Machine learning approach in diagnosing Takotsubo cardiomyopathy: the role of the combined evaluation of atrial and ventricular strain, and parametric mapping.

Abstract

Background: CMR with LGE is key diagnostic tool in differential diagnosis between non-ischemic cause of cardiac chest pain. Some patients are not eligible for a gadolinium contrast enhanced CMR; in this scenario the diagnosis remains challenging without invasive examination. Our purpose was to derive a machine learning model integrating some non-contrast CMR parameters and demographic factors to identify Takotsubo cardiomyopathy (TTC) in subjects with cardiac chest pain.

Material and methods: three groups of patients were retrospectively studied: TTC, acute myocarditis, and structural healthy controls. Global and regional left ventricular longitudinal (LS), circumferential (CS), and radial strain (RS) analysis included were assessed. Reservoir (εs), conduit (εe), and booster (εa) bi-atrial functions were evaluated by tissue-tracking. Parametric mapping values were also assessed in all the patients. Five different tree-based ensemble learning algorithms were tested concerning their ability in recognizing TTC in a fully cross-validated framework.

Results: The CMR-based ML ensemble model, by using the Extremely Randomized Trees algorithm with Elastic Net feature selection, showed sensitivity of 92% (95% CI 78 - 100), specificity of 86% (95% CI 80 - 92) and area under the ROC of 0.94 (95% CI 0.90 – 0.99) in diagnosing TTC. Among non-contrast CMR parameters, the Shapley additive explanations analysis revealed that left atrial (LA) εe and εe strain rate were the top imaging markers in identifying TTC patients.

Conclusions: Our study demonstrated that using a tree-based ensemble learning algorithm on non-contrast CMR parameters and demographic factors enables to identify subjects with TTC with good diagnostic accuracy.

Translational outlook:
Our results suggest that non-contrast CMR features can be implemented in a ML model to accurately identify TTC subjects. This model could be a valuable tool for aiding in the diagnosis of subjects with a contraindication to the contrast media. Furthermore, the left atrial conduit strain and strain rate were imaging markers had strong impact on TTC identification. Further prospective and longitudinal studies are needed to validate these findings and assess predictive performance in different cohorts, such as those with different ethnicities, social backgrounds and undergoing different treatments.

Keywords: *Myocardial strain; Takotsubo; CMR; Machine learning; Ensemble Learning; Feature Selection; Cross-validation.*

Abbreviations

CMR cardiac magnetic resonance

ML machine learning

AM acute myocarditis

LS longitudinal strain

CS circumferential strain

RS radial strain

$\boldsymbol{\varepsilon_s}$ Reservoir strain

SRs Reservoir strain rate

$\boldsymbol{\varepsilon_e}$ Conduit strain

SRe Conduit strain rate

$\boldsymbol{\varepsilon_a}$ Booster strain

SRa Booster strain rate

LVEF Left Ventricle Ejection Fraction

FS feature selection

TTC Takotsubo syndrome

LOO leave-one-out

CV cross-validation

LV left ventricle

CMR-FT cardiac magnetic resonance feature tracking

LA left atrium

RA right atrium

Key points

- CMR-derived parameters were supposed to have high discriminatory power in identifying Takotsubo cardiomyopathy.
- Many CMR-derived features were selected and implemented in a tree-based ML ensemble model to recognize TTC patients.
- Left atrial strain parameters proved to be the best non-contrast CMR markers in making TTC diagnosis.

Introduction

In the Emergency Department, acute chest pain of cardiac origin is a common compliant in daily practice which can arise either from ischemic or non-ischemic disease. The first step is consistent with diagnosing or ruling out ST elevation myocardial infarction and non-ST elevated myocardial infarction. As a second step, non-ischemic causes of cardiac chest pain should be considered as well, such as Takotsubo cardiomyopathy (TTC) and acute myocarditis (AM). Making a differential diagnosis between the two is often challenging [1,2]. However, that is crucial, because of its impact on the management and outcome of the patients [1,3]. Although the reference standard for the diagnosis of TTC is represented by coronary angiogram with ventriculography, several non-invasive imaging modalities can useful in the work-up of TTC as well [1,2,4]. Among them, cardiac magnetic resonance imaging (CMR), following the ESC guidelines[5], is progressively gaining more

and more importance [4,6,7]. The specific CMR criteria for the diagnosis of TTC were the combination of typical wall motion abnormalities, myocardial edema, and the absence of evidence of LGE[4].

In clinical practice, CMR acquired with late-gadolinium enhancement (LGE) are often contra-indicated due to concomitant renal disease, to prevent a life-threatening complication after contrast medium administration, namely nephrogenic systemic fibrosis.

In addition to established diagnostic imaging criteria, recent studies have demonstrated the importance of left ventricular and atrial strain parameters in identifying patients with TTC[2,8]. Myocardial strain is a non-contrast quantitative method that use cine images in the routinely acquired CMR examination allowing to quantify the degree of myocardial deformation in different orientations[9]. Myocardial strain imaging has been shown to be a reliable diagnostic tool for both systolic and diastolic evaluation in different cardiac disease, such as hypertrophic cardiomyopathies, dilated cardiomyopathies, myocarditis, and Takotsubo cardiomyopathy[10, 9,11,12]. Strain analysis offers additional information over the traditional ejection fraction or cardiac chamber dimension and provides a sub-clinical assessment of myocardial function[9]. In particular, strain analysis can detect myocardial abnormalities in patients with preserved/recovery EF and wall motion abnormalities, or in case of an absence of myocardial tissue alterations (i.e edema) for a CMR examination delayed due to the lack of scanner accesibility[13].

Cau et al demonstrated a difference in myocardial strain parameters between TTC and AM helping in the differential diagnosis of these two cardiac diseases[2].

Given the possible overlap of symptoms, laboratory data, and electrocardiography changes between TTC and AM[1], may be challenging. Machine learning (ML), a sub-field of Artificial Intelligence whose field of application is steadily growing with several applications in cardiovascular imaging [14,15], has the potential to identify and analyze such complex relationships.

In this paper we compared five different tree-based ensemble learning methods, which were applied to solve the problem of discriminating between TTC and AM in subjects with acute chest pain using a CMR protocol without intravenous contrast administration. A fully cross-validated

pipeline was developed to perform feature selection, training, and evaluation of ML models in an unbiased manner. We hypothesized that non-contrast CMR-derived parameters can accurately and reliably characterize the complexity of the disease and that a ML algorithm could leverage those features to diagnose TTC.

Material and Methods

The data collection and protocols used in this study were approved by the Institutional review board, and individual patient consent was waived because of the retrospective nature of the study.

Study population

In this retrospective single-center study we searched in our database all the patients who underwent CMR between March 3rd, 2017, and February 7th, 2021, because of clinical suspicion of AM or apical ballooning TTC. A total of 43 subjects were finally included in the study cohort. Of those, 18 and 14 were TTC and AM subjects respectively, and 11 were structural healthy controls.

TTC diagnosis was made using the current definition as reported in the Position Statement of the European Society of Cardiology Heart Failure Association [3]. The diagnostic criteria include regional wall motion abnormalities not limited to a single epicardial vascular distribution territory. The disease onset is usually preceded by a stressful trigger, culprit atherosclerotic coronary disease as assessed at invasive catheterization is usually absent, new ECG abnormalities are present, elevated serum natriuretic peptide and a small increase in cardiac troponin are often detected. LV dysfunction full recovery is typically noted at follow-up [3]. Conversely, the diagnosis of AM was made clinically according to that reported in the Position Statement of the European Society of Cardiology Working Group on Myocardial and Pericardial Diseases [16]. Endomyocardial biopsy was not performed. Exclusion criteria were subjects under 18 years of age, contraindication to CMR (implantable devices, severe claustrophobia), history of severe renal disease with a eGFR < 30 mL/min/1.73 m2, and coronary artery disease. The control group comprised structural healthy subjects who had CMR to rule out scar related ventricular tachycardia. Institutional Review Board

approval for this study was obtained and informed consent was waived because of its retrospective nature.

