Detection of Left Ventricular Outflow Obstruction from Standard B-Mode Echocardiogram Videos using Deep Learning

Victoria Yuan, BS[1,2], Hirotaka Ieki MD, PhD[3], Christina Binder MD, PhD[2], Yuki Sahashi MD, PhD[2], Paul C. Cheng MD PhD[3], David Ouyang MD[2,4]

1. David Geffen School of Medicine at University of California, Los Angeles
2. Department of Cardiology, Smidt Heart Institute, Cedars-Sinai Medical Center, Los Angeles, CA
3. Division of Cardiology, Department of Medicine, Stanford University, Palo Alto, CA
4. Division of Research, Kaiser Permanente Northern California, Pleasanton, CA

Corresponding author:

Dr. David Ouyang

David.ouyang@kp.org

4480 Hacienda Drive

Pleasanton, CA 94588

(832)-495-1605

Disclosures

Dr. Ouyang reports support from the National Institutes of Health (R00HL157421, R01HL173487 and R01HL173526) and Alexion, and consulting or honoraria for lectures from EchoIQ, Ultromics, Pfizer, InVision, the Korean Society of Echocardiography, and the Japanese Society of Echocardiography. The remaining authors have nothing to disclose.

Funding

Ms. Yuan gratefully acknowledges funding from the Sarnoff Cardiovascular Research Foundation.

Total Word Count: 3,241

Abstract

Background Hypertrophic cardiomyopathy (HCM) affects 20 million individuals globally, with increased risk of sudden death and heart failure. While cardiac myosin inhibitors show great promise as disease-specific treatment, current indications are for obstructive HCM. Obstruction is not always well characterized by echocardiography. Artificial intelligence (AI) might assist in the improving the underdiagnosis of left ventricular outflow tract (LVOT) obstruction.

Objective We aimed to develop a deep learning model to detect LVOT obstruction from non-Doppler B-mode echocardiography.

Methods We identified 2,396 patients with LVOT obstruction and 6,177 control patients matched by age, sex, and septal thickness. LVOT obstruction was defined as the presence of an LVOT gradient or systolic anterior motion of the mitral valve on final echocardiography report. A deep learning model was trained on non-Doppler apical-4-chamber (A4C) B-mode echocardiographic videos to detect the presence of outflow obstruction identified later in the same study by spectral Doppler. We evaluated our model on held-out test sets from Cedars-Sinai Medical Center (CSMC), Stanford Healthcare (SHC), and Kaiser Permanente Northern California (KPNC).

Results In a test set of 3,848 videos from CSMC, our model demonstrated strong performance, detecting LVOT obstruction with area under the receiver operating characteristic curve (AUC) of 0.858 (95% confidence interval [0.847, 0.870]). The model demonstrated generalizable performance in the KPNC cohort with AUC of 0.817 [0.740, 0.922] and SHC with AUC 0.836 [0.807, 0.827]. Performance was consistent across patient subgroups, including those with hyperdynamic LV function, pre-existing valvular disease, and small LV cavity size.

Conclusion In this study, we developed an AI model to detect LVOT obstruction from standard A4C videos, highlighting patients who may benefit from more detailed cardiac workup for obstructive HCM.

Keywords: Artificial intelligence; deep learning; obstructive hypertrophic cardiomyopathy; echocardiography

Abbreviations List

1. HCM – hypertrophic cardiomyopathy
2. SAM – systolic anterior motion
3. LVOT – left ventricular outflow tract
4. A4C – apical 4-chamber
5. CSMC – Cedars-Sinai Medical Center
6. IVSd – diastolic interventricular septal thickness
7. SHC – Stanford Healthcare
8. CNN – convolutional neural network
9. AUC – area under the receiver operating characteristic curve
10. Grad-CAM – Gradient-weight Class Activation Mapping

Introduction

Hypertrophic cardiomyopathy (HCM) is an autosomal dominant disease that affects 20 million individuals globally[1] and increases risk for arrhythmia, sudden cardiac death, and severe heart failure[1–3]. The underlying pathophysiology of obstructive HCM involves myocyte disorganization, systolic anterior motion (SAM) of the mitral valve, and subsequent left ventricular outflow tract

