertrophic cardiomyopathy (HCM) and hypertensive heart disease (HHD) are the most-common. However, the detection of LVH and the differentiation between HCM and HHD are often clinically challenging, especially when the hypertrophy is mild to moderate [1]. Meanwhile, the prognosis of patients with HCM varies. In severe and untreated HCM patients, complications include heart failure or even sudden cardiac death [2,3]. For HHD patients, LVH is typically observed in those with long-term poorly controlled hypertension, and without proper treatment, this can progressively develop into heart failure [4]. More importantly, the treatments for these two diseases are quite different [5]. Therefore, based on accurate LVH detection, the early and precise differential diagnosis between HCM and HHD is crucial for treatment plan decisions and prognosis improvement [2].

Compared to other imaging modalities, cardiac magnetic resonance (CMR) offers a noninvasive and multiparametric approach to evaluating cardiac morphology, function, and tissue characteristics [6–8]. Its high-resolution cine images also benefit the analysis of both the myocardium (MYO) and small myocardial structures, such as the papillary muscle (PM) [9]. Although small, the PM plays an important role in the maintenance of the mitral valve and left ventricle function [10,11]. Till now, several CMR studies have stressed the importance of the PM on LVH-related diseases: Kozor et al. demonstrated disproportionate hypertrophy of the PM in Fabry disease (FD) and HCM. They also found that the PM mass contributes 8%, 7%, 10%, and 13% to the total weight in healthy subjects, hypertensive patients, HCM patients, and FD patients, respectively [12]. Hoigne et al. confirmed that an abnormal PM was the only significant parameter to diagnose non-obstructive HCM [13]. Despite the morphological differences in the PM in various forms of LVH, the PM's characteristics have been proven to provide prognostic prediction value for patients with HCM based on image complexity [14].

Despite the direct visual inspection and parameter calculation of CMR images, radiomics provides a “microscopic” approach to analyze CMR morphology and texture characteristics [15]. Radiomics has been developed and researched for over a decade, and its practical applications have been evaluated in many CMR studies [16,17]. We noticed that most studies focused on the MYO features; only recently have some studies extracted the LV features for disease classification [16,18,19]. However, no radiomics studies have attempted to extract the PM features.

According to our clinical experience and observation of CMR images, the PM showed different morphologies, distribution patterns, and signal intensities between patients with LVH and healthy controls (HCs). In this study, we aimed to determine the value of the PM-derived radiomic features in subjects with LVH and HC. Specifically, we sought to develop LVH detection and HCM vs. HHD differentiation models in addition to the PM-derived features.

2. Materials and Methods

Figure 1 presents the flowchart of the study.

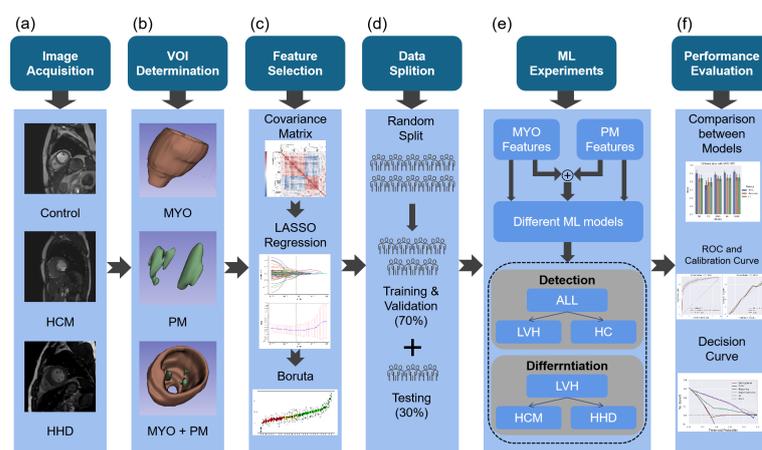


Figure 1. Schematic view of this study. (a) CMR data were collected from HCM, HHD, and HC subjects, and typical mid-ventricle cine images are shown; (b) examples of the volume of interest (VOI) for the MYO, PM, and MYO+PM together in an HC subject (visualization was performed with a 3D slicer); (c) described methods used in the feature selection; (d) shows how the data partition was performed; (e) described ML pipeline: selected MYO and PM features were evaluated for detection and differentiation performance with different ML methods; (f) results were compared between different groups and ML methods, and the evaluation was performed with the ROC curve, calibration curve, and decision curve. HCM, hypertrophic cardiomyopathy; HHD, hypertensive heart disease; VOI, volume of interest; LASSO, least absolute shrinkage and selection operator; ML, machine learning; MYO, myocardium; PM, papillary muscle; LVH, left ventricular hypertrophy.

2.1. Study Population

In this retrospective study, patients with LVH were consecutively enrolled in Renji hospitals from August 2018 to December 2021, and HC subjects were randomly selected from our database.

The HCM group inclusion criteria were as follows: (1) genetic determination of an HCM mutation; (2) LVH (end-diastolic wall thickness > 15 mm) in the absence of a clinical condition known to cause hypertrophy; (3) hypertrophy in a recognizable pattern, i.e., apical-variant HCM (AHCM) [2].

The inclusion criteria for the HHD group were as follows: (1) echocardiogram (ECG) demonstration of a hypertrophic LV (maximal LV wall thickness > 11 mm or LV-mass-to-body-surface area > 115 g/m² for men or > 95 g/m² for women) in the absence of other cardiac or systemic diseases [20]; (2) diagnosis of arterial hypertension [21].

The HC group consisted of healthy volunteers who demonstrated normal cardiac dimensions and volumes, normal cardiac function, and the absence of late gadolinium enhancement. None of the control subjects had a history of known cardiac disease, including cardiac surgery or interventions.

Exclusion criteria for all subjects were an established diagnosis of FD, cardiac amyloidosis, severe valvular disease, aortic stenosis, iron deposition, evidence of inflammatory processes in the myocardium or pericardium, history of ST segment elevation myocardial infarction, and LVH caused by athlete's heart [22].

2.2. Image Acquisition and Analysis

CMR examinations were performed in Renji hospital with a 3T MR scanner (Ingenia, Philips). Cine images were acquired using a balanced steady-state free-precession (bSSFP) sequence [23]. Other typical cine parameters were: slice thickness = 6–8 mm, echo time (TE) = 1.4 ms, repetition time (TR) = 3.0 ms, flip angle = 45°, in-plane resolution = 0.8–1.2 mm × 0.8–1.2 mm, bandwidth = 1900 Hz/pixel, number of cardiac phases = 30. For all participants, the cine stack was covered from the base to the apex.

The CMR data were exported in the Digital Imaging and Communications in Medicine (DICOM) format. Anonymization and CMR function analyses were performed on a workstation with dedicated post-processing software cvi42 (Version 5.13.0, Albert, Canada). All images were subjected to manual quality control to exclude those of inadequate quality before further assessment. Basic CMR parameters that summarize LV function (including LV mass, LV end-diastolic volume (LVEDV), and LV ejection fraction (LVEF)) were obtained using the cvi42.

2.3. Definition of Volume of Interest

Limited by copyright, we were not able to export contours directly from the cvi42 software. To export the MYO and PM contours, an open-source software itk-snap (Version 3.8.0) was used for the following analysis. All images were segmented simultaneously by two cardiologists (>4 years of CMR experience) in consensus. After that, another experienced cardiologist (>5 years of CMR experience) checked the contours and made adjustments if necessary. In this study, only ED cine images were considered for analysis due to ambiguous border between the endocardium and PM in end systole. The VOI of the MYO was defined as the area between the LV endocardium and epicardium [24]. For the VOI of the PM, we delineated the PM contours according to anatomical characteristics, movement pattern, and previous segmentation examples [12,25]. non-PM structures (e.g., papillary muscle variants, apical–basal muscle bundle etc.), and trabeculations were excluded [9,26].