A flowchart demonstrating the application of inclusion and exclusion criteria is provided in **Figure 2.**

CMR acquisition

CMR scans were performed after hospital admission with a mean delay of 4.1 ± SD 2.6 days using a 1.5 T scanner system (Philips Achieva dStream, *Philips Healthcare, Best, The Netherlands*). 8 channels anterior cardiac coil array was used. Cine-CMR examinations were electrocardiogram triggered and performed during breath-hold expiration manoeuvres. Thirty phases were derived for each cardiac cycle. CMR protocol included functional sequences, such as cine bright blood steady-state free precession (SSFP) on the short axis and long axes (2 chambers, 3 chambers and 4 chambers); and morphological and tissue characterization sequences, such as T2 Short Tau Inversion Recovery (STIR) on both short and long axes, T1 mappings and T2 mapping acquisitions. Details of CMR sequence parameters are included in the **Supplementary Methods**.

CMR image post-processing

We used the commercially available software Circle CVI42 (*CVI42, Circle Cardiovascular Imaging Inc., Calgary, Canada*) for cardiac MRI feature tracking (CMR-FT) data analysis. Offline CMR-FT analyses were conducted for evaluation of peak global longitudinal strain (GLS), global radial strain (GRS), and global circumferential strain (GCS) in a 16-segment software-generated 2D model. Concerning GLS, data on myocardial strain were derived from two-, three- and four-chambers long-axis views. Regarding GRS and GCS, data on myocardial strain was derived from apical, mid-ventricular, and basal short-axis views in all the patients. On all images, the epi- and endocardial borders were traced in end-diastole. After that, an automatic computation was triggered,

by which the applied software algorithm automatically outlined the border throughout the cardiac cycle. The quality of the tracking and contouring was visually validated and manually corrected when needed.

CMR-FT analyses of atrial deformation were conducted offline. LA and RA endocardial borders were manually traced on long axis view of the cine images when the atrium was at its minimum volume. In particular, the four-, three-, and two-chamber views were used to derive LA longitudinal strain. LA appendage and pulmonary veins were excluded from segmentation. RA longitudinal strain was based on the four-chamber view only. After manual segmentation, the software automatically tracked the myocardial borders throughout the entire cardiac cycle. The quality of the tracking and contouring was visually validated and manually corrected by a radiologist with 3 years of experience in cardiac imaging. There are three peaks in the strain curve, including reservoir, conduit, and booster strain. Accordingly, their corresponding strain rate (SR) parameters were included.

Global and regional native T1 mapping were assessed on the same commercial post-processing software (*CVI42, Circle Cardiovascular Imaging Inc., Calgary, Canada*) by manually tracing endocardial and epicardial contours. A 10% safety margin was automatically set for both borders to prevent contamination from the blood pool and neighbouring tissues. Finally, the reference point was set at the RV insertion to generate a 16-segment AHA model.

Blinded CMR interpretation

To determine the accuracy of the human diagnostic evaluation of the non-contrast CMR analysis, we performed a dedicated blinded evaluation of all enrolled patients. The non-contrast CMR images were reviewed by an experienced radiologist with 3 years of training in cardiovascular imaging who was blinded to clinical history. The reviewer categorized his interpretation based on the level of confidence for each diagnosis of TTC, rated using a 4-point Likert scale (1=no confidence, 4=high

confidence). The time required for each evaluation was also recorded and compared to the overall time taken by the model, which encompasses feature selection, training, and prediction.

Machine Learning

Machine learning models were trained to identify TTC subjects integrating multiple variables. In particular, forty-two CMR-derived variables on atrial and ventricular strain and parametric mapping and 2 demographic variables (age and gender) were available for feature selection and model derivation (full list provided in **Supplementary Table 1**).

Tree-based ensemble algorithms

Tree-based ensembles can handle high-dimensional data and classify each subject into multiple classes natively. Additionally, they can estimate any (possibly nonlinear) relationship due to their non-parametric nature.

We selected three bagging ensembles - Random Forest [17], Extremely Randomized Trees [18] and a plain bagging classifier of decision trees [19] - and two boosting ensembles, Adaptive Boosting and Extreme Gradient Boosting. In brief, ensemble learning consists in producing a strong estimator by combining the predictions of multiple, weaker estimators that are generally weaker when considered in isolation[20]. Random forests, extremely randomized trees and the plain bagging of decision trees derive the estimator by fitting decision trees on different bootstrap resamples of the data set and combine individual predictions to obtain the final ensemble prediction. Adaptive boosting and extreme gradient boosting build the ensemble sequentially; at each iteration, a weak estimator is fitted on samples having different weights according to the performance of the previous weak learner (additional details are provided in the **Supplementary Methods**). Each algorithm was used to compute individualized probability of TTC, considering multiple variables.

Model Building

Figure 1 shows the steps involved in building and validating ML models using nested leave-one-out (LOO) cross-validation. First, 42 subjects were randomly selected for constituting the training set used for variable selection and model building; the remaining subject was used for testing the resulting model; second, feature selection was performed on the training set; in greater detail, another LOO procedure was used to derive 42 feature sets (one for each subject in the training set) and the union of all sets (all features selected at least once) was computed; third, the model was built using the selected variables; fourth, the independent subject that was set aside at the start of the procedure was classified into TTC, AM or control. The LOO procedure was repeated for each subject, each time using 42 subjects for variable selection and model building and one subject for validation; this ensured that each subject was used exactly once for validation. In the end, each subject was predicted using a model trained with all other subjects. Once finished, the predictions from the 43 models were pooled and used to assess predictive performance in diagnosing TTC. This process was repeated for every ensemble algorithm. The LOO protocol is detailed in **Supplementary Figure 1**. Due to the small sample size, we refrained from tuning algorithm's hyperparameters and used default values (see **Supplementary Table 2** for specific values).

LOO maximizes data available for training and testing, hence reducing overfitting and avoiding overlap between data used to select features and build models, reduces variance in the estimation of generalization error and limits bias when estimating predictive performance of models.

Repeated leave-10-out analysis

To further assess the robustness of our methodology, we repeated the analysis 100 times, each time shuffling the dataset and setting aside roughly 10 subjects as the independent test set and using the remaining subjects for variable selection and model

building. Stratified sampling was performed to ensure each class had the same frequency in both training and test data.

Feature Selection

Feature selection was performed using regularized regression with elastic net penalty, which zeroes the coefficients of less predictive features while handling possible redundancy in the feature set[21,22]. Only features with non-zero coefficients were selected for model training (additional details are provided in the **Supplementary Methods**). Once the nested LOO procedure was finished, the final set of variables was computed as the intersection of the 43 sets; thus, we considered only features selected at every round of the CV protocol.

Feature Importance

To inspect the impact of the features included in the ML model, we analyzed the predictions from the repeated leave-10-out analysis and used the Shapley Additive Explanations (SHAP) framework to derive average importance scores. SHAP exploits a game theoretic approach to attribute individual contributions of features to model predictions[23]. Additional details are provided in the **Supplementary Methods**.

