Radiomic analysis of native T1 mapping images discriminates between MYH7 and MYBPC3-related hypertrophic cardiomyopathy

Wang, Jie; Yang, Fuyao; Liu, Wentao; Sun, Jiayu; Han, Yuchi; Li, Dong; Gkoutos, Georgios V; Zhu, Yanjie; Chen, Yucheng

DOI:
10.1002/jmri.27209

Radiomic analysis of native T1 mapping images discriminates between MYH7 and MYBPC3-related hypertrophic cardiomyopathy

Jie Wang, MD,[1,2†] Fuyao Yang, BMSci[1†], Wentao Liu, BMSci [3], Jiayu Sun, MD[4], Yuchi Han, MD, MMSc[5], Dong Li, PhD[6], Georgios V Gkoutos, PhD,[2,7] Yanjie Zhu, PhD[*8], Yucheng Chen, MD,[*1,4,9]

†Both authors contributed equally to this work

[1] Department of Cardiology, West China Hospital, Sichuan University, Chengdu, Sichuan, 610041, China

[2] College of Medical and Dental Sciences, Institute of Cancer and Genomic Sciences, University of Birmingham, B15 2TT, Birmingham, United Kingdom

[3] Medical Big Data Center, West China Hospital, Sichuan University, Chengdu, 610041, Sichuan Province, P. R. China

[4] Department of Radiology, West China Hospital, Sichuan University, Chengdu, 610041, Sichuan Province, P. R. China

[5] Department of Medicine (Cardiovascular Division), University of Pennsylvania, Philadelphia, Pennsylvania, USA

[6] Division of Hospital Medicine, Emory University School of Medicine, 101 Woodruff Cir, Atlanta, GA 30322, USA

[7] MRC Health Data Research UK (HDR UK).

[8] Paul C. Lauterbur Research Centre for Biomedical Imaging, Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences

[9] Center of Rare diseases, West China Hospital, Sichuan University, Chengdu, 610041, Sichuan Province, P. R. China

*Corresponding author:

Yucheng Chen, MD, FACC, FSCMR; Professor, Department of Cardiology, West China Hospital, Sichuan University, Guoxue Xiang No. 37, Chengdu, Sichuan, 610041, China; Tel: 86-18980602149; Fax: 86-028-85422836; E-mail: chenyucheng2003@126.com and Yanjie Zhu, PhD; Paul C. Lauterbur Research Centre for Biomedical Imaging, Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, 1068 Xueyuan Ave., Shenzhen university town, Nanshan, Shenzhen, Guangdong, China. Tel: (86)755-86392248; Email: yj.zhu@siat.ac.cn.

Grant Support:

This work was supported by a grant from 1•3•5 project for disciplines of excellence, West China Hospital, Sichuan University (ZYJC18013) and China Scholarship Council.

Conflicts of interest

All authors declare that they have no competing interests regarding this manuscript.

Running Title: Phenotypic differences via radiomics analysis in HCM

Radiomic analysis of native T1 mapping images discriminates between MYH7 and MYBPC3-related hypertrophic cardiomyopathy

ABSTRACT

Background: The phenotype via conventional cardiac MRI analysis of *MYH7* (β-myosin heavy chain) - and *MYBPC3* (β-myosin-binding protein C)-associated hypertrophic cardiomyopathy (HCM) groups is similar. Little study exists on the genotypic-phenotypic association as assessed by the Machine Learning in HCM patients.

Purpose: To explore phenotypic differences based on radiomics analysis of T1 mapping images between *MYH7* and *MYBPC3*-associated HCM subgroups.

Study Type: Prospective observational study

Subjects: 102 HCM patients with pathogenic, or likely pathogenic mutation, in *MYH7* (n=68) or *MYBPC3* (n=34) genes.

Field Strength/Sequence： Cardiac MRI was performed at 3.0 T with balanced steady-state free precession (bSSFP), phase-sensitive inversion recovery (PSIR) late gadolinium enhancement (LGE), and modified Look–Locker inversion recovery (MOLLI) T1 mapping sequences.

Assessment: All patients underwent next generation sequencing and Sanger genetic sequencing. Left ventricular native T1 and LGE were analyzed. 157 radiomic features

were extracted and modeled using a Support Vector Machine (SVM) combined with principal component analysis (PCA). Each subgroup was randomly split 4:1 (feature selection/test validation).

Statistical Tests: Mann-Whitney U tests and Student's t-tests were performed to assess differences between sub-groups. A receiver operating characteristic (ROC) curve was used to assess the model's ability to stratify patients based on radiomic features.