Supplementary Figure S1 shows a representative apical–basal muscle bundle, which is hard to distinguish from the PM with short-axis cine images solely in our datasets.

2.4. Feature Extraction and Feature Selection

Image filters (wavelet, Laplace of Gaussian (LOG), and gradient) were applied to our datasets, and 3D radiomic features were extracted for both the MYO and PM VOIs using Pyradiomics [27]. The detailed description and calculation equation are available on the Pyradiomics official website: <https://pyradiomics.readthedocs.io/en/v3.0.1/radiomics.html>, accessed on 1 October 2022.

After feature extraction, the data were randomly split into training and testing datasets in a 7:3 ratio.

Feature selection was performed on the training dataset with 3 steps: (1) The Pearson correlation coefficient (ρ) was calculated for each feature; features with $\rho > 0.8$ were excluded; (2) a least absolute shrinkage and selection operator (LASSO) regression algorithm was employed to select important features [28]; (3) features were ranked using the Boruta method [29].

2.5. Model Development

Five machine learning classifiers were evaluated: AdaBoost (AB), K-nearest neighbor (KNN), support vector machine (SVM), random forest (RF), and decision tree (DT). A five-fold cross-validation (CV) with a grid search was performed within the training dataset to determine the best combination of model parameters. The performance of the different models was compared based on the area under the curve (AUC), accuracy, precision, recall, and calibration curve.

2.6. Evaluation of MYO and PM Features

For the multi-feature analysis, we aimed to assess the value of the PM features. Therefore, three radiomics groups were designed: (1) the MYO group, including the best $2 \times N$ MYO features; (2) the PM group, including the best $2 \times N$ PM features; (3) the MYO+PM group, including the best N MYO and N PM features. Subsequently, different machine learning (ML) models were tested and evaluated in the three groups.

2.7. Comparison of Radiomic Models and Clinical Data Models

To compare our proposed radiomics models with clinical data models, an optimal ML model was selected and developed using CMR parameters (LVEF, LVEDV index, and LV mass index). We also combined the CMR parameters and radiomic features to examine the possible synergistic effects between CMR parameters and radiomics information. The development and evaluation processes were the same as those previously described.

2.8. Statistics

All statistical analyses were performed using SPSS (Version 26) and Python (Version 3.7). To compare the means, Student's *T*-test, the paired *T*-test, and the Mann–Whitney U-test were conducted as appropriate. Statistical significance was set at $p < 0.05$. Model performance was evaluated using the receiver operating characteristic (ROC) curves and AUC, accuracy, precision, and recall, and a comparison between the AUCs was performed with the Delong test [30]. The Hosmer–Lemeshow test was used for the calibration curve evaluation [31]. Weights were calculated for the HC, HCM, and HHD groups, where appropriate, to compensate the unbalanced datasets [32].

3. Results

3.1. Demographic and CMR-Based Clinical Characteristics

The detailed inclusion and exclusion pipeline is shown in Figure 2.

A total of 230 LVH (NHCM = 158, NHHD = 72) patients who satisfied the inclusion criteria and passed quality control were included. Subjects in the HC group were matched with patients with LVH in terms of age and sex ($p = 0.866$ and 0.364 , respectively) at a 1:2 ratio. Table 1 shows the demographic and CMR data of the cohort.

For the CMR parameters, subjects with LVH showed a significantly increased LV mass (144.3 g vs. 83.2 g, $p < 0.001$) and LV mass index (80.0 g/m² vs. 46.3 g/m², $p < 0.001$). However, for the HHD group, we found that their LVEFs were lower than those in the HCM and HC groups (55% vs. 68% and 65%, respectively, $p < 0.001$ for both groups), whereas the LVEDVs were higher than those in the HCM and HC groups (163 mL vs. 130 mL and 126 mL, respectively, $p < 0.001$ for both groups). This phenomenon could be partly explained by the fact that HHD patients usually have had hypertension for many years; therefore, some patients already had an enlarged left ventricle.

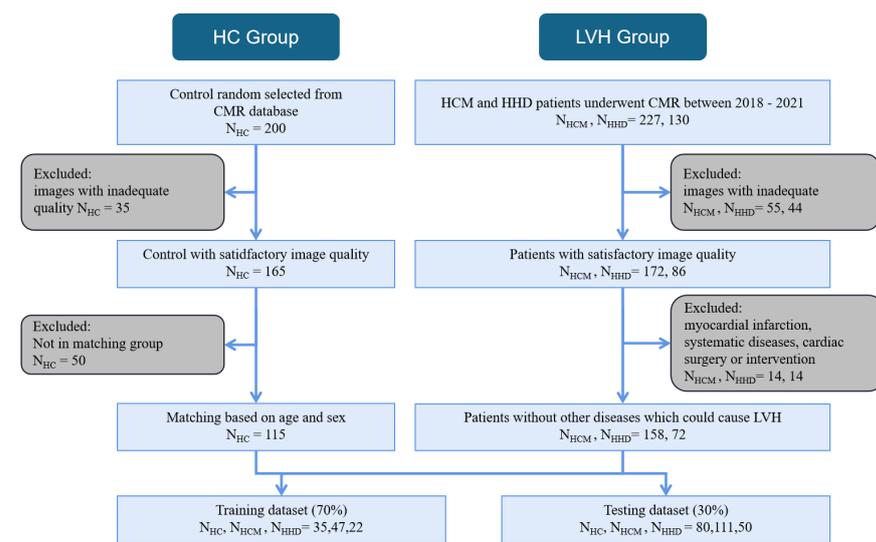


Figure 2. This figure shows the inclusion and exclusion of LVH and HC subjects. All subjects were obtained from our datasets between 2018 and 2021. For the LVH group, firstly, images with inadequate quality were excluded. Secondly, subjects with myocardial infarction, systemic diseases, or cardiac surgery were further excluded. For the HC group, images with satisfactory quality were matched to the LVH group by sex and age at a 1:2 ratio.

Table 1. Demographic and CMR characteristics.