(LVOT) obstruction[4]. Mavacamten, a cardiac myosin inhibitor, was shown in multiple clinical trials to reduce LVOT obstruction, improve heart failure symptoms, and prevent negative cardiac remodelling[5,3,6]. However, LVOT obstruction is frequently missed, with an underdiagnosis rate of nearly 90%[7,8]. In the United States alone, nearly 750,000 patients with HCM are estimated as undiagnosed, and 75% of these patients are estimated to have LVOT obstruction[4,7,9,10]. Approximately 60% of patients with obstructive HCM experience delayed diagnoses, requiring over 4 visits on average before receiving a definitive diagnosis and appropriate management[11]. Given the prevalence of obstructive HCM and its clinical burden, there is a persistently unmet need for earlier disease detection[11,12].

Transthoracic echocardiography is the primary tool to diagnose LVOT obstruction; however, evaluation requires careful expert sonographer evaluation and additional images and Doppler measurements for a thorough evaluation. Because LVOT gradients are dynamic, they may also require provocative measures such as Valsalva maneuvers or pharmacologic administration to accentuate the gradient for detection[7]. An elevated LVOT gradient is present at rest in only 20% of patients with obstructive HCM, while the rest have LVOT obstruction that becomes apparent only with provocation or stress testing[13]. Consequently, nearly half of patients with obstructive HCM are missed on resting echocardiography[11,14]. Thus, identifying individuals in standard images who may have LVOT obstruction can inform more detailed assessment and improve detection of obstructive HCM.

Deep learning analysis of echocardiography has the ability to accurately assess cardiac function and detect cardiovascular diseases with high precision[15–20]. AI models have been shown to improve measurement of common parameters such as left ventricular ejection fraction[18], as well as evaluate echocardiograms for severity of valvular disease[17,19,20], amyloidosis[15], and cardiac hypertrophy[15]. Deep learning has also shown promise to integrate information from multiple views, mimicking human evaluation of echocardiographic studies[21,22]. Existing work to diagnose HCM with deep learning leverage electrocardiograms[23–25], which have limited spatial information compared to echocardiography, and do not address obstruction[26,27]. Given the challenges in detecting obstructive HCM and the resulting underdiagnosis, machine learning can facilitate early identification of LVOT obstruction. In this study, we developed a deep learning model that predicts the presence of LVOT obstruction from standard apical four-chamber (A4C) view videos and evaluated the model in two geographically distinct cohorts.

Methods

Study Cohort

Transthoracic echocardiogram studies at Cedars-Sinai Medical Center (CSMC) from patients who received care between 21 April 2014 and 05 June 2022 were obtained for model development. Patients with LVOT obstruction were identified by analyzing the study text reports where physicians report their final findings. Case patients were defined as patients whose final echocardiogram reports found a measured LVOT gradient at rest; LVOT gradient inducible with provocative testing; intra- or mid-cavitary gradient; or SAM without a measurement for LVOT gradient. Control patients had no reported LVOT obstruction, LVOT gradient, or SAM. Patients who had a history of transcatheter aortic valve replacement or myomectomy were excluded from the cohort. Patients with incomplete data for sex, age, or diastolic interventricular septum thickness (IVSd) were also excluded. Case and control patients were matched by sex, age ± 10 years, and IVSd ± 0.2 cm in a 2:1 ratio to cases. A4C videos of each patient were then extracted. For patients who had multiple A4C videos, each video was considered an independent sample.

Geographically distinct cohorts at Stanford Healthcare (SHC) from patients who received care from 2007-2018 and at Kaiser Permanente Northern California (KPNC) from patients who received care from 2022-2024 were identified for external validation of the AI algorithm. Given the clinical protocols at SHC, cases were identified based on reported LVOT peak velocity rather than gradient. Cases were defined as patients with LVOT peak velocity greater than 274 cm/s, which is equivalent to a gradient ≥ 30 mmHg. At KPNC, cases were defined as patients with reported LVOT gradient ≥ 30 mmHg whereas negative controls had either no gradient or a measured gradient < 30 mmHg. This study was approved by the Institutional Review Boards at CSMC, SHC, and KPNC. A pre-specified study protocol was not prepared, and the study was not registered given the retrospective nature of the analysis. There was no patient or public involvement during the design, reporting, or dissemination of the study.

