New Echocardiographic Risk Score for HCM Patients Follow-up

Marion Taconné[1*], Valentina D A Corino[1], Annamaria Del Franco[2], Eleonora Insinna[2], Pietro Cerveri[1], Iacopo Olivotto[2], Luca Mainardi[1]

[1] Department of Electronics, Information and Bioengineering, Politecnico di Milano, Milan, Italy
[2] Cardiomyopathy Unit, Careggi University Hospital, Florence, Italy

Abstract

Current risk stratification in hypertrophic cardiomyopathy (HCM) often overlooks the full diagnostic potential of echocardiography and provides limited support for longitudinal patient monitoring. This study proposes a novel risk score based only on 26 conventional echocardiographic parameters, trained using machine learning to predict 5-year composite cardiovascular events. We retrospectively analyzed 1,061 HCM patients from the SHARE registry, applying logistic regression, support vector machine, random forest, and gradient boosting classifiers. The composite endpoint included heart failure progression and arrhythmic events. Logistic regression achieved the best performance, with a balanced accuracy of 73.6%, sensitivity of 72.2%, and specificity of 75.0% using nested 5-fold cross-validation. Moreover, longitudinal analysis in patients with serial echocardiographic follow-ups showed that the predicted risk score increased progressively in 83.7% of those who later experienced an event, suggesting value in dynamic risk tracking. These findings highlight the potential of echocardiography-driven machine learning tools to enhance individualized HCM management through both static risk assessment and dynamic follow-up.

1. Introduction

Hypertrophic cardiomyopathy (HCM) is a complex genetic heart disease characterized by left ventricular (LV) hypertrophy, myocardial disarray, and fibrosis. It remains a leading cause of sudden cardiac death (SCD), especially among young individuals, with an annual risk of approximately 1% [1,2]. In HCM patients, arrhythmic events arise from hypertrophied myocardial regions, where fibrosis and disarray create an arrhythmogenic substrate. Although exertion has traditionally been viewed as a common trigger, most sudden deaths occur during rest or mild activity, as shown in unselected cohorts [3].

While implantable cardioverter-defibrillators (ICDs) are effective in the primary prevention of SCD, they involve non-negligible risks, including infection, inappropriate shocks, and psychological distress [4–6]. Therefore, accurate identification of high-risk patients is critical to optimize the balance between benefit and harm.

Current international guidelines rely on clinical scores such as the European Society of Cardiology (ESC) 5-year SCD risk model [1, 5], which integrates pre-selected variables from retrospective cohorts. However, this model demonstrates moderate predictive accuracy, with a reported C-index around 0.69 [1].

Despite the central role of echocardiography in HCM management, its rich quantitative features remain underutilized in most existing risk scores. Machine learning (ML) techniques offer a data-driven, hypothesis-free approach to risk prediction that is capable of capturing more complex interaction. Moreover, little attention has been paid to the temporal dynamics of risk during patient follow-up.

In this study, we propose a novel ML-based risk score built exclusively from conventional echocardiographic features to predict a 5-year composite cardiovascular outcome in HCM patients. Beyond baseline prediction, we also investigate its temporal evolution in follow-up, aiming to support dynamic and personalized disease monitoring.

2. Methods

2.1. Study Population

Patients were retrospectively selected from a monocentric dataset within the SHARE registry. The final dataset included 1,061 HCM patients with baseline echocardiographic, clinical data, and follow-up information.

Only 34 routinely collected echocardiographic parameters were included: age, LV end-diastolic and end-systolic volumes (LVEDV and LVESV), aortic valve regurgitation and grade, aortic valve stenosis, mitral valve regurgitation and grade, mitral valve stenosis, lateral mitral annulus velocity (lat E’), septal mitral annulus velocity (sep E’), A and E wave velocity, E wave deceleration time, E/E’ ratio with lat E’, E/E’ ratio with sep E’, LV outflow tract

 ISSN: 2325-887X DOI: 10.22489/CinC.2025.083

obstruction (LVOTO) location, value at rest with hypovalue (presence or abscence of obstruction), and value with provocation with hypovalue, thickness of the intraventricular septum at diastole, left atrium diameter and volume, LV ejection fraction (LVEF), LV internal diameter at diastole and at systole, maximum wall thickness (MWT) and location, LV posterior wall thickness (PWT), aortic valve systolic anterior motion, aortic sinus dimension, tricuspid valve regurgitation, tricuspid valve stenosis.