| Clinical Data Entries | Overall (N = 345) | LVH (N = 230) | HCM (N = 158) | HHD (N = 72) | HC (N = 115) | <i>p</i> -Value ¹ |
|--|-------------------|---------------|---------------|--------------|--------------|------------------------------|
| Demographic and Clinical Features | | | | | | |
| Age, year | 51.1 ± 15.7 | 51.4 ± 15.3 | 51.6 ± 15.3 | 50.9 ± 15.4 | 50.6 ± 16.6 | 0.634 |
| Male, n (%) | 245 (71) | 165 (72) | 109 (69) | 56 (78) | 80 (70) | 0.171 |
| Weight, kg | 71.5 ± 14.3 | 72.4 ± 14.0 | 70.2 ± 12.8 | 77.3 ± 15.4 | 69.8 ± 14.7 | 0.001 |
| Height, cm | 168.5 ± 9.0 | 168.3 ± 8.4 | 168.5 ± 8.3 | 170.0 ± 8.5 | 169.1 ± 9.9 | 10.048 |
| BMI, kg/m ² | 25.1 ± 3.9 | 25.4 ± 3.8 | 24.9 ± 3.5 | 26.6 ± 4.2 | 24.3 ± 3.8 | 10.205 |
| BSA, m ² | 1.79 ± 0.22 | 1.80 ± 0.21 | 1.77 ± 0.20 | 1.87 ± 0.23 | 1.77 ± 0.23 | 10.003 |
| CMR Parameters | | | | | | |
| LVEF, % | 64.1 ± 10.6 | 63.5 ± 11.9 | 67.5 ± 7.4 | 54.8 ± 15.0 | 65.4 ± 7.2 | 1 < 0.001 |
| LVEDV, mL | 135.3 ± 40.7 | 140.2 ± 42.3 | 130.1 ± 29.6 | 162.5 ± 55.7 | 125.5 ± 35.5 | 1 < 0.001 |
| LVEDV index, mL/m ² | 75.2 ± 19.0 | 77.7 ± 20.3 | 73.8 ± 15.6 | 86.2 ± 26.2 | 70.2 ± 14.8 | 10.002 |
| LV mass, g | 123.9 ± 53.2 | 144.3 ± 52.1 | 146.4 ± 54.4 | 139.7 ± 46.6 | 83.2 ± 24.0 | 10.515 |
| LV mass index, g/m ² | 68.8 ± 27.9 | 80.0 ± 27.3 | 82.7 ± 29.1 | 74.1 ± 21.9 | 46.3 ± 9.5 | 10.031 |

Values are the mean ± standard deviation or number (percentage). ¹ *p*-values were calculated between the HCM and HHD groups. *p*-values < 0.05 are shown in bold. BMI = body mass index; BSA = body surface area; CMR = cardiac magnetic resonance; HC = healthy control; HCM = hypertrophic cardiomyopathy; HHD = hypertensive heart disease; LV = left ventricular; LVEF = left ventricular ejection fraction.

3.2. Feature Extraction and Selection

A total of 2632 features ($N_{MYO} = 1316$ and $N_{PM} = 1316$) were extracted; the detailed feature information is available in Supplementary Table S1. After correlation and LASSO feature selection, 60 features ($N_{MYO} = 28$, $N_{PM} = 32$) survived; meanwhile, for the detection task, 20 features ($N_{MYO} = 8$, $N_{PM} = 12$) survived, and the features were ranked using relative importance calculated from the Boruta method. Table 2 shows the six most-important features for the differentiation tasks derived from the MYO and PM. The gradient gray-level co-occurrence matrix (GLCM) correlation was the most-important in the MYO group, and the shape maximum 2D diameter slice was the most-important in the PM group. Complete lists of the detection and differentiation tasks are provided in Supplementary Tables S2 and S3.

Table 2. The 6 most-important features selected in the MYO and PM groups for HCM vs. HHD according to the Boruta method.

| MYO (N = 6) | Relative Importance | PM (N = 6) | Relative Importance |
|---------------------------------------|---------------------|--|---------------------|
| gradient GLCM correlation | 37.5 | original shape maximum 2D diameter slice | 10.4 |
| original shape sphericity | 4.9 | log-sigma-5-0-mm-3D first-order kurtosis | 3.6 |
| original shape elongation | 3.5 | original GLSZM ZoneEntropy | 3.2 |
| wavelet-LHL GLCM Imc1 | 3.0 | wavelet-HLL GLCM IMC2 | 3.0 |
| log-sigma-5-0-mm-3D glszm ZoneEntropy | 1.3 | log-sigma-2-0-mm-3D GLCM correlation | 2.9 |
| wavelet-LHH GLCM MCC | 1.2 | gradient GLCM IDMN | 2.4 |

The three most-important features in the MYO and PM groups were selected to construct the MYO+PM group. GLCM = gray-level co-occurrence matrix; GLSZM = gray-level size zone matrix; IDMN = inverse difference moment normalized; IMC = informational measure of correlation; MCC = maximal correlation coefficient; the detailed calculation method for each feature can be found on Pyradiomics' official website.

3.3. Multi-Feature Analysis of MYO, PM, and MYO+PM Groups

For both the detection and differentiation tasks, we used $N = 3$ for further analysis; therefore, we had six features in the MYO, PM, and MYO+PM groups. Among the six features in the MYO+PM group, there were three shape features (MYO: shape sphericity and shape elongation; PM: shape maximum 2D diameter slice).

For the detection task, different model performances on the training dataset are shown in Figure 3. After comparing the AUC, accuracy, and calibration curves (Supplementary Figure S2a), the SVM models were selected. The performance of the SVM models with the testing dataset is presented in Table 3. Among the MYO, PM, and MYO+PM groups, the model developed with the MYO features achieved the highest AUC (0.966 vs. 0.944 (MYO+PM) and 0.772 (PM), $p = 0.924$ and $p < 0.001$, respectively); however, accuracy, precision, and recall were slightly higher in the MYO+PM group. The results indicated the MYO features had a similar efficacy as the MYO+PM features in the detection task.

For the differentiation task, the z-score distribution of the MYO+PM group is shown in Figure 4. The performances of the different models on the training dataset are shown in Figure 3, and the calibration curves are shown in Supplementary Figure S2b. Similar to the detection task, SVM was selected for further analysis. The results of the SVM model showed that the MYO+PM group had a significantly higher AUC (0.935 vs. 0.875 (MYO) and 0.716 (PM), $p = 0.040$ and 0.002 , respectively), and the accuracy showed a 4.4% increase (87.0% vs. 82.6%). The ROC and calibration curves for the differentiation task with the SVM models are plotted in Figure 5, which shows excellent calibration results (all $p > 0.05$,

indicating excellent calibration results) for both the training and testing data. The clinical usefulness of the MYO, PM, and MYO+PM models is shown as decision curves in Figure 6.

Table 3. Performance of selected models on LVH detection and HCM vs. HHD differentiation with MYO, PM, and MYO+PM features.

| Group | Feature Number | AUC ¹ | Accuracy (%) | Precision | Recall | F1 Score |
|---|----------------|---------------------------|--------------|--------------|--------------|--------------|
| LVH detection task (SVM model) | | | | | | |
| MYO | 6 | 0.966 ² | 90.4 | 0.903 | 0.829 | 0.853 |
| PM | 6 | 0.772 ³ | 68.3 | 0.676 | 0.486 | 0.507 |
| MYO+PM | 3 + 3 | 0.964 | 91.3 | 0.913 | 0.829 | 0.866 |
| HCM vs. HHD differentiation task (SVM model) | | | | | | |
| MYO | 6 | 0.875 ⁴ | 82.6 | 0.831 | 0.773 | 0.739 |
| PM | 6 | 0.716 ⁵ | 73.9 | 0.811 | 0.182 | 0.308 |
| MYO+PM | 3 + 3 | 0.935 | 87.0 | 0.871 | 0.818 | 0.800 |

¹ *p*-values were calculated comparing the MYO+PM group using the Delong test. ² *p* = 0.924. ³ *p* < 0.001. ⁴ *p* = 0.040. ⁵ *p* = 0.002.