Development of Deep Learning Model

All model training, validation, and evaluation were performed in Python v3.10.12 and PyTorch v2.2. To construct the deep learning model, the dataset was split using 80% for training, 10% for validation, and 10% as a held-out test set to evaluate model performance. Patients and their associated videos were randomized to each set such that there was no data leakage between training, validation, and test sets. If a patient had multiple videos, each video was treated as an independent example in its corresponding set. Video clips were extracted at a sampling rate of 2 with a minimum frame requirement of 32 to ensure that 16 clips were extracted from video. Frames were resized to a height and width of 112 and 112 pixels, respectively, and normalized to the range [0,1]. Data augmentation was not performed. We then trained a binary classifier using an 18-layer R(2+1)D convolutional neural network (CNN)[28]. The CNN was not pre-trained, and weights were randomly initialized. Model training was performed on one NVIDIA RTX 3090 graphics processing unit with a learning rate of 0.01, batch size of 24, binary cross entropy loss, and Adam optimization with an epsilon of 1e-8 and weight decay of 1e-2. Training was performed for a maximum of 100 epochs with early

stopping if the validation loss was stable for 15 consecutive epochs. Alpha dropout layers and gradient clipping were not utilized. The model was selected based on the epoch with the minimum validation loss. A sigmoid activation function was then applied to model logits to obtain the final class probabilities.

Statistical Analysis

All statistical analysis was conducted in Python v3.10.12 with built-in packages scipy 1.12.0 and sklearn v1.2.0; visualizations were created using the matplotlib v3.8.2. To match cases and controls, sex, age, and IVSd were selected as covariates to account for the more significant confounders for obstructive phenotype. Case and control patients were matched by sex, age ± 10 years, and diastolic interventricular septum thickness (IVSd) ± 0.2 cm in a 2:1 ratio to cases with replacement. Standardized mean differences for age, sex, and IVSd between cases and controls were calculated to ensure balanced matching.

To determine a diagnostic threshold, we first determined the Youden index in the validation set as the difference between the true positive rate and the false positive rate; we then quantified sensitivity, specificity, and positive predictive value (PPV) at different thresholds of model predictions. Given the higher disease prevalence in the test cohort compared to the epidemiological prevalence of obstruction, we also investigated diagnostic thresholds and PPV in cohorts with simulated disease prevalences of 0.5%, 1.0%, 2.5%, 5.0%, 10.0% to better contextualize performance. The final diagnostic threshold was chosen to maximize specificity while maintaining a minimum sensitivity of 50%. Model performance was evaluated by calculating the area under the receiver operating curve (AUC), sensitivity, specificity, PPV, and negative predictive value (NPV) across all 3 sites. 95% confidence intervals were calculated for AUC, sensitivity, specificity, PPV, and NPV by bootstrapping with 10,000 samples. Model confidence was evaluated with calibration curve analysis and Brier score.

Subgroup analysis was conducted to characterize model performance in patients with diverse presentations of obstruction and similar phenotypes. We evaluated the performance of our model on subgroups of patients with mitral regurgitation, aortic stenosis, hyperdynamic LV function, and small LV cavity size. Patients were also subset into cohorts of resting gradients, 366s detected with only provocative maneuvers, or intracavitary gradients using their final echocardiography reports. Finally, model performance was analyzed by severity of LV hypertrophy as reported by the interventricular septum thickness. Patient demographics and comorbidities were extracted from electronic health records using ICD-10 codes. Subgroups were identified using echocardiography parameters from clinical measurements and clinicians' final study reports. Patients with LV hypertrophy and the degree of hypertrophy were identified using final echocardiography reports. Patients with small LV cavity size were similarly identified based on clinicians' final echocardiography reports and study sex-specific thresholds for left ventricular end-diastolic volume index ($LVEDV_i$) measured in

echocardiography studies. Finally, we utilized Gradient-weighted Class Activation Mapping (Grad-CAM) to conduct saliency mapping for model interpretability. Briefly, Grad-CAM leverages the gradients in the final layer of our CNN model to compute feature activation maps[29]. We applied Grad-CAM to each individual frame of an A4C video to identify key features highlighted by the model.

Data and Code Availability

Our code and model weights are available on GitHub at https://github.com/echonet/obstruction.