Results: There were no significant differences between *MYH7*- and *MYBPC3*-associated HCM subgroups based on traditional native T1 values (global, basal and middle short-axis slice native T1; P=0.760, 0.914 and 0.178, respectively). However, the SVM model combined with PCA achieved an accuracy and area under the curve (AUC) of 92.0% and 0.968 (95% confidence interval [CI]:0.968-0.971), respectively. For the test validation dataset, the accuracy and AUC were 85.5% and 0.886 (95%CI: 0.881-0.901), respectively.

Data Conclusion: Radiomic analysis of native T1 mapping images may be able to discriminate between *MYH7*- and *MYBPC3*-associated HCM patients exceeding the performance of conventional native T1 values.

Key words: Magnetic Resonance Imaging; Cardiomyopathy, Hypertrophic; Machine Learning; Support Vector Machine; Human Genetics.

INTRODUCTION

Hypertrophic cardiomyopathy (HCM) is the most common primary cardiomyopathy which manifests as unexplained left ventricular (LV) myocardial hypertrophy. The prevalence of HCM is 1:500 in the general adult population [1]. It is an autosomal dominant genetic disease with 40-60% patients having mutations of sarcomere-related genes. Of those, *MYH7* (β-myosin heavy chain) and *MYBPC3* (β-myosin-binding protein C) are the most commonly affected genes [2, 3].

The relationship between genotype and phenotype amongst HCM patients is clinically important. In previous studies, researchers aimed to identify different genotypes by looking for specific imaging markers in HCM patients [2, 3]. Specific association of genotype and imaging-phenotypes could then be established for advanced risk stratification strategies. In addition, genetic testing was able to provide additional information on risk stratification and clinical management within HCM patients carrying different mutated genes. However, the relationship of genotype and phenotype in *MYH7* and *MYBPC3* gene mutations is unclear. Some studies have reported phenotypic differences between patients with *MYH7 and MYBPC3* gene mutations [4, 5], whereas others have failed to establish such associations [6-9].

Cardiac magnetic resonance imaging (MRI) offers accurate morphological, functional, and myocardial tissue characterization by late gadolinium enhancement (LGE) and T1 mapping in HCM [10]. LGE characteristics have been demonstrated to differentiate different etiologies of left ventricular hypertrophy [11] and T1 mapping potentially offers further discrimination criteria [12]. However, little data exists on exploring

genotype-phenotype associations based on cardiac MRI assessment of patients presenting with HCM phenotypic characteristics. A multicenter registry [2] involving 5 tertiary HCM centers in North America and Europe reported a lack of phenotypic differences between *MYH7*- and *MYBPC3*- associated HCM when assessed by cardiac MRI. Another, small sample sized, study by Ellims et al reported that, in addition to LGE, T1 mapping did not identify different phenotypic manifestations between patients with *MYBPC3* and *MYH7* genes[7]. However, traditional analysis of cardiac MRI images, such as LGE and T1 mapping, is mainly based on global or mean image signal intensity levels and ignores subtle changes in signal inhomogeneity and variable distribution of abnormalities. Radiomics analysis, which is based on quantitative image analyses, and typically relies on biomedical imaging containing information reflecting disease-specific processes, has the potential to identify and extract such information [13]. In addition, cardiac MRI-based texture analyses have been previously demonstrated to provide useful discrimination criteria between HCM and hypertension patients and to provide incremental prognostic value in HCM patients [14, 15]. Our hypothesis was that radiomics and machine learning could better discriminate different phenotypic characteristics related to different genotypes. In this study, we aimed to explore the potential relationship between myocardial tissue characteristics by T1 mapping and two most common genotypes, the *MYH7* mutation and the *MYBPC3* mutation.

MATERIALS AND METHODS

Study population

The study population was prospectively recruited from our clinical cardiac MRI registry and consisted of 372 HCM patients, who were referred for cardiac MRI on a 3.0-T scanner between 2011-2017[16]. Only patients with a single pathogenic mutation in either the *MYH7* or *MYBPC3* gene were recruited for the study. The HCM diagnoses were in accordance with the latest guidelines and based on the presence of LV wall thickness ≥15mm in one or more myocardial segments (or ≥13 mm in a first degree relative of an index patient with HCM) measured by any imaging modality (including echocardiography and cardiac MRI) in the absence of secondary causes of hypertrophy [17]. All clinical and baseline data including demographic characteristics, symptoms of heart failure, brain natriuretic peptide (BNP), arrhythmias, and medications were assessed. This study was approved by West China Hospital Institutional Ethics Committee, Sichuan University and written informed consent was obtained from each participant.