Statistical analysis

In this study, continuous data were described as the mean ± SD. Homogeneity of variance across groups was assessed through Levene's test. Normality of residuals was assessed through Shapiro-Wilks's test. One-way ANOVA, Kruskal-Wallis H and Mann-Whitney U tests were used for comparison between groups of continuous variables, as appropriate. For categorical variables, we used the Chi-squared test. The overall performance of ML models was assessed with the area under the receiver operating characteristics (AUROC) and precision-recall (AUPR) curves.

Sensitivity and specificity were also reported (additional details on performance evaluation are given in the **Supplementary Methods**). The DeLong test[24] was used to test for statistically significant differences between AUC. Calibration was assessed through Brier score and observed vs. predicted plot. Multiple testing correction was done via Benjamini-Hochberg procedure to control the false discovery rate. A P_{adj} value < .05 was considered significant. Confidence intervals at 95% significance level were computed using the efficient Wilson score method with continuity correction[30] for proportions and the percentile method for pooled mean scores of the repeated analysis. All statistical analyses were performed with the R software version 4.1.0 (R Foundation for Statistical Computing, Vienna, Austria) and Python version 3.9.

Results

Patient demographics and CMR parameters.

A total of 32 symptomatic subjects were screened. Of those, 18 were TTC (17 females, mean age 69±10 years.) and 14 patients were AM (5 females, 43±16 years). Additionally, 11 structural healthy subjects (7 females, 50±9 years) who underwent CMR imaging to rule our scar related tachycardia were included as controls (**Figure 2**). Baseline characteristics of included patients are summarized in **Table 1**.

Model comparison and predictive performance

AUCs for the repeated leave-10-out analysis are reported in **Figure 1**. All models showed excellent capabilities in identifying TTC patients among AM and structural controls, except the adaptive boosting. Of note, the extremely randomized trees outperformed all other models in terms of AUROC (0.94, 95% CI [0.90 – 0.99], P < .001 vs all) and sensitivity (91.6%, 95% CI [78 – 100], P < .001 vs all, **Figure 3**). The AUPR was also excellent (0.92, 95% CI [0.85 – 0.98]). The model

also showed good calibration with a Brier score of 0.11 (95% CI [0.08 – 0.13], **Supplementary Figure 2**). For this reason, it was selected as the final model.

The reader's AUCs were worse than the models' (AUROC of 0.52 and AUPR of 0.43). Reader's calibration was substandard, with a Brier score of 7.3 (95% CI [5.6 – 9]). Using a reader's diagnostic evaluation of 2 (Youden's index) or greater as a threshold to classify all subjects as TTC, the sensitivity was 83% (95% CI [58 – 96]) and specificity was 24% (95% CI [10 – 46]). At the Youden's index, the extremely randomized trees model had a sensitivity of 93 (95% CI [92 – 94]) and a specificity of 82%, 95% CI [80 – 83]. In the leave-one-out analysis the performance were similar, except for the random forests' AUROC and sensitivity that were higher than those of extremely randomized trees, although not significant (0.97 vs 0.96, P = .66 and 100% vs 94%, P = 1.0 respectively, **Supplementary Figure 3** and **4**).

To further assess the effectiveness of the feature selection, we evaluated the extremely randomize trees model in the repeated leave-10-out analysis without performing feature selection. The AUROC for the reduced feature set was significantly greater than that for the whole feature set (0.94 vs 0.93, P < .001, **Supplementary Figure 5A**). Sensitivity and specificity were similar, although the model derived using all features had slightly higher sensitivity (91.8 vs 91.6, P < .001, **Supplementary Figure 5B**).

Finally, we compared the time needed to produce a diagnosis. On average, the reader needed 586 s (SD 52) compared to 0.26 s (SD 0.004) for the model, which encompasses the time needed for feature selection, training and prediction (**Supplementary Figure 6**).

Important variables for identifying TTC

The 13 features identified by the feature selection are shown in **Figure 4A** ranked by decreasing impact on identifying TTC patients according to SHAP values. Box-whisker plots showing inter-group differences and pairwise comparisons are reported in **Supplementary Figure 7.**

Overall, age, left atrial (LA) conduit strain and strain rate had the most impact on identifying TTC. Greater age and LA conduit rate and smaller LA Conduit rate values had more impact on identifying TTC subjects (Figure 4B). The S-shaped relationships between LA conduit strain and strain rate (x-axis) and SHAP values (y-axis) for LA conduit and LA conduit rate are shown in **Figure 4C** and **4D**.

Discussion

In the present study, we have demonstrated that a tree-based ML ensemble algorithm trained with non-contrast CMR measurements and demographic information was able to identify patients with TTC accurately when they are mixed with patients with AM and structural healthy controls.

The incremental diagnostic performance of an ML-based model in the diagnosis of cardiac diseases has been demonstrated in previous studies[25–27,28]. Baeßler et al investigated texture analysis of non-contrast T1-weighted CMR images using ML-based approaches in patients with hypertrophic cardiomyopathy showing a sensitivity of 94% and a specificity of 90% of the proposed model in distinguish patients with hypertrophic cardiomyopathy from controls[28]. Gopalakrishnan et al. evaluated a ML-based model that uses CMR parameters to differentiate cardiomyopathies from a cohort of 83 pediatric patients demonstrating an accuracy of 80,72 % with an area under the curve of 0.80[29].

To the best of our knowledge, this is the first work focused on an ML-based model that combines demographic and non-contrast CMR parameters (including atrial and ventricular strain measurements and parametric mapping features) to discriminate between TTC and AM.

In the light of the ESC guidelines, CMR with LGE has acquired a growing role to facilitate the differential diagnosis between ischemic and non-ischemic origin of chest pain, identifying the underlying cause and reducing the need of invasive coronary angiography[5]. In the clinical setting, some patients are not eligible for a gadolinium contrast enhanced CMR; therefore, a non-contrast CMR may be of diagnostic support in the differential diagnosis of these patients. The ability of the

ML model to correctly classify patients with TTC could allow a diagnosis in individuals with a contraindication to the contrast media and reduced associated costs.

Among non-contrast CMR parameters, the SHAP analysis revealed that left atrial (LA) conduit strain and strain rate were the imaging markers with highest impact for TTC identification. The LA plays a key role in maintaining left ventricular filling. Several studies have highlighted the LA significant pathophysiological contribution in different cardiomyopathies [30–32]. Hinojar et Al. investigated LA contractile function using CMR-FT in seventy-five patients with hypertrophic cardiomyopathy [30]. They demonstrated that LA longitudinal function is impaired also in HCM patients with normal LA volume and LV filling pressure in comparison with the structural healthy control group, thus suggesting that LA deformation parameters are sensitive parameters since the early state of the disease [30]. LA strain impairment was reported also in patients with myocarditis [31,32]. Similar LA strain impairment was reported in patients with TTC [8,33]. In this respect, Backhaus et Al. demonstrated altered LA reservoir and conduit functions in TTC patients [33]. In spite of the pathophysiology of the disease not being fully explained, LA dysfunction is reported in TTC patients. It may be caused either by direct atrial involvement or may be secondary to LV dysfunction [34]. Our data are consistent with that reported in literature [8,33,35]. They show that LA conduit strain and strain rate may identify TTC patients among those with acute chest pain via an ML model. LA conduit represents the most sensitive marker of exercise intolerance in patients with heart failure and preserved ejection fraction, reflecting an increased stiffness and decreased elastance of LA, as shown by Von Roeder et Al. [36]. In addition, an association between LA conduit function and early LV filling was noted [36]. A potential “atrial myopathy” in TTC patients, beyond the well-known LV involvement, was intrinsically suggested [36].