Results

Study Cohort

From CSMC, 2,396 patients were identified as reported to have LVOT obstruction; the case cohort thus consisted of 3,622 echocardiography studies with 16,426 A4C videos. The control cohort included 4,859 patients matched by IVSd (standardized mean difference 0.02, Kolmogorov-Smirnov test statistic = 0.01, p-value = 0.65), age (standardized mean difference 0.02), and sex (standardized mean difference 0.04) without reported LVOT obstruction, resulting in 5,316 echocardiography studies with 22,180 A4C videos (Supplementary Figure 1). Despite the balance by Kolmogorov-Smirnov test, we recognize that there were residual differences in LVEF (58.1% vs 70.4%) and comorbidity prevalence between the cases and controls. Of the 65% of case patients with a measured LVOT gradient, 57% had resting gradients < 30 mmHg. Case and control patients had similar racial backgrounds and rates of mitral regurgitation and left atrial dilation with left atrial volume index (LAV_i) > 34 mL/m^2. Case patients had higher mean LVEF and lower rates of comorbidities such as diabetes, hypertension, atrial fibrillation, and coronary artery disease (Table 1). The data was randomly split on the patient level into training, validation, and test sets (Table 2). Training, validation, and test sets had similar distributions of age, sex, prevalence of cases, and comorbidities.

To externally validate our model at SHC, we identified a cohort consisting of 4 case patients with 4 studies and 4 videos and 362 control patients with 362 studies and 362 videos. At KPNC, the case cohort comprised of 297 patients with 325 echocardiography studies and 2,523 A4C videos and 591 patients with 648 studies and 4,570 A4C videos. While case and control patients were similar in age, the controls had more males and higher rates of comorbidities including hypertension, diabetes, and atrial fibrillation. Compared to the CSMC cohort, patients at KPNC had lower rates of diabetes but higher rates of comorbidities such as hypertension, atrial fibrillation, and coronary artery disease.

Model Performance across Three Centers

Our model demonstrated strong performance across three institutions in detecting of LVOT obstruction from standard A4C videos from the same study. In the internal test set of CSMC patients, the model showed an AUC 0.858 (95% CI [0.847, 0.870]) with a sensitivity of 63.0% [60.6%, 65.3%] and specificity of 88.5% [87.2%, 89.8%] (Central Illustration, Table 3).On an external dataset from

KPNC, the model performed similarly well with an AUC of 0.817 (95% CI [0.807, 0.827]), sensitivity of 71.1% (95% CI [69.3%, 72.9%]), and specificity of 76.5% (95% CI [75.3%, 77.8%]). At SHC, where the prevalence of obstruction was 1.0%, model AUC was 0.836 (95% CI [0.740, 0.922]), with a sensitivity of 100.0% (95% CI [100.0%, 100.0%]) and specificity of 56.9% (95% CI [51.9%, 62.0%]) (Central Illustration, Table 3). Across simulated disease prevalences of 0.5% - 10.0% in the CSMC test cohort, strong AUCs were maintained at 0.830 – 0.957 (Supplementary Table 2). We also evaluated the model's positive predictive value (PPV) across simulated disease prevalences and model thresholds to define positive detection; we noted that population prevalence of rare diseases is dominant when assessing PPV (Supplementary Table 3, Supplementary Figure 3). Calibration curve analysis demonstrated a calibration slope of 0.969, intercept of 0.038, and Brier score of 0.15 (Supplementary Figure 4), further supporting strong model performance and confidence.

The model performed consistently well in patient subgroups stratified by hyperdynamic LV function, LV hypertrophy, and mitral regurgitation (Supplementary Figure 2, Supplementary Table 1). However, while AUCs were robust in the CSMC subgroups with and without aortic stenosis and with and without small LV cavity; performance decreased in the KPNC subset with co-morbid aortic stenosis. Such variability could be related to the challenge in diagnosis of HCM in patients with aortic stenosis, as aortic stenosis alone can cause increased wall thickness given high afterload. Overall, AUCs across all patient subgroups ranged from 0.831 – 0.908 and AUCs of 0.657 – 0.881 in CSMC and KPNC, respectively. The AUC for patients with LV hypertrophy were 0.834 (0.814, 0.854) in the CSMC cohort and 0.808 [0.793, 0.823] in the KPNC cohort. Model performance to detect obstruction in the CSMC cohort was robust across the spectrum of wall thickness, measured gradient, and across phenotypes of resting, provoked, and intracavitary gradients (Supplementary Tables 4, 5, and 6).