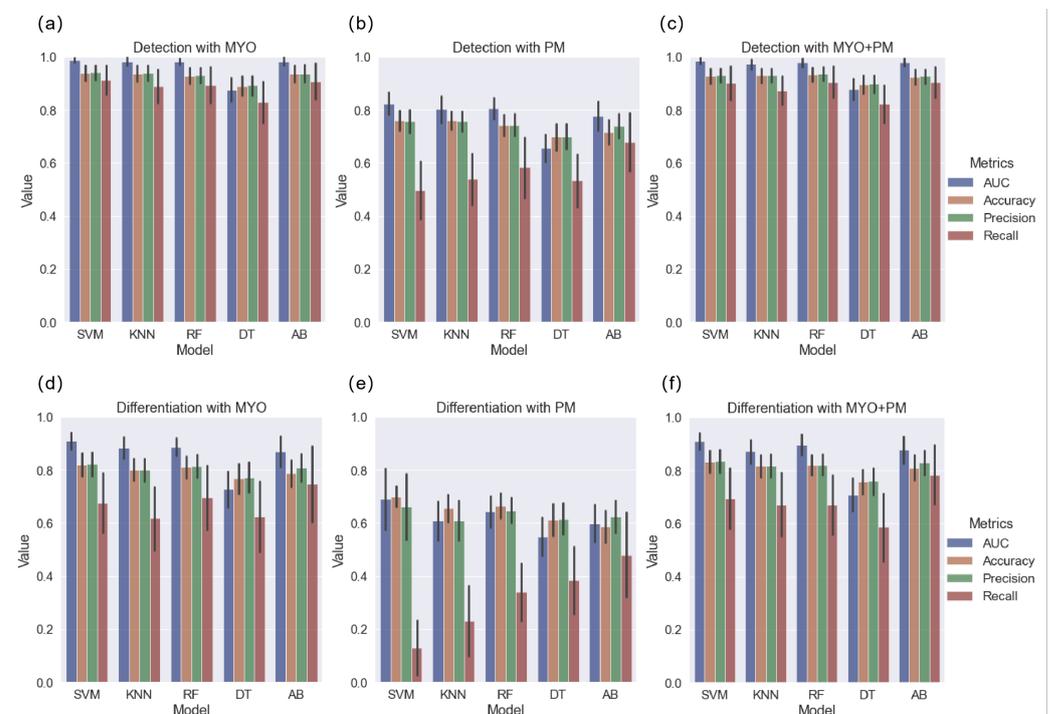


Figure 3. Performance (AUC, accuracy, precision, and recall) of different models on the detection and differentiation tasks with 3 radiomics subgroups; (a–c) detection task with different ML methods using MYO, PM, and MYO+PM groups, respectively; (d–f) differentiation task with different ML methods using MYO, PM, and MYO+PM groups, respectively. SVM, support vector machine; KNN, K-nearest neighbor; RF, random forest; DT, decision tree; AB, AdaBoost.

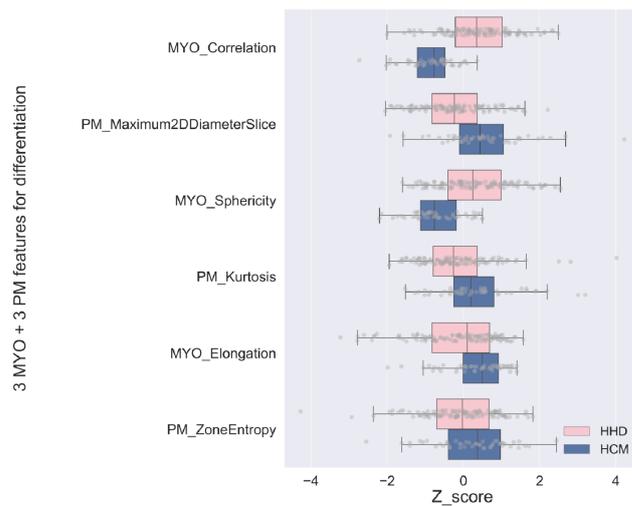


Figure 4. Box plot for the best 3 MYO and 3 PM features in differentiation task. The pink and blue boxes represent the range between the first quartile to the third quartile, while the gray dots represent each single patient in the training dataset. Only features in the MYO+PM group for the differentiation task are exhibited.

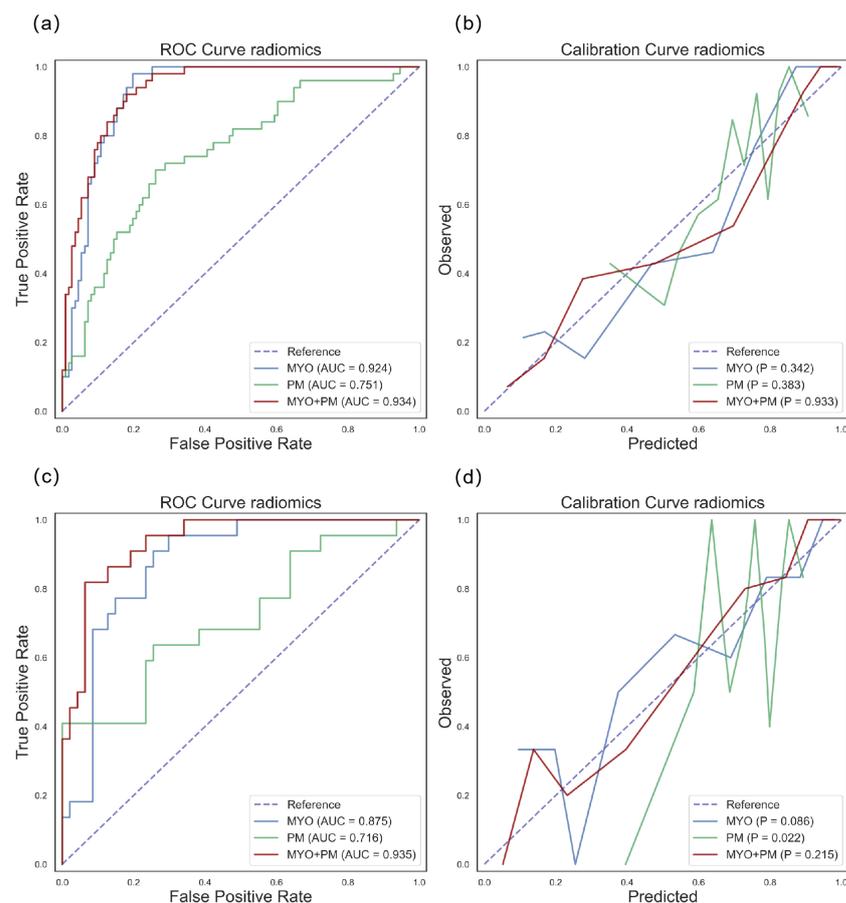


Figure 5. ROC curves and calibration curves for differentiation task with MYO, PM, and MYO+PM groups using SVM models. Blue, green, and red lines represent the MYO, PM, and MYO+PM groups, respectively; the AUC for the ROC curves and Hosmer–Lemeshow *p*-values are denoted in the figure legends. The first row shows the results on the training dataset: (a) ROC curve and (b) calibration curve. The second row shows the results on the testing dataset: (c) ROC curve and (d) calibration curve.