Genetic testing

Next generation sequencing (NGS) genetic testing based on a panel including 117 cardiomyopathy-related genes [18] and *TTR* gene sequencing were performed. Methods for genetic sequencing and *in silico* analysis were consistent with previous report [19]. Exon-enriched DNA was sequenced using a Hiseq2000 Sequencing System (Illumina, San Diego, CA, USA). After sequencing, the raw data were saved in a FASTQ format. Illumina sequencing adapters and low quality reads (<80bp) were filtered by cutadapt[20]. After quality control, the clean reads were mapped to the University of

California Santa Cruz (UCSC) hg19 human reference genome using Burrows-Wheeler Alignment. Duplicated reads were removed using the picard tools and mapping reads were used for variations detection. Variations that included single-nucleotide variants (SNVs) and small insertions or deletions (indels) were identified using both the VarScan 2.2.7 software package [21] as well as the variant quality score recalibration protocol in the Genome Analysis Toolkit [22]. These were further filtered using recommended threshold values (mapping quality > 30, base quality >15, and read numbers > 3) [19, 22, 23]. Then, SNVs available within dbSNP130 (hg19) as well as those reported by the 1000 Genomes Projectm, were filtered out from the output files using the ANNOVAR software tool [24].

Variant calling was performed separately for each individual sample. The functional effect of non-synonymous SNVs was assessed by PolyPhen-2, SIFT, and MutationTaster [25-27]. Non-synonymous SNVs with a SIFT score of <0.05, Polyphen-2 score of >0.85 or MutationTaster score of >0.85 were considered to be significant. To sort potentially deleterious variants from benign polymorphisms, perl scripts were used to filter the SNVs against those of dbSNP135 (hg19). Any SNV recorded in dbSNP135 with a reported minor allele frequency of ≥1% in Chinese populations within the 1000 genome database were considered benign polymorphisms and therefore removed from subsequent analysis. Moreover, pathogenicity determination of gene mutation was carried out based on the American College of Medical Genetics and Genomics recommendations (ACMG) [28]. A patient was considered as genotype positive when he or she carried mutations classified as likely

pathogenic (class IV) or pathogenic (class V). Variants were also considered pathogenic if published as causative HCM mutations in at least two independent peer-reviewed studies. Genetic results were subsequently confirmed by Sanger sequencing.

Cardiac MRI

ECG gated cardiac MRI was performed on a 3.0-T scanner (MAGNETOM Skyra, Siemens Healthcare Ltd., Erlangen, Germany) with a 30-channel phased-array receiver coil. Balanced Steady-state free precession (bSSFP) cine images of the entire LV from the base to the apex in consecutive short-axis views were acquired during breath-holds, with the following parameters: repetition time, 3.4ms; echo time, 1.3ms; temporal resolution, 42 ms; flip angle (FA), 50 degrees; field of view, 320-340mm; matrix size, 256×144. The reconstructed in plane spatial resolution was 1.4mm*1.3mm and the slice thickness was 8mm with no gap. Native T1 mapping was performed before injection of gadolinium by a motion-corrected (MOCO) modified Look-Locker inversion recovery (MOLLI) sequence with a 5b (3b) 3b acquisition scheme. Parameters for MOLLI were as follows: TR 740ms, TE 1.06ms, FA 35°, bandwidth 930 Hz/pixel, initial TI 100ms, with 80ms increments, parallel imaging factor 2, matrix size 256*144, in-plane spatial resolution of 2.4*1.8mm, and a total acquisition time of 11 heart beats. LGE images were acquired 10–15 minutes after intravenous administration of 0.15 mmol/kg of gadopentetate dimeglumine (Gd-DTPA, Magvist, Bayer Healthcare), using a phase-sensitive inversion recovery (PSIR) sequence with short axis views (TR, 700 ms; TE, 1.56 ms; FA, 20°; matrix

size, 256 × 144). The inversion time (TI) was individually optimized to null normal myocardial signal using a TI scout sequence.

Cardiac MRI image post-procession analyses

All functional analysis and T1 measurements were performed by experienced radiologists (perform over 2000 cases per year) using commercially available software (Qmass 8.1; Medis Medical Imaging Systems, Leiden, the Netherlands). The tracing method of endo- and epi-cardial borders for ventricular function and mass was in accordance with a previous study [29]. The maximal wall thickness (MWT) was defined as the greatest dimension anywhere within the LV myocardium. For myocardial T1 calculation, the endocardial and epicardial contours were manually traced on the pre-contrast T1-MOCO images, using the Qmass 8.1 software. The average myocardial T1 was obtained for each imaging slice and global native T1 was calculated as the average of basal, middle, and apical slices. LGE was defined automatically as a myocardial signal intensity of 6 standard deviations (SD) from the normal myocardium. The total extent of LGE was calculated by summing LGE percent in each slice as a proportion of total LV myocardium (%LGE).