Investigating Model Explainability

Saliency mapping demonstrated that the obstruction model focuses on the left ventricular septum at the location of obstruction (Figure 1). Because control and case patients were matched by IVSd, these results imply that there may be additional characteristics of the septum beyond thickness that contribute to LVOT obstruction, such as septal motion. Additionally, Grad-CAM identifies the medial mitral annular plane as a location with significant activation signal, which reflects the pathophysiology of obstruction. These findings on interpretability suggests that our model is learning clinically relevant features of obstruction.

Discussion

In this work, we developed and validated a deep learning-based model to identify LVOT obstruction from A4C videos, potentially serving as an aid for sonographers to identify patients that should undergo further evaluation for HCM. Our model demonstrated strong performance across two geographically distinct clinical centers with different protocols for measuring obstruction (AUCs

≥0.858). Model performance was consistent across a range of LVEF, valvular disorders, and different cardiac geometries with AUCs of 0.809 - 0.952. Interpretability analysis emphasized that our model utilizes features related to the pathophysiology of obstruction for its predictions.

Epidemiological studies emphasize that obstructive HCM is an underrecognized, rather than a rare disease. Patients often present with nonspecific symptoms and require multiple specialty visits before receiving a definitive diagnosis[11]. Guideline-based evaluation includes interrogation for LVOT obstruction, which requires additional time and effort during echocardiography[30]. Though the underdiagnosis of obstructive HCM is multifactorial, nearly 50% of patients with obstruction are not detected with resting echocardiography, requiring additional diagnostic measures[11,13,14]. Such challenges contribute to delayed diagnoses and even inappropriate treatment of patients with obstructive HCM[11]. Our findings suggest that deep learning can facilitate efficient recognition of patients who may benefit from further workup. Our work offers an assistive tool that can be deployed in real-time on the ultrasound machine to alert the sonographer with the probability of obstruction and thus facilitate more detailed interrogation of obstruction, such as obtaining additional Doppler images with and without provocative maneuvers. When combined with clinical histories, our deep learning model can help address diagnostic gaps in obstructive HCM and assist in connecting patients to appropriate clinical care.

Previous work leverages machine learning to detect HCM from echocardiography[26,27], distinguish HCM from other cardiomyopathies[15,31,32], and predict cardiovascular complications of the disease[33–35]. While these models showed strong performance, they do not distinguish between obstructive and non-obstructive HCM. Our deep learning model expands upon these diagnostic efforts by detecting obstruction, which can more thoroughly characterize patient's disease state. Moreover, to facilitate integration into clinical workflows, we trained our model using standard B-mode A4C videos. Our model can be combined with others that diagnose HCM from A4C videos[15] to provide complementary avenues for detection. Parallel contemporary efforts to detect obstruction include a single-view model by Park, et al.[36] that utilizes parasternal long-axis (PLAX) videos and a multi-view model by Holste, et al.[22] that detected elevated LVOT pressures. While Park, et al. developed their model using data from 722 patients and Holste, et al. using data from 249 patients with elevated pressure, our work leverages echocardiography videos from 2,936 patients with obstruction. With a larger training set and validation in two independent clinical sites, our model may be more robust and generalizable to diverse phenotypes of obstruction. Our model can be combined with others that diagnose HCM from A4C videos to provide complementary avenues for detection or with models based on different views to integrate multi-view information. Similarly, incorporation of our model with those that predict elevated LVOT gradients can facilitate more accurate detection.

Prior to clinical deployment, external validation of models is essential to thoroughly characterize performance in diverse care centers. In contrast to studies that are limited to only one

institution[23,25,27], we externally validated our model at KPNC and SHC, emphasizing that our AI algorithm may be robust to distinct clinical protocols and patient populations. Overall, strong performance across subgroups suggests that our model is robust to diverse patient phenotypes and that it learned features of obstruction shared amongst these groups. Finally, we conducted saliency mapping to further address the need for explainability of models in healthcare settings. Our model identified the left ventricular septum, which is consistent with current studies[26,27], in addition to the mitral plane as pixels significant for prediction. Both imaging features are relevant to the pathophysiology of obstruction, suggesting that our model is grounded in clinical reality.