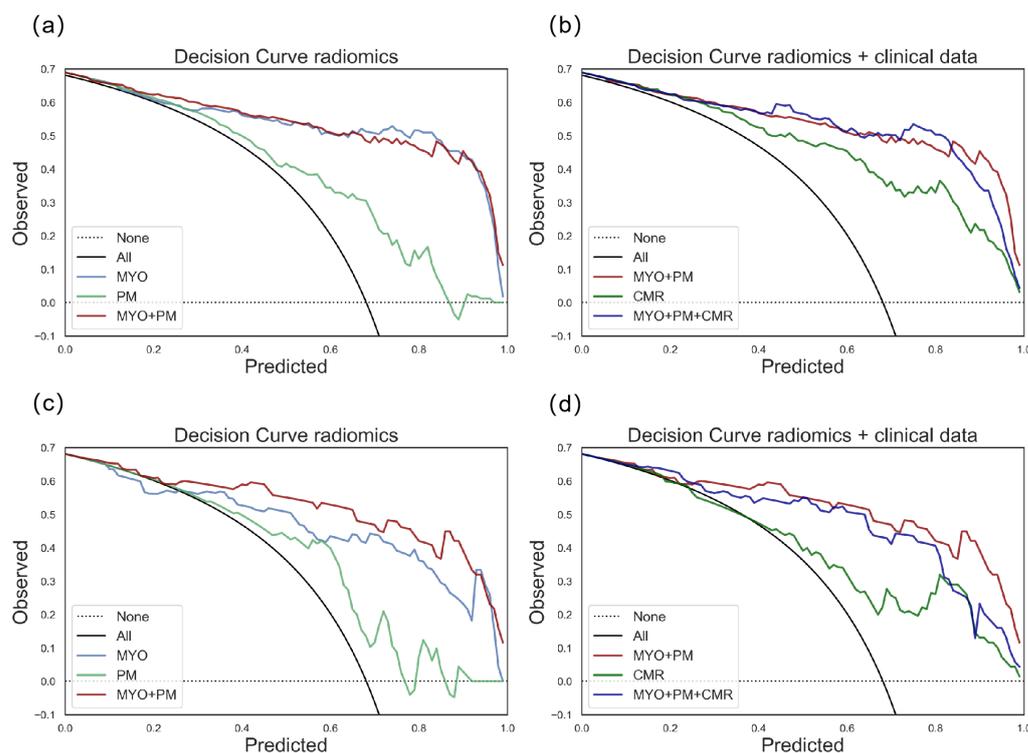


Figure 6. Decision curves for clinical usefulness evaluation. (a) shows the decision curves for the training dataset with the MYO, PM, and MYO+PM groups; (b) shows the decision curves for the training dataset with the radiomics and clinical data groups; (c) shows the decision curves for the testing dataset with the MYO, PM, and MYO+PM groups; (d) shows the decision curves for the testing dataset with the radiomics and clinical data groups.

3.4. Comparison of Radiomics Models to CMR Parameter Models

As shown in Table 4, the radiomics model showed significant improvements compared with the CMR parameter models (AUC: 0.935 vs. 0.774, $p = 0.002$), and no significant improvement was observed when comparing the radiomics + CMR parameter model with the radiomics model (AUC: 0.935 vs. 0.906, $p = 0.117$). We also observed that the CMR parameters group showed biased results (precision = 0.693, recall = 0.409, and F1 = 0.474) in the differentiation task. The calibration curves for the models developed with the CMR parameters, radiomics, and radiomics + CMR parameters are exhibited in Figure 7. The clinical usefulness of the radiomics, CMR, and radiomics+CMR models is shown as the decision curves in Figure 6.

Table 4. SVM model accuracy of clinical-based features (including demographic and CMR features) and radiomics-based features on testing dataset.

| Evaluation Metrics | CMR Parameters (LVEF + LVEDV Index + LV Mass Index) | Radiomics (MYO + PM) | Radiomics + CMR Parameters |
|--------------------|--|-------------------------|-------------------------------|
| AUC ¹ | 0.774 ² | 0.935 | 0.906 ³ |
| Accuracy (%) | 71.0 | 87.0 | 85.5 |
| Precision | 0.693 | 0.871 | 0.860 |
| Recall | 0.409 | 0.818 | 0.818 |
| F1-score | 0.474 | 0.800 | 0.783 |

¹ p -values for the AUC were calculated comparing the radiomics group using the Delong test. ² $p = 0.002$. ³ $p = 0.117$.

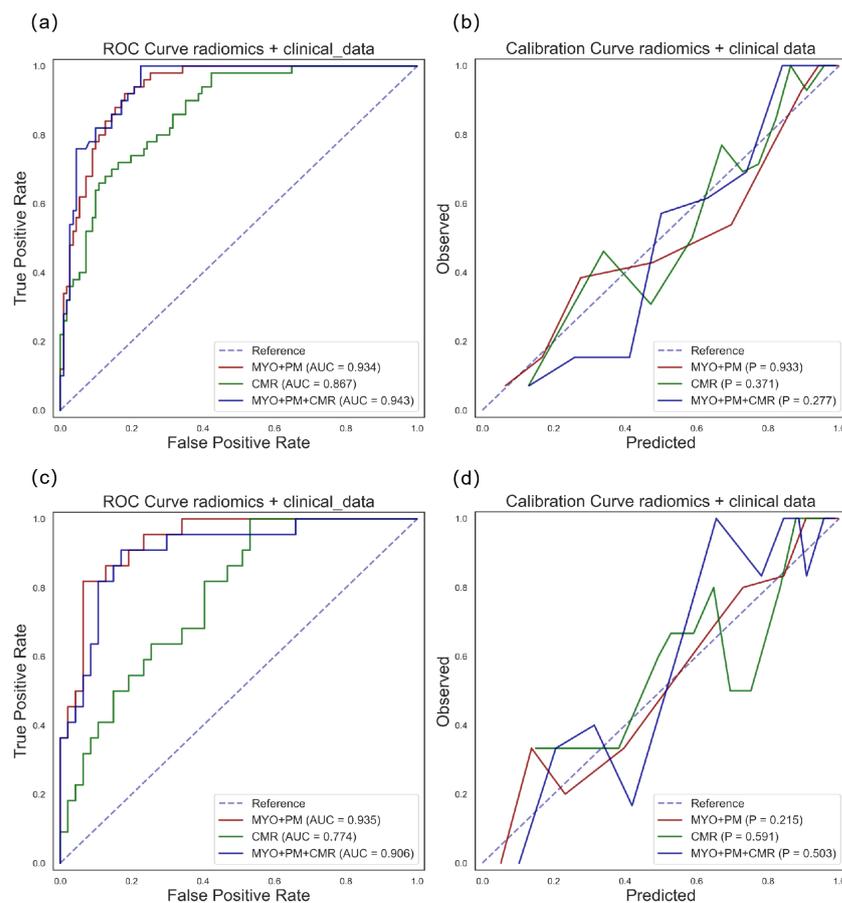


Figure 7. ROC curve and calibration curve for differentiation task with radiomics and clinical data using SVM model. Red, green, and blue lines represent radiomics, CMR parameters, and radiomics + CMR parameters groups, respectively. The AUC for the ROC curves and Hosmer–Lemeshow p-values are denoted in the figure legends. The first row shows the results on the training dataset: ROC curve (a) and calibration curve (b). The second row shows the results on the testing dataset: ROC curve (c) and calibration curve (d).

4. Discussion

In this study, we investigated the LVH detection and HCM vs. HHD differentiation ability of the MYO, PM, and MYO+PM radiomic features; we also compared the MYO+PM group's results with CMR parameters group's results.

4.1. Summary of Main Findings

The main findings of our study were as follows:

- The MYO and MYO+PM groups showed great LVH detection based on the AUC and accuracy;
- The MYO+PM group outperformed the MYO group on the HCM vs. HHD differentiation task;
- Our proposed radiomics models showed significantly better performance than the CMR parameter models.
- Our methods showed excellent calibration results and high clinical usefulness, as shown by the calibration curves and decision curves.