Feature extraction and selection of T1 images

A total of 157 quantitative features [30-34] were extracted from the T1 MOCO images from 5 sets of distinct texture descriptors (Table 1). The process of image filtering and feature extraction was performed with MATLAB 2014 (Mathworks, Natick, MA, USA), which is consistent with previous studies [14]. This task was performed blinded

to the genetic information associated with each patient. To increase the ratio between the number of training samples and the number of features, a principal component analysis (PCA) was applied to the training dataset before training the SVM classifier. Each subgroup was randomly split 4:1 (feature selection/test validation). The SVM classifier was trained on the feature selection datasets (80%) and tested in the test validation datasets (20%).

Statistical analysis

Statistical analysis was performed using MedCalc (version 13.0; Ostend, Belgium) software. The Shapiro-Wilk test was used to test normality of variables. Unpaired Student's T-tests or Mann Whitney U tests were performed, depending on whether the data was normally distributed, to compare variables between subgroups. Continuous variables were expressed as mean ± SD and median (interquartile range (IQR)). Categorical variables were expressed as N (%). Model accuracy, sensitivity and specificity were calculated using probabilities derived from the linear SVM classifier. A receiver operating characteristic (ROC) curve was used to assess the performance of the model using radiomic features. $P < 0.05$ was considered statistically significant.

RESULTS

Baseline data

The baseline clinical and cardiac MRI characteristics of the *MYH7* and *MYBPC3* subgroups are shown in Table 1. There were no significant differences in baseline

clinical characteristics [age, sex, body mass index, body surface area, blood pressure, symptoms of heart failure, BNP, arrhythmias, medications, cardiac function, MWT and LV mass index] between the two subgroups (all P > 0.05, Table 2).

Conventional analysis of T1 mapping and LGE

There was no significant difference in LGE burden, when quantified as %LGE (10.74±0.08 % versus 11.08±9.11 %; P=0.797) in HCM patients with *MYH7* gene mutation compared to patients with *MYBPC3* gene mutation. Moreover, mean native T1 values showed no significant differences either on the global or slice level in HCM patients with *MYH7* gene mutation compared to patients with *MYBPC3* gene mutation. (Global T1: 1275.38 ms (1236.11, 1305.95)versus 1283.42 ms (1260.39, 1294.69), p=0.760; Basal T1: 1274.50 ms (1227.28, 1318.80) versus 1272.00 ms (1214.07, 1331.45), p=0.914; Middle T1: 1260.87 ms (1223.05, 1292.33) versus 1267.03 ms (1243.77, 1322.08), p=0.178; Table 1).

Features extracted and performance of selected features in identifying HCM with mutation of MYH7 from mutation of MYBPC3

157 radiomic features were extracted from 102 subjects (Table 1). Table 3 shows the prediction accuracy of the model developed to discriminate *MYH7* and *MYBPC3* subgroups. In the feature selection datasets, the SVM plus PCA model achieved an accuracy of 92% for global native T1, 92.3% for basal slice native T1, and an accuracy of 86.9% for middle slice native T1 for radiomics features. ROC analysis using radiomic features of T1 images demonstrated a differential performance with AUC of 0.986 [95% CI 0.968-0.971] and sensitivity of 82.3% and specificity of 97.2%

(Table 3, Figure 1). The model's ability to discriminate between the two subgroups based on radiomics features of slice-based T1 [AUC: 0.969(basal slice), 0.940(middle slice)],is reported in Table 3. In the test validation dataset, ROC analysis using radiomics features of T1 images had an AUC of 0.886 [95% CI 0.881-0.901] and a sensitivity of 75% and specificity of 91.1% (Table 3, Figure 1).When radiomics features were used for discriminating between the two subgroups in the basal and middle slice alone, the AUC was 0.887 and 0.849 respectively (Table 3).

According to Gain calculate the contribution for each parameter in the prediction model [35], image features of local binary pattern (LBP)-7, LBP-37 and run percentage (RP).2 contributed most to the model.

DISCUSSION

Main findings

The present study demonstrated the ability of phenotype-based differentiation of the two most common genotypes of *MYH7* and *MYBPC3* HCM patients by using radiomics analysis of T1 images. It also further confirmed that there were no significant phenotypic differences between *MYH7*- and *MYBPC3*-associated HCM groups when using conventional cardiac MRI analysis, such as cardiac function, LGE and native T1. The extracted image features from radiomics analysis of T1 mapping images could provide robust discriminating power between *MYH7* and *MYBPC3* subgroups.