The composite outcome at 5 years comprises heart failure progression (including cardiac transplant, device implantation, LVEF $< 35\%$, or worsening NYHA class to III-IV) and arrhythmic events (SCD, aborted cardiac arrest, appropriate ICD therapy). Among the 1,061 patients, 342 (32.2%) experienced the composite outcome within the 5-year follow-up period.

For patients with multiple echocardiographic exams, the first exam was used if no event occurred during follow-up; for patients who experienced an event, the exam immediately preceding the event was selected.

2.2. Feature Selection and Model Training

The objective was to classify patients according to the presence or absence of a composite outcome at 5 years. Features with pairwise correlation greater than 0.75 were removed based on clinical input. The remaining features were subjected to sequential floating forward selection (SFFS) [7], implemented in Scikit-Learn [8]. Feature selection was integrated into a nested cross-validation framework to ensure unbiased performance estimation and prevent data leakage. In each training fold of the outer loop (5-fold stratified cross-validation), an inner loop (3-fold stratified cross-validation) was used for hyperparameter tuning via grid search and for applying SFFS. Four classifiers were evaluated: random forest, logistic regression, support vector machine, and gradient boosting, each with a dedicated hyperparameter grid. The pipeline included median imputation for missing value, standardization, and random undersampling to address class imbalance.

Performance was assessed on the outer test folds using F1 score, sensitivity, specificity, accuracy, and balanced accuracy. ROC curves were averaged across folds to compute the mean AUC with standard deviation.

2.3. Feature Importance Analysis

To enhance model interpretability, SHAP (SHapley Additive exPlanations) analysis was used to quantify the contribution of each feature to individual predictions [9]. For the following classifiers: random forest, support vector machines, and gradient boosting, SHAP values were computed across all five outer cross-validation folds using the best-performing model for each fold. For the logistic regression, feature importance was directly inferred from the model coefficients, which reflect the direction and magnitude of each variable's effect on the predicted log-odds.

2.4. Longitudinal Risk Follow-Up

To explore temporal patterns in predicted risk, we conducted a longitudinal analysis in patients with repeated echocardiographic assessments. The model trained on the first outer fold was applied to all earlier exams available for each patient in the corresponding test set, generating a probability of experiencing the composite outcome at 5 years for each timepoint. This enabled the reconstruction of individual risk trajectories. We then analyzed the evolution of predicted probabilities in patients who experienced the outcome within 5 years after the last exam to assess whether the model captured progressive risk increase over time.

3. Results

3.1. Model Performance

All four classifiers demonstrated comparable performance in identifying patients at risk of experiencing the composite outcome at 5 years (Table 1). 26 features remained after the correlation filter and were subjected to SFFS. Logistic regression achieved the highest overall accuracy (74.1% ± 1.4) and F1 score (64.2% ± 1.5), while support vector machines and random forests offered slightly lower sensitivity and specificity balances. Gradient boosting yielded the lowest overall scores across most metrics. A summary of the cross-validated performance for each model is presented in Table 1. To visualize discriminative ability, Figure 1 shows the mean ROC curves for the four models across the five outer folds, along with the mean ROC curve over the same folds of the ESC risk score for comparison.

3.2. Feature Importance Analysis

Feature importance was assessed for all outer folds based on the absolute values of logistic regression coefficients. Figure 2 reports results from the first test outer fold, which is used in the longitudinal analysis.

The most predictive features included LV diameter, maximum wall thickness, LVOT gradient and location, and age, variables that are part of the ESC risk score computaion. LVEF, featured in the AHA guidelines [2], also ranked highly. Additional influential parameters included mitral and tricuspid valve abnormalities (regurgitation and stenosis), parameter associated with diastolic dysfunction such as sep E', LVEDV, and posterior wall thickness.

Table 1. Cross-validated performance metrics (%) for each classifier. Values are reported as mean ± standard deviation across the five outer folds.

Metric	Random Forest	Logistic Regression	SVM	Gradient Boosting
Sensitivity	73.4 ± 6.7	72.2 ± 4.4	68.4 ± 3.5	70.1 ± 4.7
Specificity	72.6 ± 4.6	75.0 ± 3.5	75.2 ± 3.5	70.7 ± 4.5
Balanced Accuracy	73.0 ± 2.6	73.6 ± 1.2	71.8 ± 1.5	70.4 ± 1.6
F1 Score	63.5 ± 3.0	64.2 ± 1.5	62.1 ± 1.8	60.5 ± 2.0