The main findings are summarized in Table 5.

Table 5. Summary table of main findings in this research.

| Task Name | Matrices | MYO | PM | MYO+PM | CMR | MYO+PM+CMR |
|-----------------------------|----------|--------------|-------|--------------|-------|-------------|
| LVH detection | AUC | 0.966 | 0.772 | 0.964 | 0.908 | 0.965 |
| | Accuracy | 90.4 | 68.3 | 91.3 | 81.7 | 91.3 |
| HCM vs. HHD differentiation | AUC | 0.875 | 0.716 | 0.935 | 0.774 | 0.906 |
| | Accuracy | 82.6 | 73.9 | 87.0 | 71.0 | 85.5 |

Bold numbers indicate the highest AUC and accuracy with the selected feature group.

4.2. Discussion Based on Results

Radiomics analysis has been widely applied for the differentiation of cardiomyopathies. However, to our knowledge, this is the first study focusing on the radiomics of the PM. Previous radiomics studies on cardiomyopathy classification were mainly based on the MYO features. Ulf et al. extracted radiomic features from native T1 MYO and selected six texture features; their final model achieved a 0.80 accuracy and 0.89 AUC on the test data [24]. Xu et al. combined deep learning (DL) and radiomics methods for ECG-based LVH aetiology differentiation; their final results showed an AUC of 0.839 for HCM vs. HHD classification [33]. Izquierdo et al. compared their radiomics models to CMR-index-based models and found no significant differences [19]. Although the model performance varied with the datasets, in this study, our models showed satisfactory HCM vs. HHD classification results (accuracy = 87.0% vs. 80.0%, AUC = 0.94 vs. 0.89) compared with previous studies. Using the CMR parameter model as the baseline, our radiomics model showed significant improvement (AUC:0.935 vs. 0.774, $P = 0.002$). Based on these results, we demonstrated the effectiveness and superiority of the proposed radiomics models.

In addition, our models achieved excellent results with limited features (three MYO + three PM features), and the number of features selected was also consistent with several previous studies [24,34]. Among the six features selected, three belonged to the shape feature class (MYO: shape sphericity and shape elongation; PM: shape maximum 2D diameter slice), which made our concise MYO+PM radiomics model more explainable. Radiologically, hypertrophy in HCM patients usually show a concentric pattern, while the hypertrophy of HHD is mainly affected by high blood pressure, which made “MYO sphericity” and “MYO elongation” more reasonable from a clinical perspective. “PM maximum 2D diameter slice” represented the largest distance between the surface mesh vertices in the slice level, which could reflect the different distribution patterns of the PM in HCM and HHD. These results also validated our observation that the PM’s morphology can facilitate the classification between HCM and HHD.

From a clinical point of view, no diagnosis relies purely on radiomic features; therefore, we also developed CMR parameter models and radiomics + CMR combined models. Consistent with a previous study, our CMR-parameter-based model showed unsatisfactory prediction accuracy of only 81.7% and 71.0% on LVH detection and HCM vs. HHD differentiation, respectively [1]. After the incorporation of radiomic features, the performance improved in both tasks, but for HCM vs. HHD, the AUC and accuracy were still worse than the models using solely radiomic features (Table 5).

4.3. Technical Perspectives

Despite these promising findings, we considered the possibility of over-fitting. Over-fitting is a common problem in radiomics studies. Therefore, we implemented calibration curves for different ML models to facilitate both model selection (Supplementary Figure S2) and model performance evaluation (Figures 5 and 7). Although the SVM and KNN methods both showed excellent calibration results, the SVM models showed higher performance than the KNN models in both the detection and differentiation tasks (Figure 4). Therefore, the SVM models were selected for further analysis. The following experimental results validated that the SVM models performed well and did not show an over-fitting tendency with the six radiomic features as the inputs (Figure 5).

From a technical perspective, to deal with unbalanced datasets, class weights were calculated for each group when appropriate. Although we did not compare the performance of our “balanced” dataset with that of an unbalanced dataset, our results showed the following: except for the PM group, the MYO and MYO+PM groups exhibited relatively high precision (all > 0.8) and recall (all > 0.75) on both the training (Figure 3) and testing (Tables 3 and 4) datasets.

4.4. Clinical Perspectives

The implementation of both the MYO and MYO+PM features showed excellent LVH detection performance; however, the MYO+PM group outperformed the MYO group on the HCM vs. HHD differentiation ($p = 0.040$). We also validated that the incorporation of the CMR parameters into the radiomic information did not facilitate HCM vs. HHD differentiation. Our results suggest to clinicians that the PM is a potential useful diagnostic tool for LVH. In addition, we suggest that the calibration curve should be examined for every machine learning algorithm, if appropriate.

4.5. Limitations

This study had some limitations.

First is the immature DL-based PM automatic segmentation results of the PM (DSC = 0.79 in previous studies) [25,35]. In this study, we performed manual segmentation and quality control of selected VOIs. To increase the segmentation accuracy and minimize interobserver variance, two cardiologists completed the segmentation together in consensus. Although time consuming, accurate manual segmentation is crucial for study reliability and repeatability. We also look forward to developing a precise and stable DL segmentation model for further implementation of deep PM radiomics studies.

Second, LVH is not a singular disease entity. The different pathologies of LVH exhibit various clinical behaviors, treatment responses, and prognoses. HCM and HHD are two most-common causes of LVH. Therefore, only patients with HCM or HHD were included in this study. If the sample size permits, we would also like to validate our methods using other available LVH-related disease datasets.

5. Conclusions

To the best of our knowledge, this is the first study on PM-derived radiomic features. Our results showed that the proposed PM radiomics exhibited excellent performance in LVH detection and HCM vs. HHD differentiation and outperformed clinical data models. Moreover, through our experiments, we hypothesized that PM radiomics have great potential for the classification of other cardiac diseases and could facilitate clinicians' clinical decision-making.

Supplementary Materials: The following Supporting Information can be downloaded at: <https://www.mdpi.com/article/10.3390/diagnostics13091544/s1>: Figure S1: An example of apical basal muscle bundle.; Figure S2: Comparison of models using calibration curves.; Table S1: Full list of extracted features.; Table S2: Full list of features ranked using Boruta method for detection task.; Table S3: Full list of features ranked using Boruta method for differentiation task. Refs. [27,36] are cited in the supplementary materials.

Author Contributions: Conceptualization, Q.L. (Qiming Liu) and Q.L. (Qifan Lu); methodology, Q.L. (Qiming Liu); software, Q.L. (Qiming Liu); validation, Q.L. (Qiming Liu), Q.L. (Qifan Lu) and Y.C.; formal analysis, Q.L. (Qiming Liu) and Q.L. (Qifan Lu); investigation, Z.T. and Q.W.; resources, M.J. and J.P.; data curation, Z.T. and Q.W.; writing—original draft preparation, Q.L. (Qiming Liu) and Q.L. (Qifan Lu); writing—review and editing, Q.L. (Qiming Liu); visualization, Q.L. (Qiming Liu) and Q.L. (Qifan Lu); supervision, J.P.; project administration, M.J. and J.P.; funding acquisition, M.J. All authors have read and agreed to the published version of the manuscript.

Funding: This research received funding from National Science Fund for National Natural Science Foundation of China (U21A20341 and 81971570); the Shanghai Academic/Technology Leader Program (21XD1432100); the Shanghai Science and Technology Commission Program (20Y11910500); the Clinical Research Plan of SHDC (SHDC2020CR2025B); Shanghai Jiaotong University (YG2019ZDA13); and the University of Shanghai for Science and Technology (10-20-302-425).