Previous studies exploring the association between genotype and phenotype based on conventional clinical characteristics in HCM patients

Wang *et al.* (n=70) and Song *et al*. (n=68) identified phenotypic differences in HCM patients with *MYH7* mutation compared to HCM patients carrying *MYBPC3* mutation in Chinese population cohorts [4, 5]. A meta-analysis also reported HCM patients with *MYH7* mutation had higher cardiac conduction disease, ventricular arrhythmia and heart transplantation rate in comparison to those with MYBPC3 mutation ($p < 0.05$) [36].In contrast, Van Driest *et al.* reported that HCM patients with *MYBPC3* mutations (n=63) didn't differ from those patients with a single *MYH7* or light chains gene mutation (n = 61) with respect to age at diagnosis and MWT (22.5 ± 5 mm vs. 23.5 ± 7 mm) [6].

Previous studies exploring the association between genotype and cardiac MRI-phenotype based on conventional cardiac MRI analysis in HCM patients

Currently, there is a limited number of studies comparing differences in the burden of myocardial fibrosis by cardiac MRI between patients carrying *MYBPC3* and *MYH7* gene mutations. Ellims *et al.* found HCM patients with *MYH7* (n=11) or *MYBPC3* (n=17) mutations had similar LGE burden and T1 times [7]. Adaya *et al.* also reported that there were no significant differences in prevalence or degree of myocardial fibrosis quantified by LGE between the two genotypes [2]. . In addition, according to Miller et al, there was no significant difference in LGE burden in patients with *MYH7* compared to *MYBPC3* variants [37]. Our global and slice models further confirmed, via native T1 assessment, that there are no significant differences in fibrosis between

patients carrying these variants. Previous studies reported that increased LGE burden was associated with an increase in sustained ventricular tachycardia or cardioverter-defibrillator shock. However, data from the sarcomeric human cardiomyopathy (SHARE) registry [38] indicated that the incidence of cardiac arrest was higher in patients with *MYBPC3* variants. Furthermore, Ho et al. [38] also reported that patients with *MYBPC3* variants demonstrated a trend towards an increase in arrhythmic events even though no significant difference in LGE burden between the two genotypes was found. Our findings may provide potential insights for these studies. Native T1 is a surrogate marker of fibrosis and our study characterized disease-specific differences in fibrosis patterns by radiomic analysis of T1 maps. It was further suggested the fabricate substrate of *MYH7* may be different from that of *MYBPC3*, further resulting in different incidences of cardiac arrest between patients with MYH7 and *MYBPC3* variants. As a result, new studies on exploring mechanism of cardiac fibrosis between two genotypes should be further explored.

The advantages and clinical value of study based on radiomics analysis from T1 maps to explore the association between genotype and phenotype in HCM patients

Our results have identified that the highest contributing radiomic features were LBP-7, LBP-37 and RP.2. The former (LBP-7 and LBP-37) represent a rotation invariant image descriptor computed from discrete Fourier transforms of LBP histograms. The latter represents run-length matrix of RP histograms. It showed that each feature contained different information. Therefore, T1 mapping-based radiomic studies may

be better suited than simple mean value measurements of T1 maps since significant T1 mapping information cannot be recovered from conventional T1 mapping analysis.

Radiomics analysis is normally considered as a statistical assessment technique of pixel signal intensity distributions and relationships between neighboring pixels. Not only does it allow quantification of image structure but it also provides ways for pattern identification, which can help enhance diagnostic accuracy and improve risk stratification in medical imaging [39], while traditional native T1 analysis was hard to detect these difference. For example, the scar heterogeneity, quantified by entropy, was reported that it could predict appropriate implantable cardioverter-defibrillator therapy in cardiomyopathies [40].Our study employs this radiomics based approach to explore the association of genotype (*MYH7* and *MYBPC3)* and phenotype in HCM cohorts, demonstrating that such an approach can provide a feasible and powerful tool for the detection of subtle phenotypic differences between these variants, and hence an in-depth mechanistic study would be useful. Therefore, this approach could aid identification, characterization and exploration of genotypic-phenotypic associations in HCM patients with different genotypes and may lead to individualized precise risk stratification.

Limitations

This was a single center study and it is important for the results to be validated at other sites using other types of scanners, field strength and T1 mapping sequences.

Conclusion

Unlike conventional LGE and native T1 analysis, radiomic analysis of T1 images is able to discriminate phenotypic differences between *MYH7*- and *MYBPC3* -associated HCM subgroups.

ACKNOWLEDGMENTS

JW acknowledges support from China Scholarship Council and Genesky Biotechnologies Inc., Shanghai, China. GVG acknowledges support from support from the NIHR Birmingham ECMC, NIHR Birmingham SRMRC, Nanocommons H2020-EU (731032) and the NIHR Birmingham Biomedical Research Centre and the MRC HDR UK (HDRUK/CFC/01), an initiative funded by UK Research and Innovation, Department of Health and Social Care (England) and the devolved administrations, and leading medical research charities. The views expressed in this publication are those of the authors and not necessarily those of the NHS, the National Institute for Health Research, the Medical Research Council or the Department of Health.