Finally, our model outperformed clinical reader’s diagnoses with an average increase in AUROC of 0.42 (80%), in sensitivity of 0.08 (10%) and in specificity of 0.618 (257%). Model calibration was also better (0.11 vs 7.3), but still suboptimal due to slight overestimations of the observed frequencies in the first deciles and slight underestimations in the last deciles. The model

was also much faster, taking, on average, 0.26 seconds for feature selection, training and making the prediction, against 560 seconds needed by the clinical reader to make a diagnosis.

The following study limitations should be acknowledged and addressed in future research before the presented method can be employed in clinical practice. First, the relatively small sample size and the cross-sectional nature of the study. However, we enrolled exclusively very homogenous TTC patients with the classic apical ballooning. Second, although we have taken precautions when training models and estimating generalization performance, our model may still have overfit the data. To guard against overfitting, we employed nested cross-validation to select features with diagnostic value, train algorithms and evaluate performance on the same cohort, by considering non-overlapping subsets of the data and thus reducing the bias in performance estimation[37]. The promising results of our study could prompt further prospective trials including a larger number of patients to confirm our findings. In fact, deriving a ML model that performs well in the general population would likely require a much larger cohort to capture the heterogeneity and nuances of takotsubo syndrome. Third, in our study the predictive value of strain and parametric mapping parameters for adverse cardiovascular events was not assessed at follow-up. Further longitudinal studies are warranted to evaluate the association of these CMR parameters with patient outcome. Fourth, given the retrospective nature of our work, the control group includes patient with a negative CMR examination and considered "structurally" healthy. Further prospective studies that include healthy volunteers are warranted to confirm our results. Fifth, the impairment in strain and parametric mapping measurements in patients with TTC would have probably been different if CMR had been performed within a shorter period of time, ideally the same day of hospital admission. This may have potentially influenced the ML algorithms' performance and should be further tested in the future. Finally, although our model only took a fraction of the time needed by the clinical reader to make a diagnosis, the variable extraction process is lengthy (30 minutes on average). We compared the time to diagnosis of the model trained on variables extracted from non-contrast CMR examinations to that of the clinical reader that directly interpreted the CMR

examinations. The rationale for this choice was to acquire reader's diagnoses in an environment that was as similar as possible to ordinary clinical practice. Considering the time needed to extract variables, the overall model time to diagnosis is, on average, 1800 seconds, almost three times slower than the reader. However, the model's performance are significantly higher. In future studies, a deep learning system able to analyse and automatically extract variables from non-contract CMR scans could be developed and compared to classical ML algorithms trained on variables extracted by the radiologists.

Conclusion

We demonstrated that a tree-based ML ensemble algorithm integrating demographic factors with non-contrast CMR parameters could accurately identify TTC in patients with cardiac chest pain, and as such, it could be an additional support to the clinician in identifying TTC in subjects with cardiac chest pain.

Disclosure

All authors agreed with the content and gave consent to submit.

All authors contributed equally to the work.

The authors state that this work is not under consideration elsewhere and none of its contents have been published previously.

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

All authors read and approved the final manuscript.

Some of the patients under analysis were published in one of our previous studies.

The scientific guarantor of this publication is the corresponding author.

The authors declare that they have no competing interests.

Acknowledgments

We thank Rodrigo Salgado for help with the current study.

References

1. Ghadri J-R, Wittstein IS, Prasad A, et al. International Expert Consensus Document on Takotsubo Syndrome (Part I): Clinical Characteristics, Diagnostic Criteria, and Pathophysiology. *Eur Heart J*. 2018;39(22):2032-2046. doi:10.1093/eurheartj/ehy076
2. Cau R, Bassareo P, Deidda M, et al. Could CMR Tissue-Tracking and Parametric Mapping Distinguish Between Takotsubo Syndrome and Acute Myocarditis? A Pilot Study. *Acad Radiol*. Published online 2021. http://www.sciencedirect.com/science/article/pii/S1076633221000155
3. Lyon AR, Bossone E, Schneider B, et al. Current state of knowledge on Takotsubo syndrome: a Position Statement from the Taskforce on Takotsubo Syndrome of the Heart Failure Association of the European Society of Cardiology. *Eur J Heart Fail*. 2016;18(1):8-27. doi:https://doi.org/10.1002/ejhf.424
4. Eitel I, von Knobelsdorff-Brenkenhoff F, Bernhardt P, et al. Clinical Characteristics and Cardiovascular Magnetic Resonance Findings in Stress (Takotsubo) Cardiomyopathy. *JAMA*. 2011;306(3). doi:10.1001/jama.2011.992
5. Collet J-P, Thiele H, Barbato E, et al. 2020 ESC Guidelines for the management of acute coronary syndromes in patients presenting without persistent ST-segment elevation: The Task Force for the management of acute coronary syndromes in patients presenting without persistent ST-segment elevation. *Eur Heart J*. 2021;42(14):1289-1367. doi:10.1093/eurheartj/ehaa575
6. Eitel I, Lücke C, Grothoff M, et al. Inflammation in takotsubo cardiomyopathy: insights from cardiovascular magnetic resonance imaging. *Eur Radiol*. 2010;20(2):422-431.

doi:10.1007/s00330-009-1549-5

7. Cau R, Solinas C, De Silva P, et al. Role of cardiac MRI in the diagnosis of immune checkpoint inhibitor-associated myocarditis. *Int J cancer*. Published online June 2022. doi:10.1002/ijc.34169

8. Cau R, Bassareo P, Caredda G, Suri JS, Esposito A, Saba L. Atrial Strain by Feature-Tracking Cardiac Magnetic Resonance Imaging in Takotsubo Cardiomyopathy. Features, Feasibility, and Reproducibility. *Can Assoc Radiol J = J l'Association Can des Radiol*. Published online October 2021:8465371211042497. doi:10.1177/08465371211042497

9. Scatteia A, Baritussio A, Bucciarelli-Ducci C. Strain imaging using cardiac magnetic resonance. *Heart Fail Rev*. 2017;22(4):465-476. doi:10.1007/s10741-017-9621-8

10. Palmisano V, Cossa S, Esposito A, et al. Obstructive and Nonobstructive Hypertrophic Cardiomyopathy: Differences in Global and Segmental Myocardial Strain by Cardiac Magnetic Resonance Feature Tracking. *J Thorac Imaging*. 2022;37(1):49-57. doi:10.1097/RTI.0000000000000612

11. Cau R, Loewe C, Cherchi V, et al. Atrial Impairment as a Marker in Discriminating Between Takotsubo and Acute Myocarditis Using Cardiac Magnetic Resonance. *J Thorac Imaging*. 2022;Publish Ah(00):1-7. doi:10.1097/rti.0000000000000650

12. Cau R, Bassareo P, Suri JS, Pontone G, Saba L, Saba L. The emerging role of atrial strain assessed by cardiac MRI in different cardiovascular settings : an up-to-date review Feature tracking. 2022;(1).