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki and approved by the clinical trial NCT03271385 and Renji Hospital Ethics Committee (20170406J, 6 April 2017).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: No new data were created, and the data are unavailable due to privacy or ethical restrictions.

Acknowledgments: We would like to thank Chongwen Wu in Shanghai Renji Hospital for technical support.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of the data; in the writing of the manuscript; nor in the decision to publish the results.

Abbreviations

The following abbreviations are used in this manuscript:

| | |
|-------|---|
| AB | AdaBoost |
| AUC | area under the curve |
| CV | cross-validation |
| CMR | cardiovascular magnetic resonance |
| DL | deep learning |
| FD | Fabry disease |
| HC | healthy control |
| HCM | hypertrophic cardiomyopathy |
| HHD | hypertensive heart disease |
| LASSO | least absolute shrinkage and selection operator |
| LV | left ventricular/left ventricle |
| LVEDV | left ventricular end diastole volume |
| LVEF | left ventricular ejection fraction |
| LVH | left ventricular hypertrophy |
| ML | machine learning |
| MYO | myocardium |
| KNN | K-nearest neighbor |
| PM | papillary muscle |
| RF | random forest |
| SVM | support vector machine |
| VOI | volume of interest |

References

1. Giusca, S.; Steen, H.; Montenbruck, M.; Patel, A.R.; Pieske, B.; Erley, J.; Kelle, S.; Korosoglou, G. Multi-parametric assessment of left ventricular hypertrophy using late gadolinium enhancement, T1 mapping and strain-encoded cardiovascular magnetic resonance. *J. Cardiovasc. Magn. Reson.* **2021**, *23*, 92. [[CrossRef](#)] [[PubMed](#)]
2. Writing Committee Members; Ommen, S.R.; Mital, S.; Burke, M.A.; Day, S.M.; Deswal, A.; Elliott, P.; Evanovich, L.L.; Hung, J.; Joglar, J.A.; et al. 2020 AHA/ACC Guideline for the Diagnosis and Treatment of Patients with Hypertrophic Cardiomyopathy: Executive Summary: A Report of the American College of Cardiology/American Heart Association Joint Committee on Clinical Practice Guidelines. *Circulation* **2020**, *142*, e533–e557. [[CrossRef](#)]
3. Sen-Chowdhry, S.; Jacoby, D.; Moon, J.C.; McKenna, W.J. Update on hypertrophic cardiomyopathy and a guide to the guidelines. *Nat. Rev. Cardiol.* **2016**, *13*, 651–675. [[CrossRef](#)] [[PubMed](#)]
4. Nwabuo, C.C.; Vasan, R.S. Pathophysiology of Hypertensive Heart Disease: Beyond Left Ventricular Hypertrophy. *Curr. Hypertens. Rep.* **2020**, *22*, 11. [[CrossRef](#)] [[PubMed](#)]

5. Goto, S.; Solanki, D.; John, J.E.; Yagi, R.; Homilius, M.; Ichihara, G.; Katsumata, Y.; Gaggin, H.K.; Itabashi, Y.; MacRae, C.A.; et al. Multinational Federated Learning Approach to Train ECG and Echocardiogram Models for Hypertrophic Cardiomyopathy Detection. *Circulation* **2022**, *146*, 755–769. [[CrossRef](#)]
6. Zhang, N.; Yang, G.; Gao, Z.; Xu, C.; Zhang, Y.; Shi, R.; Keegan, J.; Xu, L.; Zhang, H.; Fan, Z.; et al. Deep Learning for Diagnosis of Chronic Myocardial Infarction on Nonenhanced Cardiac Cine MRI. *Radiology* **2019**, *291*, 606–617. [[CrossRef](#)] [[PubMed](#)]
7. Cardim, N.; Galderisi, M.; Edvardsen, T.; Plein, S.; Popescu, B.A.; D'Andrea, A.; Bruder, O.; Cosyns, B.; Davin, L.; Donal, E.; et al. Role of multimodality cardiac imaging in the management of patients with hypertrophic cardiomyopathy: An expert consensus of the European Association of Cardiovascular Imaging Endorsed by the Saudi Heart Association. *Eur. Hear. J.-Cardiovasc. Imaging* **2015**, *16*, 280–280. [[CrossRef](#)] [[PubMed](#)]
8. Dweck, M.R.; Williams, M.C.; Moss, A.J.; Newby, D.E.; Fayad, Z.A. Computed Tomography and Cardiac Magnetic Resonance in Ischemic Heart Disease. *J. Am. Coll. Cardiol.* **2016**, *68*, 2201–2216. [[CrossRef](#)]
9. Rajiah, P.; Fulton, N.L.; Bolen, M. Magnetic resonance imaging of the papillary muscles of the left ventricle: Normal anatomy, variants, and abnormalities. *Insights Imaging* **2019**, *10*, 83. [[CrossRef](#)]
10. Farhan, S.; Silbiger, J.J.; Halperin, J.L.; Zhang, L.; Dukkipati, S.R.; Vogel, B.; Kini, A.; Sharma, S.; Lerakis, S. Pathophysiology, Echocardiographic Diagnosis, and Treatment of Atrial Functional Mitral Regurgitation. *J. Microbiol. Methods* **2018**, *80*, 2314–2330. [[CrossRef](#)]
11. Teo, E.P.; Teoh, J.G.; Hung, J. Mitral valve and papillary muscle abnormalities in hypertrophic obstructive cardiomyopathy. *Curr. Opin. Cardiol.* **2015**, *30*, 475–482. [[CrossRef](#)] [[PubMed](#)]
12. Kozor, R.; Nordin, S.; Treibel, T.A.; Rosmini, S.; Castelletti, S.; Fontana, M.; Captur, G.; Baig, S.; Steeds, R.P.; Hughes, D.; et al. Insight into hypertrophied hearts: A cardiovascular magnetic resonance study of papillary muscle mass and T1 mapping. *Eur. Heart J. Cardiovasc. Imaging* **2017**, *18*, 1034–1040. [[CrossRef](#)] [[PubMed](#)]
13. Hoigné, P.; Attenhofer Jost, C.; Duru, F.; Oechslin, E.; Seifert, B.; Widmer, U.; Frischknecht, B.; Jenni, R. Simple criteria for differentiation of Fabry disease from amyloid heart disease and other causes of left ventricular hypertrophy. *Int. J. Cardiol.* **2006**, *111*, 413–422. [[CrossRef](#)] [[PubMed](#)]
14. Wang, J.; Li, Y.; Yang, F.; Bravo, L.; Wan, K.; Xu, Y.; Cheng, W.; Sun, J.; Zhu, Y.; Zhu, T.; et al. Fractal Analysis: Prognostic Value of Left Ventricular Trabecular Complexity Cardiovascular MRI in Participants with Hypertrophic Cardiomyopathy. *Radiology* **2021**, *298*, 71–79. [[CrossRef](#)] [[PubMed](#)]
15. Lambin, P.; Leijenaar, R.T.; Deist, T.M.; Peerlings, J.; de Jong, E.E.; van Timmeren, J.; Sanduleanu, S.; Larue, R.T.; Even, A.J.; Jochems, A.; et al. Radiomics: The bridge between medical imaging and personalized medicine. *Nat. Rev. Clin. Oncol.* **2017**, *14*, 749–762. [[CrossRef](#)] [[PubMed](#)]
16. Fahmy, A.S.; Rowin, E.J.; Arafati, A.; Al-Otaibi, T.; Maron, M.S.; Nezafat, R. Radiomics and deep learning for myocardial scar screening in hypertrophic cardiomyopathy. *J. Cardiovasc. Magn. Reson.* **2022**, *24*, 40. [[CrossRef](#)]
17. Raisi-Estabragh, Z.; Jaggi, A.; Gkontra, P.; McCracken, C.; Aung, N.; Munroe, P.B.; Neubauer, S.; Harvey, N.C.; Lekadir, K.; Petersen, S.E. 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. [[CrossRef](#)]
18. Cetin, I.; Sanroma, G.; Petersen, S.E.; Napel, S.; Camara, O.; Ballester, M.A.G.; Lekadir, K. A Radiomics Approach to Computer-Aided Diagnosis with Cardiac Cine-MRI. In *International Workshop on Statistical Atlases and Computational Models of the Heart*; Springer: Cham, Switzerland, 2018; Volume 10663, pp. 82–90. [[CrossRef](#)]
19. Izquierdo, C.; Casas, G.; Martin-Isla, C.; Campello, V.M.; Guala, A.; Gkontra, P.; Rodríguez-Palomares, J.F.; Lekadir, K. Radiomics-Based Classification of Left Ventricular Non-compaction, Hypertrophic Cardiomyopathy, and Dilated Cardiomyopathy in Cardiovascular Magnetic Resonance. *Front. Cardiovasc. Med.* **2021**, *8*, 764312. [[CrossRef](#)]
20. Lang, R.M.; Badano, L.P.; Mor-Avi, V.; Afilalo, J.; Armstrong, A.; Ernande, L.; Flachskampf, F.A.; Foster, E.; Goldstein, S.A.; Kuznetsova, T.; et al. Recommendations for Cardiac Chamber Quantification by Echocardiography in Adults: An Update from the American Society of Echocardiography and the European Association of Cardiovascular Imaging. *J. Am. Soc. Echocardiogr.* **2015**, *28*, 1–39.e14. [[CrossRef](#)]
21. Chobanian, A.V.; Bakris, G.L.; Black, H.R.; Cushman, W.C.; Green, L.A.; Izzo, J.L.; Jones, D.W.; Materson, B.J.; Oparil, S.; Wright, J.T.; et al. Seventh Report of the Joint National Committee on Prevention, Detection, Evaluation, and Treatment of High Blood Pressure. *Hypertension* **2003**, *42*, 1206–1252. [[CrossRef](#)]
22. Prakken, N.H.; Velthuis, B.K.; Teske, A.J.; Mosterd, A.; Mali, W.P.; Cramer, M.J. Cardiac MRI reference values for athletes and nonathletes corrected for body surface area, training hours/week and sex. *Eur. J. Prev. Cardiol.* **2010**, *17*, 198–203. [[CrossRef](#)]
23. Bieri, O.; Scheffler, K. Fundamentals of balanced steady state free precession MRI. *J. Magn. Reson. Imaging* **2013**, *38*, 2–11. [[CrossRef](#)] [[PubMed](#)]
24. Neisius, U.; El-Rewaidy, H.; Nakamori, S.; Rodriguez, J.; Manning, W.J.; Nezafat, R. Radiomic Analysis of Myocardial Native T1 Imaging Discriminates Between Hypertensive Heart Disease and Hypertrophic Cardiomyopathy. *JACC Cardiovasc. Imaging* **2019**, *12*, 1946–1954. [[CrossRef](#)] [[PubMed](#)]
25. Bartoli, A.; Fournel, J.; Bentatou, Z.; Habib, G.; Lalande, A.; Bernard, M.; Boussel, L.; Pontana, F.; Dacher, J.N.; Ghattas, B.; et al. Deep Learning-based Automated Segmentation of Left Ventricular Trabeculations and Myocardium on Cardiac MR Images: A Feasibility Study. *Radiol. Artif. Intell.* **2021**, *3*, e200021. [[CrossRef](#)] [[PubMed](#)]