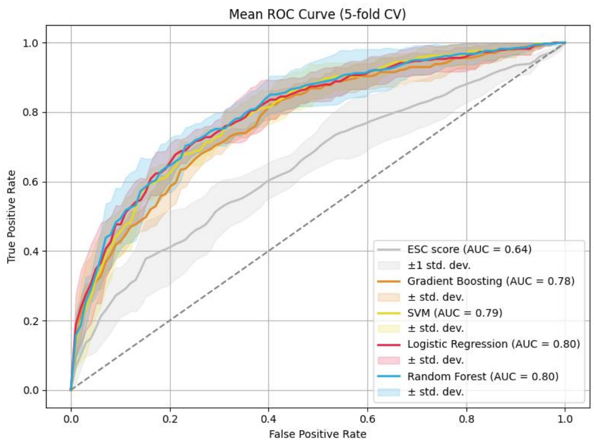

Figure 1. ROC curves of the 5-CV models with area under the curve (AUC), compare to the ESC risk score.

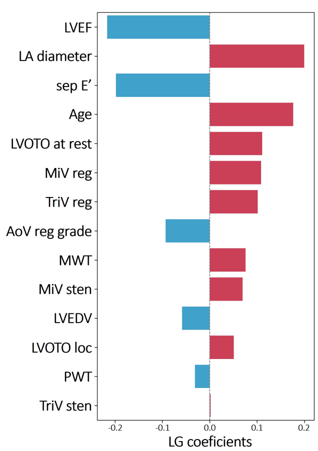

Figure 2. Feature importance in the first fold LG model.

3.3. Longitudinal Analysis

Longitudinal risk trajectories were evaluated on the test set of the first outer fold using the logistic regression model, which achieved the best overall performance. Among the 69 patients who experienced a composite event at 5 years in this test set, 49 (71%) had more than one echocardiographic examination and were included in the dynamic risk analysis. For each of these patients, the model was applied retrospectively to all available exams, yielding serial predictions of the 5-year composite outcome probability.

Across the 49 patients, the predicted risk increased on average by 3.0% ± 5.5 per year. Notably, 83.7% of these patients exhibited a positive slope in their predicted risk score over time, indicating a consistent upward trend prior to the outcome. Figure 3 illustrates these trajectories, displaying model-predicted risk as a function of time, along with the dot line indicating the timing of event occurrence.

4. Discussion

This study introduces a novel echocardiography-based risk score for HCM patients, relying exclusively on routinely acquired features. The proposed model outperformed the ESC 5-year SCD risk score in cross-validated comparisons (p-value = 0.06), reflecting enhanced predictive accuracy for long-term cardiovascular outcomes. Notably, the predicted risk increased progressively in the years preceding an event for the majority of patients, underscoring its potential as a tool for individualized and dynamic disease monitoring.

Feature importance analysis from the logistic regression model trained on the first outer fold revealed key predictors such as LV diameter, maximum wall thickness, LVOT gradient and location, age, and LVEF, variables well aligned with current clinical risk models, including the ESC and AHA scores [1, 2]. Valvular abnormalities, posterior wall thickness, and diastolic markers also emerged as relevant features. Similar patterns of feature importance were observed across other folds. These results suggest that the model integrates both established risk factors and additional echocardiographic markers with potential prognostic relevance, offering a broader characterization of patient risk profiles.

A distinctive aspect of this work is the longitudinal evaluation of risk over time, leveraging repeated echocardiographic examinations to reconstruct individual risk trajectories. In many cases, the follow-up spanned over two to three decades, providing a rare view of risk progression across the history of HCM patients. The observed trend of rising predicted risk preceding clinical events underscores the potential utility of the model not only for baseline stratification but also for ongoing, patient-specific monitoring.