13. Xu J, Yang W, Zhao S, Lu M. State-of-the-art myocardial strain by CMR feature tracking: clinical applications and future perspectives. *Eur Radiol*. 2022;32(8):5424-5435. doi:10.1007/s00330-022-08629-2

14. Cau R, Cherchi V, Micheletti G, et al. Potential Role of Artificial Intelligence in Cardiac Magnetic Resonance Imaging. *J Thorac Imaging*. 2021;Publish Ah(3):142-148. doi:10.1097/rti.0000000000000584

15. Cau R, Flanders A, Mannelli L, et al. Artificial Intelligence in Computed Tomography Plaque Characterization: A Review. *Eur J Radiol*. Published online 2021:109767. doi:https://doi.org/10.1016/j.ejrad.2021.109767

16. Caforio ALP, Pankuweit S, Arbustini E, et al. Current state of knowledge on aetiology, diagnosis, management, and therapy of myocarditis: a position statement of the European Society of Cardiology Working Group on Myocardial and Pericardial Diseases. *Eur Heart J*. 2013;34(33):2636-2648, 2648a-2648d. doi:10.1093/eurheartj/eht210

17. Breiman L. Random Forests. *Mach Learn*. 2001;45(1):5-32. doi:10.1023/A:1010933404324

18. Geurts P, Ernst D, Wehenkel L. Extremely randomized trees. *Mach Learn*. 2006;63(1):3-42. doi:10.1007/s10994-006-6226-1

19. Breiman L. Bagging predictors. *Mach Learn*. 1996;24(2):123-140. doi:10.1007/BF00058655

20. Opitz DW, Shavlik JW. Actively Searching for an Effective Neural Network Ensemble. *Conn Sci*. 1996;8(3-4):337-354. doi:10.1080/095400996116802

21. Zou H, Hastie T. Regularization and Variable Selection via the Elastic Net. *J R Stat Soc Ser B (Statistical Methodol*. 2005;67(2):301-320. http://www.jstor.org/stable/3647580

22. Brown PJ, Zhong J, Frood R, et al. Prediction of outcome in anal squamous cell carcinoma using radiomic feature analysis of pre-treatment FDG PET-CT. *Eur J Nucl Med Mol Imaging*. 2019;46(13):2790-2799. doi:10.1007/s00259-019-04495-1

23. Lundberg SM, Lee S-I. A Unified Approach to Interpreting Model Predictions. In: *Proceedings of the 31st International Conference on Neural Information Processing Systems*. NIPS'17. Curran Associates Inc.; 2017:4768–4777.

24. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics*. 1988;44(3):837-845.

25. Lossnitzer D, Klenantz S, Andre F, et al. Stable patients with suspected myocardial ischemia: comparison of machine-learning computed tomography-based fractional flow reserve and

stress perfusion cardiovascular magnetic resonance imaging to detect myocardial ischemia. *BMC Cardiovasc Disord*. 2022;22(1):34. doi:10.1186/s12872-022-02467-2

26. Sundaram DSB, Arunachalam SP, Damani DN, et al. Natural Language Processing Based Machine Learning Model Using Cardiac MRI Reports to Identify Hypertrophic Cardiomyopathy Patients. Published online April 12, 2021. doi:10.1115/DMD2021-1076

27. Overmars LM, van Es B, Groepenhoff F, et al. Preventing unnecessary imaging in patients suspect of coronary artery disease through machine learning of electronic health records. *Eur Hear J - Digit Heal*. 2022;3(1):11-19. doi:10.1093/ehjdh/ztab103

28. 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. doi:10.1016/j.ejrad.2018.03.013

29. Gopalakrishnan V, Menon PG, Madan S. cMRI-BED: A novel informatics framework for cardiac MRI biomarker extraction and discovery applied to pediatric cardiomyopathy classification. *Biomed Eng Online*. 2015;14 Suppl 2(Suppl 2):S7. doi:10.1186/1475-925X-14-S2-S7

30. Hinojar R, Zamorano JL, Fernández-Méndez M [a]A, et al. Prognostic value of left atrial function by cardiovascular magnetic resonance feature tracking in hypertrophic cardiomyopathy. *Int J Cardiovasc Imaging*. 2019;35(6):1055-1065. doi:10.1007/s10554-019-01534-8

31. Dick A, Schmidt B, Michels G, Bunck AC, Maintz D, Baeßler B. Left and right atrial feature tracking in acute myocarditis: A feasibility study. *Eur J Radiol*. 2017;89:72-80. doi:10.1016/j.ejrad.2017.01.028

32. Doerner J, Bunck AC, Michels G, Maintz D, Baeßler B. Incremental value of cardiovascular magnetic resonance feature tracking derived atrial and ventricular strain parameters in a comprehensive approach for the diagnosis of acute myocarditis. *Eur J Radiol*.

2018;104(May):120-128. doi:10.1016/j.ejrad.2018.05.012

33. Backhaus SJ, Stiermaier T, Lange T, et al. Atrial mechanics and their prognostic impact in Takotsubo syndrome: a cardiovascular magnetic resonance imaging study. *Eur Hear J - Cardiovasc Imaging*. 2019;20(9):1059-1069. doi:10.1093/ehjci/jey219

34. Stiermaier T, Graf T, Möller C, et al. Transient left atrial dysfunction is a feature of Takotsubo syndrome. *J Cardiovasc Magn Reson*. 2017;19(1):15. doi:10.1186/s12968-017-0328-8

35. Stiermaier T, Lange T, Chiribiri A, et al. Left ventricular myocardial deformation in Takotsubo syndrome: a cardiovascular magnetic resonance myocardial feature tracking study. *Eur Radiol*. 2018;28(12):5160-5170. doi:10.1007/s00330-018-5475-2

36. von Roeder M, Rommel K-P, Kowallick JT, et al. Influence of Left Atrial Function on Exercise Capacity and Left Ventricular Function in Patients With Heart Failure and Preserved Ejection Fraction. *Circ Cardiovasc Imaging*. 2017;10(4):e005467. doi:10.1161/CIRCIMAGING.116.005467

37. Emerson RW, Adams C, Nishino T, et al. Functional neuroimaging of high-risk 6-month-old infants predicts a diagnosis of autism at 24 months of age. *Sci Transl Med*. 2017;9(393). doi:10.1126/scitranslmed.aag2882

Figure legends

Figure 1. **Summary of steps involved in machine learning (ML) analysis and performance of ML models**. **(A)** ML analysis involved a leave-one-patient-out (LOO) cross-validation (1) where 42 patients were used to perform feature selection (2) using regularized regression with elastic net penalty, (3) ith model building and (4) prediction of TTC and AM for the validation patient. (5) Each patient was used exactly once as test and overall predictions for the 43 patients are used to evaluate prediction performance. CMR indicates cardiovascular magnetic resonance. **(B-C)** Receiver-operating characteristics and precision-recall curves reporting performance of reader and

ML models on identifying TTC patients from repeated leave-10-out testing. AdaBoost indicates adaptive boosting; Bagging, bagging of decision trees; XGBoost, extreme gradient boosting; RF, random forests and ExtraTrees, extremely randomized trees.

Figure 2. **Patient flowchart**. Inclusion and exclusion criteria for patients included in the study. CMR indicates cardiac magnetic resonance; TTC, takotsubo syndrome and AM, acute myocarditis.

Figure 3. **Diagnostic performance of ML models from repeated leave-10-out testing.** Sensitivity and specificity (solid bars) and 95% CI (whiskers) are presented for both ML models and the reader. Abbreviations as in **Figure 1**.

Figure 4. Feature importance for the ML model based on the reduced feature set for identifying TTC patients according to SHAP values. Plots showing the contribution of each feature of the final set to model's predictions according to SHAP values. **(A)** Global feature importance; features are sorted decreasingly by impact according to the mean absolute SHAP, for each variable (solid bars). **(B)** Relationships between features values and impact on model prediction (violin plots). Higher values of LA Conduit Rate contributed positively to predicting TTC, whereas the opposite was observed for LA Conduit. **(C-D)** Detailed trends of impact on TTC prediction for LA Conduit and LA Conduit rate. Smaller values of LA Conduit and higher values of LA Conduit Rate have greater impact on predicting TTC.