26. Gruner, C.; Chan, R.H.; Crean, A.; Rakowski, H.; Rowin, E.J.; Care, M.; Deva, D.; Williams, L.; Appelbaum, E.; Gibson, C.M.; et al. Significance of left ventricular apical–Basal muscle bundle identified by cardiovascular magnetic resonance imaging in patients with hypertrophic cardiomyopathy. *Eur. Hear. J.* **2014**, *35*, 2706–2713. [[CrossRef](#)] [[PubMed](#)]
27. van Griethuysen, J.J.; Fedorov, A.; Parmar, C.; Hosny, A.; Aucoin, N.; Narayan, V.; Beets-Tan, R.G.; Fillion-Robin, J.C.; Pieper, S.; Aerts, H.J. Computational Radiomics System to Decode the Radiographic Phenotype. *Cancer Res.* **2017**, *77*, e104–e107. [[CrossRef](#)]
28. Tibshirani, R. Regression Shrinkage and Selection via The Lasso: A Retrospective. *J. R. Stat. Soc. Ser. Stat. Methodol.* **2011**, *73*, 273–282. [[CrossRef](#)]
29. Kursu, M.B.; Rudnicki, W.R. Feature Selection with the Boruta Package. *J. Stat. Softw.* **2010**, *36*, 2–4. [[CrossRef](#)]
30. DeLong, E.R.; DeLong, D.M.; Clarke-Pearson, D.L. Comparing the areas under two or more correlated receiver operating characteristic curves: A nonparametric approach. *Biometrics* **1988**, *44*, 837–845. [[CrossRef](#)]
31. Paul, P.; Pennell, M.L.; Lemeshow, S. Standardizing the power of the Hosmer–Lemeshow goodness of fit test in large data sets. *Stat. Med.* **2013**, *32*, 67–80. [[CrossRef](#)]
32. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; et al. Scikit-learn: Machine Learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825–2830.
33. Xu, Z.; Yu, F.; Zhang, B.; Zhang, Q. Intelligent diagnosis of left ventricular hypertrophy using transthoracic echocardiography videos. *Comput. Methods Programs Biomed.* **2022**, *226*, 107182. [[CrossRef](#)] [[PubMed](#)]
34. Zhou, K.; Shang, J.; Guo, Y.; Ma, S.; Lv, B.; Zhao, N.; Liu, H.; Zhang, J.; Xv, L.; Wang, Y.; et al. Incremental diagnostic value of radiomics signature of pericoronary adipose tissue for detecting functional myocardial ischemia: A multicenter study. *Eur. Radiol.* **2023**, *33*, 3007–3019. [[CrossRef](#)] [[PubMed](#)]
35. Fournel, J.; Bartoli, A.; Bendahan, D.; Guye, M.; Bernard, M.; Rauseo, E.; Khanji, M.Y.; Petersen, S.E.; Jacquier, A.; Ghattas, B. Medical image segmentation automatic quality control: A multi-dimensional approach. *J. Microbiol. Methods* **2018**, *74*, 102213. [[CrossRef](#)]
36. Zwanenburg, A.; Vallieres, M.; Abdalah, M.A.; Aerts, H.J.W.L.; Andrearczyk, V.; Apte, A.; Ashrafinia, S.; Bakas, S.; Beukinga, R.J.; Boellaard, R.; et al. The Image Biomarker Standardization Initiative: Standardized Quantitative Radiomics for High-Throughput Image-based Phenotyping. *Radiology* **2020**, *295*, 328–338. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.