REFERENCES

1. Maron BJ, Gardin JM, Flack JM, Gidding SS, Kurosaki TT, Bild DE. Prevalence of Hypertrophic Cardiomyopathy in a General Population of Young Adults Echocardiographic Analysis of 4111 Subjects in the CARDIA Study. Circulation 1995; 92:785-789.

2. Weissler-Snir A, Hindieh W, Gruner C, et al (2017). Lack of Phenotypic Differences by Cardiovascular Magnetic Resonance Imaging in MYH7 (β-Myosin Heavy Chain) - Versus MYBPC3 (Myosin-Binding Protein C)-Related Hypertrophic Cardiomyopathy. Circ Cardiovasc Imaging 2017; 10:e005311.

3. Hinojar R, Varma N, Child N, et al. T1 Mapping in Discrimination of Hypertrophic Phenotypes: Hypertensive Heart Disease and Hypertrophic Cardiomyopathy: Findings From the International T1 Multicenter Cardiovascular Magnetic Resonance Study. Circ Cardiovasc imaging 2015; 8:e3285.

4. Wang S, Zou Y, Fu C, et al. Worse prognosis with gene mutations of beta-myosin heavy chain than myosin-binding protein C in Chinese patients with hypertrophic cardiomyopathy. Clin Cardiol 2008; 31:114-118.

5. Song L, Zou Y, Wang J, et al. Mutations profile in Chinese patients with hypertrophic cardiomyopathy. Clin Chim Acta 2005; 351:209-216.

6.Van Driest SL, Vasile VC, Ommen SR, et al. Myosin binding protein C mutations and compound heterozygosity in hypertrophic cardiomyopathy. J Am Coll Cardiol 2004; 44:1903-1910.

7. Ellims AH, Iles LM, Ling L, et al. A comprehensive evaluation of myocardial

fibrosis in hypertrophic cardiomyopathy with cardiac magnetic resonance imaging: linking genotype with fibrotic phenotype. Eur Heart J Cardiovasc Imaging 2014; 15:1108-1116.

8. Weissler-Snir A, Hindieh W, Gruner C, et al. Lack of Phenotypic Differences by Cardiovascular Magnetic Resonance Imaging in MYH7 (β-Myosin Heavy Chain)-Versus MYBPC3 (Myosin-Binding Protein C)-Related Hypertrophic Cardiomyopathy. Circ Cardiovasc Imaging 2017; 10:e005311.

9. Viswanathan SK, Sanders HK, McNamara JW, et al. Hypertrophic cardiomyopathy clinical phenotype is independent of gene mutation and mutation dosage. Plos One 2017; 12:e187948.

10. White SK, Sado DM, Fontana M, et al. T1 mapping for myocardial extracellular volume measurement by CMR: bolus only versus primed infusion technique. JACC Cardiovasc Imaging 2013; 6:955-962.

11. Rudolph A, Abdelaty H, Bohl S, et al. Noninvasive detection of fibrosis applying contrast-enhanced cardiac magnetic resonance in different forms of left ventricular hypertrophy relation to remodeling. J Am Coll Cardiol 2009; 53:284-291.

12. Hinojar R, Varma N, Child N, et al. T1 Mapping in Discrimination of Hypertrophic Phenotypes: Hypertensive Heart Disease and Hypertrophic Cardiomyopathy: Findings From the International T1 Multicenter Cardiovascular Magnetic Resonance Study. Circ Cardiovasc Imaging 2015; 8:e3285.

13. Gillies RJ, Kinahan PE, Hricak H. Radiomics: Images Are More than Pictures, They Are Data. Radiology 2016; 278:563-577.

14. Neisius U, El-Rewaidy H, Nakamori S, Rodriguez J, Manning WJ, Nezafat R. Radiomic Analysis of Myocardial Native T1 Imaging Discriminates Between Hypertensive Heart Disease and Hypertrophic Cardiomyopathy. JACC: Cardiovasc Imaging 2019; 12:1946-1954.

15. Baeßler B, Mannil M, Maintz D, Alkadhi H, Manka R. Texture analysis and machine learning of non-contrast T1-weighted MR images in patients with hypertrophic cardiomyopathy—Preliminary results. Eur J Radiol 2018; 102:61-67.

16. Yang F, Wang J, Li Y, et al. The prognostic value of biventricular long axis strain using standard cardiovascular magnetic resonance imaging in patients with hypertrophic cardiomyopathy. Int J Cardiol. 2019; 294:43–49.

17. Elliott PM, Anastasakis A, Borger MA, et al. 2014 ESC Guidelines on diagnosis and management of hypertrophic cardiomyopathyThe Task Force for the Diagnosis and Management of Hypertrophic Cardiomyopathy of the European Society of Cardiology (ESC). Eur Heart J 2014; 35:2733-2779.