Tables

	Takotsubo	Myocarditis	Control	*P*	P_{adj}
Demographics					
Age (mean ± std) (yr)	69 ± 11	44 ± 16	50 ± 10	<.001	**<.001**
Gender (female = 29)	94.4 (17/18)	64.3 (9/14)	63.6 (7/11)	0.052	0.076
Ventricle Functions					
LVEF	58.71 ± 8.9	58.11 ± 5.03	59.24 ± 4.88	0.79	0.79
EDV LV/BSA	72.37 ± 16.03	90.42 ± 18.14	78.84 ± 12.49	0.011	**0.021**
ESV LV/BSA	29.86 ± 9.63	38.57 ± 11.85	32.12 ± 6.13	0.049	0.073
SV LV/BSA	42.49 ± 11.14	51.88 ± 7.37	46.74 ± 8.75	0.029	**0.049**
RVEF	59.36 ± 5.85	55.77 ± 4.17	55.66 ± 2.96	0.056	0.076
EDV RV/BSA	57.12 ± 12.61	87.76 ± 18.77	75.84 ± 19.27	<.001	**<.001**
ESV RV/BSA	23.36 ± 6.1	38.85 ± 10.35	33.97 ± 9.67	<.001	**<.001**
SV RV/BSA	32.56 ± 7.33	48.86 ± 9.98	42.29 ± 11.36	<.001	**<.001**
Ventricular Strain (L)					
Basal RS	42.76 ± 11.15	35.7 ± 10.64	29.86 ± 17.7	0.021	**0.037**
Mid RS	31.3 ± 8.81	27.21 ± 6.9	27.61 ± 16.14	0.304	0.346
Apical RS	24.64 ± 9.25	27.42 ± 12.99	40.68 ± 23.86	0.012	**0.022**
Global RS	30.31 ± 7.35	28.44 ± 7.32	36.14 ± 6.59	0.031	**0.05**
Basal CS	-20.07 ± 2.63	-18.1 ± 3.02	-19.86 ± 3.11	0.142	0.17
Mid CS	-18.07 ± 3.39	-18.17 ± 2.45	-19.62 ± 2.16	0.322	0.355
Apical CS	-17.61 ± 4.85	-20.35 ± 5.37	-24.61 ± 2.81	0.001	**0.003**
Global CS	-16.02 ± 9.13	-18.46 ± 2.61	-20.56 ± 2.17	0.055	0.076
Basal LS	-16.76 ± 3.75	-10.96 ± 3.17	-16.78 ± 2.16	<.001	**<.001**
Mid LS	-13.07 ± 2.99	-11.66 ± 2.34	-18.0 ± 2.91	<.001	**<.001**
Apical LS	-14.03 ± 2.66	-14.88 ± 2.21	-17.57 ± 3.15	0.006	**0.013**
Global LS	-12.91 ± 2.62	-12.45 ± 2.24	-17.76 ± 1.82	<.001	**<.001**
Atrial Strain (L)					
LA Reservoir	24.89 ± 5.92	30.15 ± 7.19	35.6 ± 3.95	<.001	**<.001**
LA Reservoir Rate	1.11 ± 0.32	1.47 ± 0.35	1.53 ± 0.25	0.001	**0.003**
LA Conduit	10.66 ± 4.46	16.32 ± 6.27	21.49 ± 4.91	<.001	**<.001**
LA Conduit Rate	-0.98 ± 0.41	-1.92 ± 0.72	-1.88 ± 0.4	<.001	**<.001**
LA Booster	14.71 ± 6.21	12.39 ± 3.86	13.19 ± 2.13	0.774	0.79
LA Booster Rate	-1.65 ± 0.49	-1.52 ± 0.43	-1.74 ± 0.32	0.459	0.494
Atrial Strain (R)					
RA Reservoir	35.91 ± 22.69	28.26 ± 9.61	38.67 ± 9.19	0.084	0.107
RA Reservoir Rate	2.19 ± 1.48	1.59 ± 0.5	1.95 ± 0.62	0.486	0.51
RA Conduit	21.31 ± 15.12	17.85 ± 6.0	24.3 ± 9.8	0.236	0.276
RA Conduit Rate	-1.68 ± 1.0	-1.45 ± 0.35	-2.2 ± 0.8	0.046	0.072
RA Booster	14.93 ± 8.77	9.68 ± 5.53	13.37 ± 4.82	0.114	0.141
RA Booster Rate	-2.06 ± 1.23	-1.26 ± 0.7	-1.45 ± 0.6	0.076	0.1
Mapping					
Global T1	1141.78 ± 67.45	1058.3 ± 95.58	1031.26 ± 76.59	<.001	**<.001**
Basal T1	1110.42 ± 83.4	1046.14 ± 84.06	1030.83 ± 78.27	0.004	**0.009**
Mid T1	1138.15 ± 67.62	1055.02 ± 108.76	1017.18 ± 78.13	0.001	**0.003**
Apical T1	1209.82 ± 99.69	1080.54 ± 109.37	1053.01 ± 82.27	<.001	**<.001**
Global T2	63.85 ± 4.47	60.08 ± 5.92	54.82 ± 3.9	<.001	**<.001**
Basal T2	58.7 ± 4.17	60.73 ± 5.74	54.68 ± 4.08	0.01	**0.02**

Mid T2	63.18 ± 5.15	58.17 ± 6.5	53.31 ± 2.79	<.001	**<u><.001</u>**
Apical T2	70.38 ± 5.91	61.35 ± 6.61	56.47 ± 6.92	<.001	**<u><.001</u>**

Table 1. Baseline characteristics of takotsubo, myocarditis and structural healthy control patients. Both raw and adjusted P values of inter-group comparison tests are reported, using bold and underlined notation for P values that remained statistically significant after adjustment. EDV indicates end-diastolic volume; LV, left ventricle; BSA, body surface area; ESV, end-systolic volume; SV, systolic volume; LVEF, left ventricle ejection fraction; RVEF, right ventricle ejection fraction; RS, radial strain; CS, circumferential strain; LS, longitudinal strain; LA, left atrium; RA, right atrium; Padj indicates adjusted P value after multiple testing correction.