Our AI model diagnosed LVOT obstruction with high fidelity, complementing existing research to address gaps in detecting HCM. By utilizing A4C videos, our work can inform clinical decisions for further diagnostic assessment for patients, such as provocative measures or other imaging modalities such as cardiac MRI. Our model may be able to assist in screening for LVOT obstruction and thus earlier detection of HCM. With the advent of disease-specific treatments such as mavacamten, our model can help identify and connect patients to more timely and appropriate management. To this end, we have made our code and model weights publicly available for advancement of medical AI research, comparison to existing models, and increased transparency.

Future work can focus on integration of such AI algorithms with ultrasound carts, as additional imaging is needed to interrogate LVOT obstruction. Given the need for real-time feedback for sonographers to initiate further imaging, we anticipate the integration of such an algorithm at the point of scanning for maximal value. By improving disease detection, patients can be connected to disease-specific treatment.

Limitations

Limitations of our study include the recruitment and demographics of our study cohort. Both CSMC and SHC are academic tertiary care centers, introducing biases of patient selection. While the model performed well in patient subgroups by LVEF, the prevalence of comorbidities and mean LVEF differed between the case and control cohorts, suggesting potential residual confounding. Additionally, the limited sample size of cases in the SHC cohort in turn limits sensitivity analysis at this clinical site. While our algorithm was tested across geographically distinct centers, evaluation in additional care centers across the country and in cohorts with a larger number of cases should be performed.

Conclusion

Our deep learning model can detect LVOT obstruction with strong performance across distinct care centers and diverse patient subgroups, thus enhancing early detection of obstruction HCM. Further work involves validation in additional centers and integration of our model with ultrasound carts. Future work also includes developing a multi-view model to detect obstruction and comparing its performance to a single-view model.

References

1. Al Khatib MD Saif Yamin MD Hamzeh Faisal Mohammed MD Own Ma’en Abu Rumman MD Haya Amjad Abida MD Hritvik Jain MBBS Aman Goyal MBBS, S. A. S. M. D. A. E. S. M. D. A. N. Mavacamten in hypertrophic obstructive cardiomyopathy: Prospects for AI integration and mitigating healthcare disparities. https://www.sciencedirect.com/science/article/pii/S0146280624004225#bib0002.

2. Nielsen, S. K. *et al.* Frequency of misdiagnosis in hypertrophic cardiomyopathy. *Eur. Heart J. Qual. Care Clin. Outcomes* qcae031 (2024).

3. Olivotto, I. *et al.* Mavacamten for treatment of symptomatic obstructive hypertrophic cardiomyopathy (EXPLORER-HCM): a randomised, double-blind, placebo-controlled, phase 3 trial. *Lancet* **396**, 759–769 (2020).

4. Maron, B. J. Clinical course and management of hypertrophic cardiomyopathy. *N. Engl. J. Med.* **379**, 655–668 (2018).

5. FDA approves new drug to improve heart function in adults with rare heart condition. *U.S. Food and Drug Administration* https://www.fda.gov/drugs/news-events-human-drugs/fda-approves-new-drug-improve-heart-function-adults-rare-heart-condition (2022).

6. Masri, A. *et al.* Long-term safety and efficacy of mavacamten in symptomatic obstructive hypertrophic cardiomyopathy: Interim results of the PIONEER-OLE study. *J. Am. Heart Assoc.* **13**, e030607 (2024).

7. Maron, B. J. *et al.* Diagnosis and evaluation of hypertrophic cardiomyopathy: JACC state-of-the-art review. *J. Am. Coll. Cardiol.* **79**, 372–389 (2022).

8. Sulaiman, S. A. *et al.* Mavacamten in hypertrophic obstructive cardiomyopathy: Prospects for AI integration and mitigating healthcare disparities. *Curr. Probl. Cardiol.* **49**, 102786 (2024).

9. Maron, M. S., Hellawell, J. L., Lucove, J. C., Farzaneh-Far, R. & Olivotto, I. Occurrence of clinically diagnosed hypertrophic cardiomyopathy in the United States. *Am. J. Cardiol.* **117**, 1651–1654 (2016).