18. Arbustini E, Narula N, Tavazzi L, et al. The MOGE(S) classification of cardiomyopathy for clinicians. J Am Coll Cardiol 2014; 64:304-318.

19. Wang J, Wan K, Sun J, et al. Phenotypic diversity identified by cardiac magnetic resonance in a large hypertrophic cardiomyopathy family with a single MYH7 mutation. Sci Rep-Uk 2018; 8:973.

20. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal. 2011; 17.

21. Koboldt DC, Zhang Q, Larson DE, et al. VarScan 2: somatic mutation and copy

number alteration discovery in cancer by exome sequencing. Genome Res 2012; 22:568-576.

22. McKenna A, Hanna M, Banks E, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. 2010; 20:1297-1303.

23. Xu Z, Xie C, Xia L, et al. Targeted exome sequencing identifies five novel loci at genome-wide significance for modulating antidepressant response in patients with major depressive disorder. Transl Psychiatry. 2020; 10:30.

24. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res 2010; 38:e164.

25. Ng PC, Henikoff S. SIFT: Predicting amino acid changes that affect protein function. Nucleic Acids Res 2003; 31:3812-3814.

26. Adzhubei IA, Schmidt S, Peshkin L, et al. A method and server for predicting damaging missense mutations. Nat Methods 2010; 7:248-249.

27. Schwarz JM, Rödelsperger C, Schuelke M, Seelow D. MutationTaster evaluates disease-causing potential of sequence alterations. Nat Methods 2010; 7:575-576.

28. Richards S, Aziz N, Bale S, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genet Med 2015; 17:405-424.

29. Wang J, Li W, Sun J, et al. Improved segmental myocardial strain reproducibility using deformable registration algorithms compared with feature tracking cardiac MRI

and speckle tracking echocardiography. J Magn Reson Imaging 2018; 48:404-414.

30. Haralick RM, Shanmugam K, Dinstein I. Textural features for image classification, IEEE Transactions on Systems, Man and Cybernetics 1973; 3:610-621.

31. Haralick RM, Shapiro LG. Computer and Robot Vision. Addison-Wesley1992; 1:460.

32. Ahonen T, Matas J, He C, et al. Rotation Invariant Image Description with Local Binary Pattern Histogram Fourier Features. Image Analysis, 16th Scandinavian Conference, SCIA 2009, Oslo, Norway, June 15-18, 2009. Proceedings. DBLP, 2009.

33. Bianconi F, Alberto Álvarez-Larrán, Antonio Fernández. Discrimination between tumour epithelium and stroma via perception-based features. Elsevier Science Publishers B. V. 2015.

34. Tang X. Texture Information in Run-Length Matrices, IEEE Transactions on image Processing 1998; 7:1602-1609.

35. Jin X, Zheng X, Chen D, et al. Prediction of response after chemoradiation for esophageal cancer using a combination of dosimetry and CT radiomics. Eur Radiol. 2019; 29:6080-6088.

36. Sedaghat-hamedani F, Kayvanpour E, Tugrul OF, et al. Clinical outcomes associated with sarcomere mutations in hypertrophic cardiomyopathy: a meta-analysis on 7675 individuals. Clin Res Cardiol 2017: 30-41.

37. Miller RJH, Heidary S, Pavlovic A, et al. Defining genotype-phenotype relationships in patients with hypertrophic cardiomyopathy using cardiovascular magnetic resonance imaging. Plos One 2019; 14:e217612.

38. Ho CY, Day SM, Ashley EA, et al. Genotype and Lifetime Burden of Disease in Hypertrophic Cardiomyopathy: Insights from the Sarcomeric Human Cardiomyopathy Registry (SHaRe). Circulation 2018; 138:1387-1398.

39. van Rosendael AR, Maliakal G, Kolli KK, et al. Maximization of the usage of coronary CTA derived plaque information using a machine learning based algorithm to improve risk stratification; insights from the CONFIRM registry. J Cardiovasc Comput 2018; 12:204-209.

40. Gould J, Porter B, Claridge S, et al. Mean entropy predicts implantable cardioverter-defibrillator therapy using cardiac magnetic resonance texture analysis of scar heterogeneity. Heart Rhythm. 2019; 16:1242-1250.

Table 1. Overview of extracted features and the feature extraction methods.