Figures

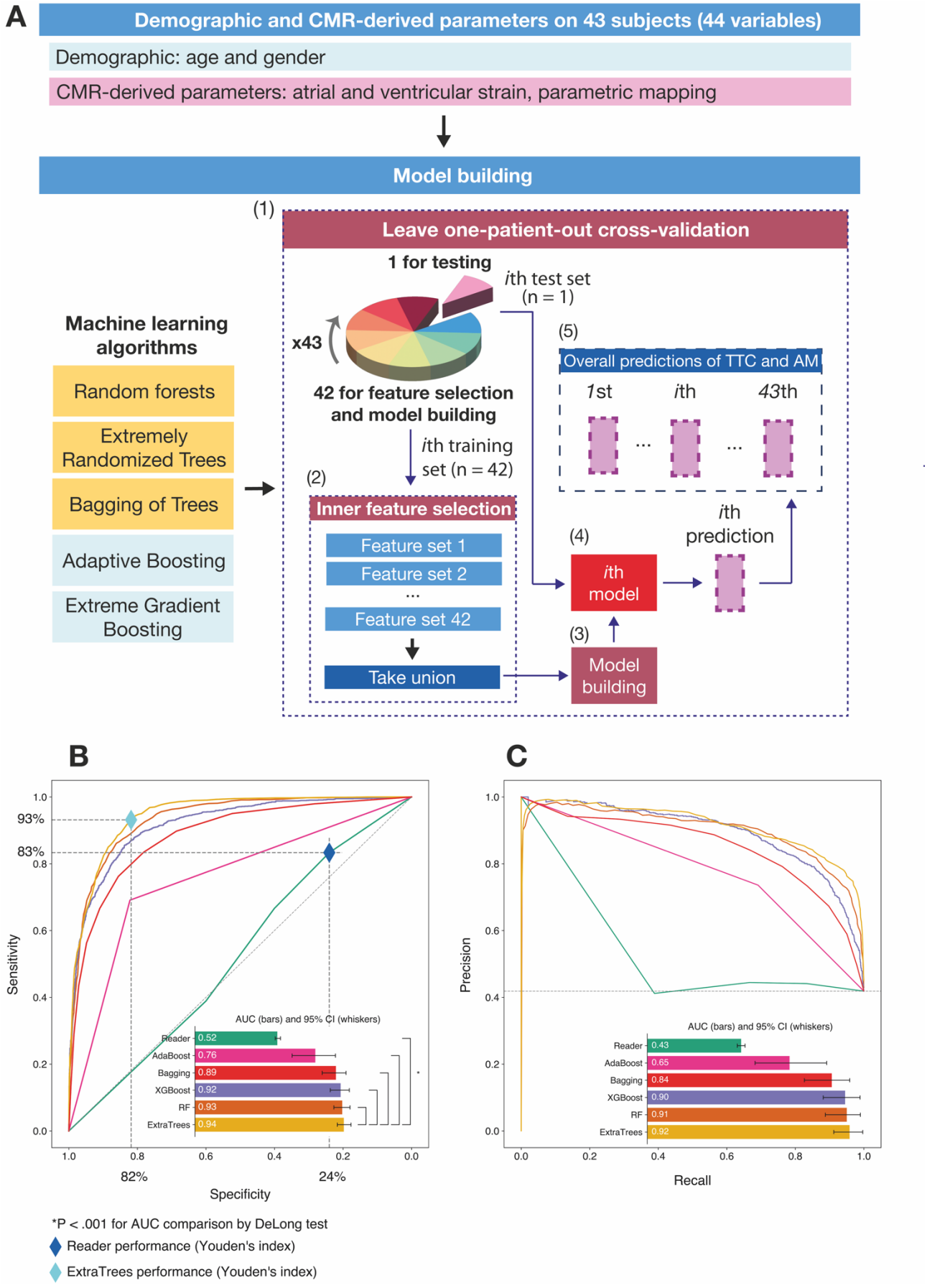
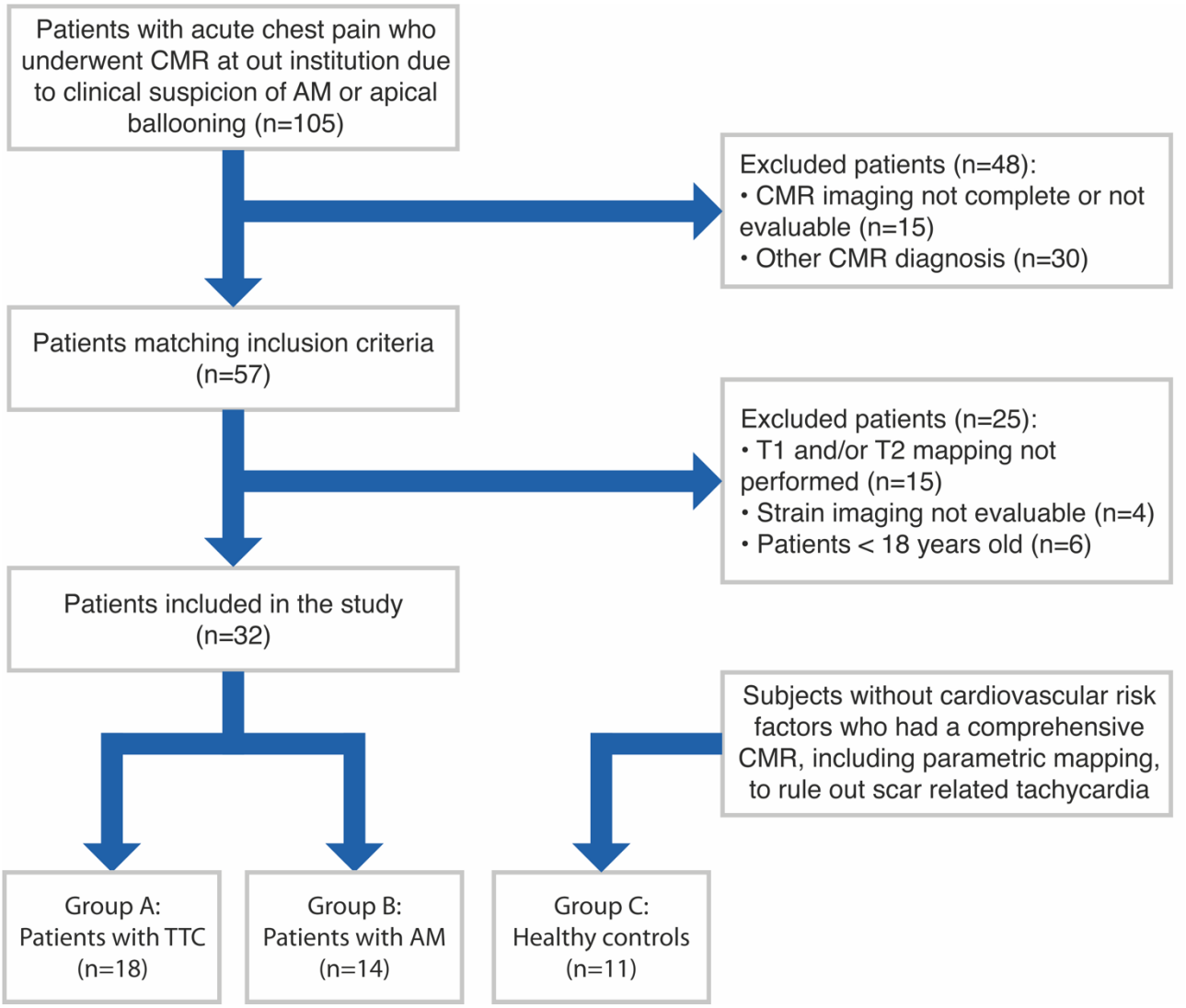
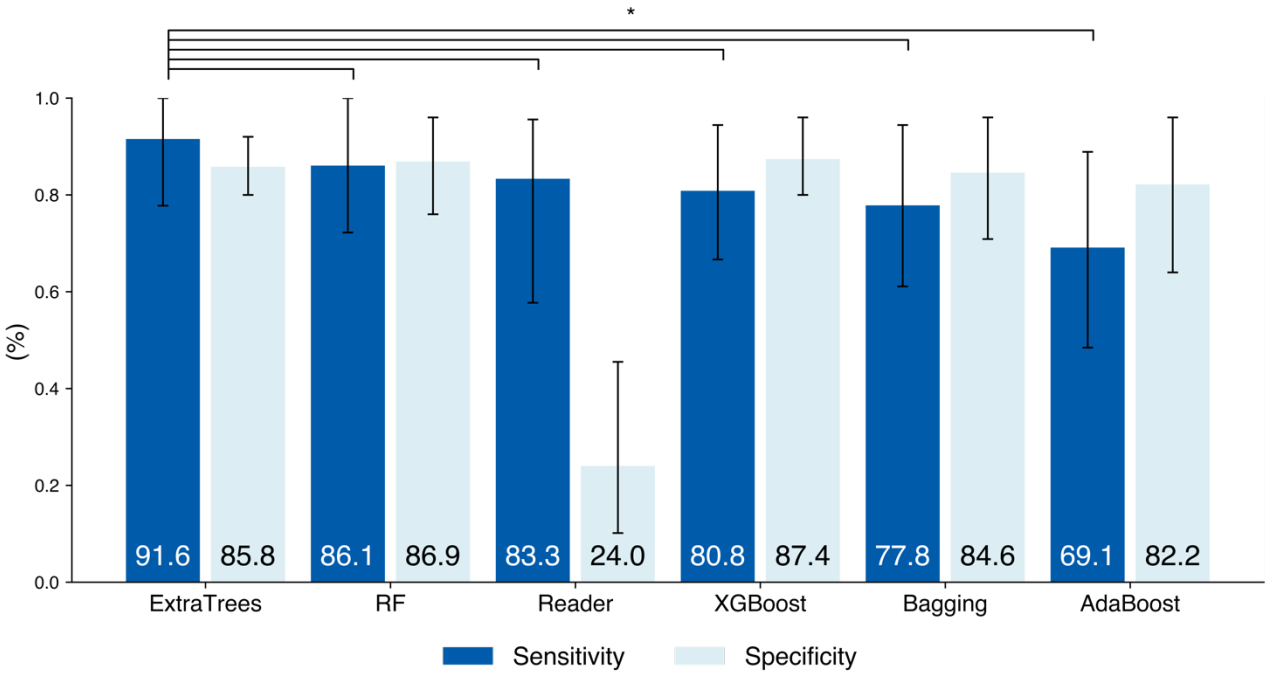