10. Abbasi, M. *et al.* Obstruction in hypertrophic cardiomyopathy: Many faces. *J. Am. Soc. Echocardiogr.* **37**, 613–625 (2024).

11. Naidu, S. S. *et al.* Frequency and clinicoeconomic impact of delays to definitive diagnosis of obstructive hypertrophic cardiomyopathy in the United States. *J. Med. Econ.* **26**, 682–690 (2023).

12. Butzner, M. *et al.* Stable rates of obstructive hypertrophic cardiomyopathy in a contemporary era. *Front. Cardiovasc. Med.* **8**, 765876 (2021).

13. Vaglio, J. C., Jr, Ommen, S. R., Nishimura, R. A., Tajik, A. J. & Gersh, B. J. Clinical characteristics and outcomes of patients with hypertrophic cardiomyopathy with latent obstruction. *Am. Heart J.* **156**, 342–347 (2008).

14. Cheitlin, M. D. Hypertrophic cardiomyopathy is predominantly a disease of left ventricular outflow tract obstruction. *Year B. Cardiol.* **2008**, 446–448 (2008).

15. Duffy, G. *et al.* High-Throughput Precision Phenotyping of Left Ventricular Hypertrophy With Cardiovascular Deep Learning. *JAMA Cardiol* **7**, 386–395 (2022).

16. Ghorbani, A. *et al.* Deep learning interpretation of echocardiograms. *NPJ Digit Med* **3**, 10 (2020).

17. Vrudhula, A. *et al.* High-throughput deep learning detection of mitral regurgitation. *Circulation* **150**, 923–933 (2024).

18. Ouyang, D. *et al.* Video-based AI for beat-to-beat assessment of cardiac function. *Nature* **580**, 252–256 (2020).

19. Holste, G. *et al.* Severe aortic stenosis detection by deep learning applied to echocardiography. *Eur. Heart J.* **44**, 4592–4604 (2023).

20. Long, A. *et al.* Deep learning for echo analysis, tracking, and evaluation of mitral regurgitation (DELINEATE-MR). *Circulation* **150**, 911–922 (2024).

21. Vukadinovic, M. *et al.* Comprehensive echocardiogram evaluation with view primed vision language AI. *Nature* **650**, 970–977 (2026).

22. Holste, G. *et al.* Complete AI-enabled echocardiography interpretation with multitask deep learning. *JAMA* **334**, 306–318 (2025).

23. Ko, W.-Y. *et al.* Detection of hypertrophic cardiomyopathy using a convolutional neural network-enabled electrocardiogram. *J. Am. Coll. Cardiol.* **75**, 722–733 (2020).

24. Lampert, J. *et al.* Calibration of ECG-based deep-learning algorithm scores for patients flagged as high risk for hypertrophic cardiomyopathy. *NEJM AI* **2**, (2025).

25. Tison, G. H., Zhang, J., Delling, F. N. & Deo, R. C. Automated and interpretable patient ECG profiles for disease detection, tracking, and discovery. *Circ. Cardiovasc. Qual. Outcomes* **12**, e005289 (2019).

26. Goto, S. *et al.* Multinational federated learning approach to train ECG and echocardiogram models for hypertrophic cardiomyopathy detection. *Circulation* **146**, 755–769 (2022).

27. Vemulapalli, S. *et al.* Machine learning computer vision point of care decision support of echocardiographic identification of hypertrophic cardiomyopathy. *JACC Adv.* **4**, 101746 (2025).

28. Tran, D. *et al.* A Closer Look at Spatiotemporal Convolutions for Action Recognition. *arXiv [cs.CV]* (2017).

29. Selvaraju, R. R. *et al.* Grad-CAM: Visual explanations from deep networks via Gradient-based localization. *arXiv [cs.CV]* (2016).

30. Writing Committee Members *et al.* 2020 AHA/ACC guideline for the diagnosis and treatment of patients with hypertrophic cardiomyopathy: A report of the American College of Cardiology/ American Heart Association Joint Committee on Clinical Practice Guidelines. *J. Thorac. Cardiovasc. Surg.* **162**, e23–e106 (2021).