Extracted features
Mean, variance, skewness, kurtosis, 5^{th} to 10^{th} central moments of the images (N =10).
Gray-level co-occurrence matrix (GLCM) features: Contrast, correlation, energy, homogeneity, angular second moment and variance of the GLCM with 10 offsets (N= 60) [30, 31].
Local binary pattern (LBP) feature: Rotation invariant image description with LBP histogram Fourier features with sample number=8 and radius=3 (N= 38) [32].
Perception-based features: coarseness, contrast, directionality, line-likeliness and roughness (N = 5) [33]
Gray level run length (GLRL) features: Short- and long-run emphasis, gray-level nonuniformity, run-length nonuniformity, run percentage, low and high gray-level run emphasis, short-run low and high gray-level emphasis, long-run low and high level emphasis of the GLRL matrix with 4 offsets (N= 44) [34].

N, the numbers of features.

Table 2. Comparisons of demographic and cardiac MRI characteristics between *MYH7* and *MYBPC3*-associated HCM patients.

Variables	*MYH7* (n=68)	*MYBPC3* (n=34)	P
Age(years)	41.43±16.55	44.10±12.48	0.433
Male gender，n (%)	32 (47.8)	20 (66.7)	0.132
BMI(kg/m^2)	22.60±3.10	24.15±4.31	0.047
BSA (m^2)	1.62±0.17	1.69±0.03	0.111
SBP (mmHg)	119.00 (111.00, 126.00)	123.50 (112.00, 135.00)	0.179
DBP (mmHg)	72.34±10.16	74.37±10.92	0.378
Asymptomatic, n (%)	32 (47.1)	17(50.0)	0.779
Signs of heart failure such as tachypnea, n (%)	14(20.6)	6(17.6)	0.725
Chest pain, n (%)	18(26.5)	8(23.5)	0.748
Others (Palpitation, dizziness, etc.)	27(39.7)	11(32.4)	0.469
NSVT, n (%)	7(10.3)	3(8.8)	0.815
AF	7(10.3)	4(11.8)	0.821
BNP	1712.8±2612	1668.2±2561	0.849
β blocker, n(%)	61(89.7)	30(88.2)	0.821
ACEI inhibitors or ARB, n (%)	13(19.1)	5(14.7)	0.581

Spironolactone, n (%)	11(16.2)	5(14.7)	0.847
LVEF (%)	62.94 (56.53, 67.80)	61.00 (55.65, 66.00)	0.286
LVEDV(mL)	129.13±28.28	139.81±29.19	0.097
RVEF (%)	58.19±10.59	59.01±9.61	0.722
RVEDV(mL)	99.90±28.87	110.10±25.74	0.106
LVMassi (g/m^2)	69.78 (57.19, 90.02)	75.06 (58.74, 92.29)	0.080
MWT(mm)	21.89±5.69	23.56±10.16	0.441
Tissue characteristics			
LGE %	10.74±0.08 %	11.08±9.11 %	0.797
Mapping			
Global native-T1 (ms)	1275 (1236, 1306)	1283 (1260, 1295)	0.760
Basal native-T1 (ms)	1275 (1227, 1319)	1272 (1214, 1331)	0.914
Middle native-T1 (ms)	1261(1223, 1292)	1267 (1244, 1322)	0.178

*$p<0.05$; cardiac MRI, cardiac magnetic resonance imaging; BMI, body mass index; BSA, body surface area; SBP, systolic blood pressure; DBP, diastolic blood pressure; NSVT, non-sustained ventricular tachycardia; AF, atrial fibrillation; BNP, brain natriuretic peptide; ACEI, angiotensin converting enzyme inhibitors; ARB, angiotensin-receptor blockers; LV, left ventricle; RV, right ventricle; EF, ejection fraction; EDV, end diastolic volume; ESV, end systolic volume; LVMassi, LV mass index; MWT, maximal wall thickness; LGE, late gadolinium enhancement; Continuous variables that are normally distributed are presented as mean ±SD, whereas non-normally distributed variables are presented as median (IQR).

Table 3. Accuracy of radiomic analysis of native T1 images for the discrimination between *MYH7*- and *MYBPC3*-associated HCM patients

	Radiomic analysis of global native T1		Radiomic analysis of basal slice native T1		Radiomic analysis of middle slice native T1	
	Feature selection	Test validation	Feature selection	Test validation	Feature selection	Test validation
Sensitivity (%)	82.3%	75.0%	92.6%	92.6%	89.8%	70.4%
Specificity (%)	97.2%	91.1%	92.2%	80.4%	85.3%	88.2%
Accuracy	92.0%	85.5%	92.3%	84.6%	86.9%	82.1%
AUC	0.968	0.886	0.969	0.887	0.940	0.849
95% CI	0.968-0.971	0.881-0.901	0.969-0.971	0.883-0.902	0.940-0.942	0.843-0.866

AUC, area under the curve; CI, confidence interval.

Figure Legend

Figure 1. Performance of the model's (using global T1 radiomic features) ability to discriminate between *MYH7*- and *MYBPC3* HCM patients.