A
Demographic and CMR-derived parameters on 43 subjects (44 variables)
Demographic: age and gender
CMR-derived parameters: atrial and ventricular strain, parametric mapping
Model building
(1)
Leave one-patient-out cross-validation
1 for testing
ith test set (n = 1)
x43
42 for feature selection and model building
ith training set (n = 42)
Machine learning algorithms
Random forests
Extremely Randomized Trees
Bagging of Trees
Adaptive Boosting
Extreme Gradient Boosting
(2)
Inner feature selection
Feature set 1
Feature set 2
...
Feature set 42
Take union
(3)
Model building
(4)
ith model
ith prediction
(5)
Overall predictions of TTC and AM
1st
ith
43th
B
Sensitivity
Specificity
93%
83%
82%
24%
AUC (bars) and 95% CI (whiskers)
Reader 0.52
AdaBoost 0.76
Bagging 0.89
XGBoost 0.92
RF 0.93
ExtraTrees 0.94
C
Precision
Recall
AUC (bars) and 95% CI (whiskers)
Reader 0.43
AdaBoost 0.65
Bagging 0.84
XGBoost 0.90
RF 0.91
ExtraTrees 0.92
*P < .001 for AUC comparison by DeLong test
Reader performance (Youden's index)
ExtraTrees performance (Youden's index)

Figure 2.


Patients with acute chest pain who underwent CMR at out institution due to clinical suspicion of AM or apical ballooning (n=105)
Excluded patients (n=48):
• CMR imaging not complete or not evaluable (n=15)
• Other CMR diagnosis (n=30)
Patients matching inclusion criteria (n=57)
Excluded patients (n=25):
• T1 and/or T2 mapping not performed (n=15)
• Strain imaging not evaluable (n=4)
• Patients < 18 years old (n=6)
Patients included in the study (n=32)
Subjects without cardiovascular risk factors who had a comprehensive CMR, including parametric mapping, to rule out scar related tachycardia
Group A: Patients with TTC (n=18)
Group B: Patients with AM (n=14)
Group C: Healthy controls (n=11)
Figure 2.



*P < .001 for sensitivity comparison by McNemar test

Figure 3.

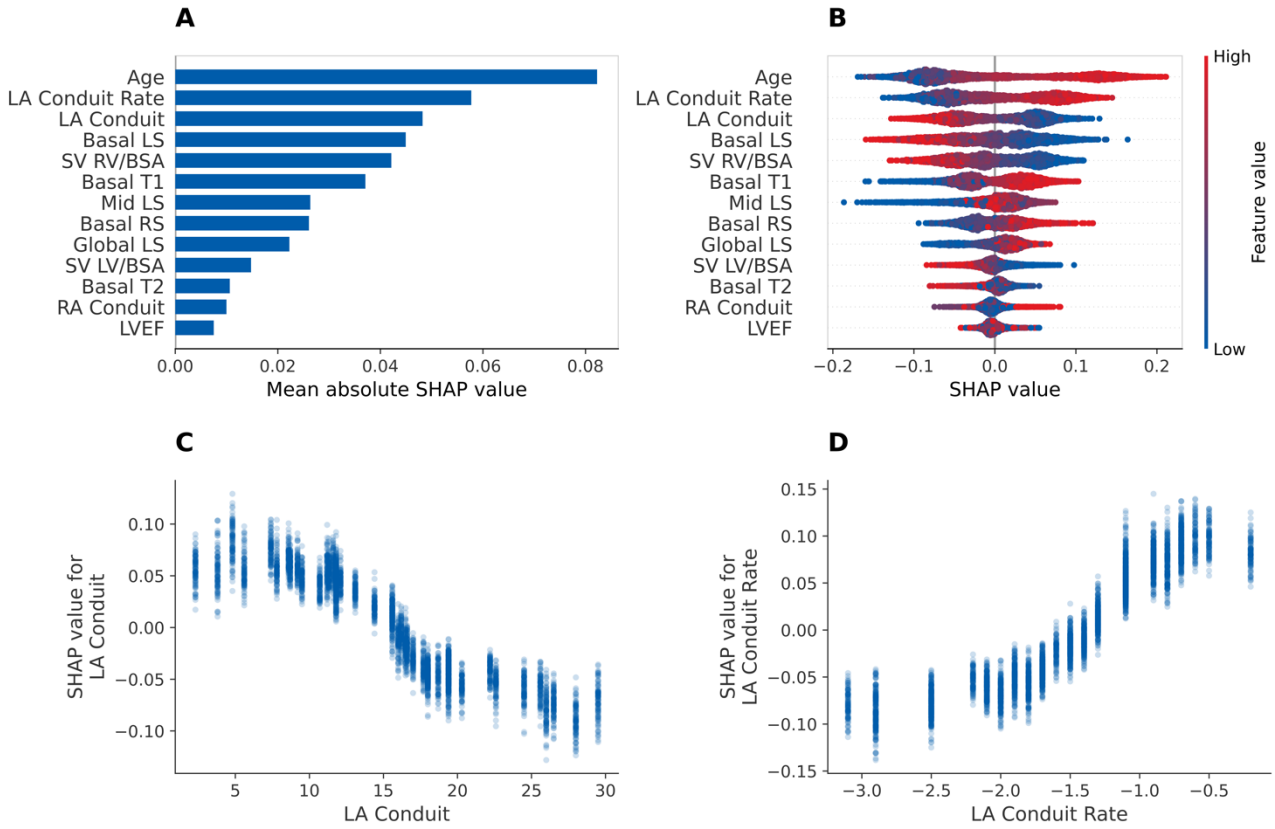

A
Age
LA Conduit Rate
LA Conduit
Basal LS
SV RV/BSA
Basal T1
Mid LS
Basal RS
Global LS
SV LV/BSA
Basal T2
RA Conduit
LVEF
0.00
0.02
0.04
0.06
0.08
Mean absolute SHAP value
B
Age
LA Conduit Rate
LA Conduit
Basal LS
SV RV/BSA
Basal T1
Mid LS
Basal RS
Global LS
SV LV/BSA
Basal T2
RA Conduit
LVEF
−0.2
−0.1
0.0
0.1
0.2
SHAP value
High
Feature value
Low
C
SHAP value for
LA Conduit
0.10
0.05
0.00
−0.05
−0.10
5
10
15
20
25
30
LA Conduit
D
SHAP value for
LA Conduit Rate
0.15
0.10
0.05
0.00
−0.05
−0.10
−0.15
−3.0
−2.5
−2.0
−1.5
−1.0
−0.5
LA Conduit Rate

Figure 4.