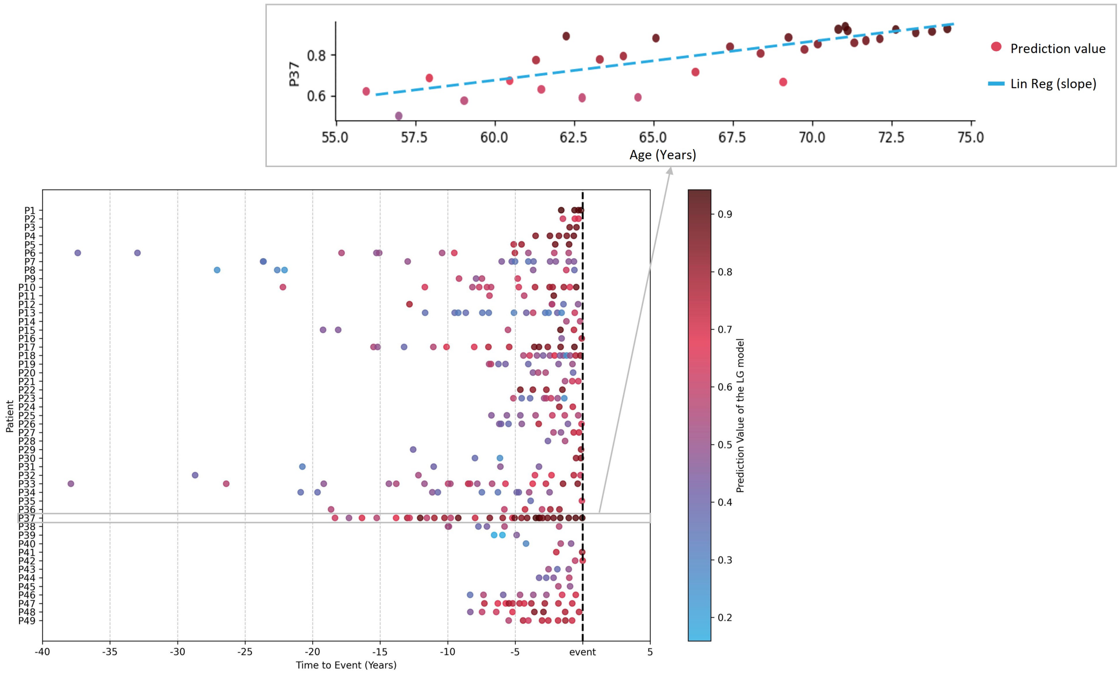

Figure 3. Predicted risk of the best model through follow-up of patients (P1-P49) with event in the first test set fold.

These findings support further investigation into the integration of echocardiographic data into risk models. Future work will include external validation on multicenter datasets and assessment of clinical utility in prospective settings. The proposed tool may ultimately assist clinicians in identifying high-risk individuals and adapting follow-up intensity or therapeutic strategies in a timely and personalized manner.

Acknowledgment

This research study is part of the project INSIGHT-LV and SMASH-HCM funded by the European Union under the Marie Skłodowska-Curie GA #101198472 and GA #101137115 respectively.

References

[1] O'Mahony C, et al. A novel clinical risk prediction model for sudden cardiac death in hypertrophic cardiomyopathy (HCM Risk-SCD). European Heart Journal 2014;35(30):2010–2020. ISSN 15229645.

[2] Ommen SR, et al. 2020 AHA/ACC Guideline for the diagnosis and treatment of patients with wypertrophic cardiomyopathy: A report of the American college of cardiology/American heart sssociation joint committee on clinical practice guidelines. Journal of the American College of Cardiology 2020;76(25):e159–e240. ISSN 15583597.

[3] Maron BJ, Olivotto I, et al. Epidemiology of hypertrophic cardiomyopathy–related death. Circulation 2000; 102(8):858–864. ISSN 0009-7322.

[4] Priori SG, et al. 2015 ESC Guidelines for the management of patients with ventricular arrhythmias and the prevention of sudden cardiac death the task force for the management of patients with ventricular arrhythmias and the prevention of sudden cardiac death of the european society of cardiology (ESC). European Heart Journal 2015;36(41):2793–2867l. ISSN 15229645.

[5] Zeppenfeld K, et al. 2022 ESC Guidelines for the management of patients with ventricular arrhythmias and the prevention of sudden cardiac death: Developed by the task force for the management of patients with ventricular arrhythmias and the prevention of sudden cardiac death . European Heart Journal 2022;43(40):3997–4126. ISSN 0195-668X.

[6] Arbelo E, et al. 2023 ESC Guidelines for the management of cardiomyopathies: Developed by the task force on the management of cardiomyopathies of the European Society of Cardiology (ESC). European Heart Journal 2023; 44(37):3503–3626. ISSN 15229645.

[7] Ferri FJ, Pudil P, Hatef M, Kittler J. Comparative study of techniques for large-scale feature selection. Machine intelligence and pattern recognition 1994;16:404–413.

[8] Pedregosa F, et al. Scikit-learn: Machine learning in Python. Journal of Machine Learning Research 2011;12(9):2825–2830.

[9] Lundberg SM, Lee Si. A unified approach to interpreting model predictions. arXiv preprint arXiv170507874 2017;.

Address for correspondence:

Marion Taconné (marionhelene.taconne@polimi.it)
Politecnico di Milano, Building 21,
Via Camillo Golgi 39, 20133, Milan, Italy