31. Alimadadi, A. *et al.* Machine learning-based classification and diagnosis of clinical cardiomyopathies. *Physiol. Genomics* **52**, 391–400 (2020).

32. Antonopoulos, A. S. *et al.* Machine learning of native T1 mapping radiomics for classification of hypertrophic cardiomyopathy phenotypes. *Sci. Rep.* **11**, 23596 (2021).

33. Lyon, A. *et al.* Distinct ECG phenotypes identified in hypertrophic cardiomyopathy using machine learning associate with arrhythmic risk markers. *Front. Physiol.* **9**, 213 (2018).

34. Smole, T. *et al.* A machine learning-based risk stratification model for ventricular tachycardia and heart failure in hypertrophic cardiomyopathy. *Comput. Biol. Med.* **135**, 104648 (2021).

35. Zhao, K. *et al.* Machine learning in hypertrophic cardiomyopathy: Nonlinear model from clinical and CMR features predicting cardiovascular events. *JACC Cardiovasc. Imaging* **17**, 880–893 (2024).

36. Park, J. *et al.* Single-view echocardiographic analysis for left ventricular outflow tract obstruction prediction in hypertrophic cardiomyopathy: A deep learning approach. *J. Am. Soc. Echocardiogr.* **38**, 1115–1126 (2025).

Figure 1. Model Interpretability with Saliency Mapping Saliency maps with Gradient-weighted Class Activation Mapping (Grad-CAM) for cases from the CSMC cohort, with yellow pixels being more significant for model prediction and darker purple pixels being less important. Created in BioRender. Yuan, V. (2026) https://BioRender.com/p216ki7

Central Illustration. Deep learning algorithm to detect LVOT obstruction from A4C with AUC, sensitivity, and specificity (top) Representative apical 4-chamber images for selected DL-predicted cases (middle) and controls (bottom). Created in BioRender. Yuan, V. (2026) https://BioRender.com/4sbs3ze

Table 1. Demographic information and comorbidities of study cohorts at Cedars-Sinai Medical Center (CSMC), Stanford Healthcare (SHC), and Kaiser Permanente Northern California (KPNC)

	CSMC Controls (n = 4,859)	**CSMC Cases (n = 2,396)**	**SHC Controls (n = 362)**	**SHC Cases (n = 4)**	**KPNC Controls (n = 591)**	**KPNC Cases (n= 297)**
Age	68.7±14.6	68.9±15.3	62.5±16.5	67.0±15.4	68.0±13.4	68.5±13.8

Sex						
Male	2,170 (44.8%)	1,355 (56.6%)	218 (60.2%)	1 (25.0%)	230 (38.9%)	115 (38.7%)
Female	2,671 (55.2%)	1,041 (43.4%)	144 (39.8%)	3 (75.0%)	361 (61.0%)	182 (61.3%)
Race and Ethnicity						
American Indian or Alaska Native	9 (0.19%)	5 (0.21%)	0 (0.0%)	0 (0.0%)	66 (11.2%)	23 (7.7%)
Asian	339 (7.0%)	193 (8.1%)	56 (15.5%)	0 (0.0%)	88 (14.9%)	58 (19.5%)
Black or African American	919 (19.0%)	330 (13.8%)	36 (9.9%)	0 (0.0%)	111 (18.8%)	31 (10.4%)
Native Hawaiian/Pacific Islander	11 (0.23%)	5 (0.21%)	3 (0.8%)	0 (0.0%)	4 (0.7%)	0 (0.0%)
Hispanic	587 (12.1%)	588 (12.1%)	40 (11.0%)	1 (25.0%)	0 (0.0%)	0 (0.0%)
Non-Hispanic	4,149 (85.7%)	4165 (85.7%)	296 (81.8%)	2 (50.0%)	591 (100.0%)	297 (100.0%)
White	3,139 (64.8%)	1,639 (68.4%)	179 (49.4%)	2 (50.0%)	305 (51.6%)	182 (61.2%)
Other	319 (6.6%)	168 (7.0%)	58 (16.0%)	1 (25%)	1 (0.2%)	1 (0.3%)
Unknown	107 (2.2%)	56 (2.3%)	24 (6.6%)	1 (25%)	6 (1.0%)	3 (1.0%)
Left ventricular ejection fraction %	58.1±13.6	70.4±8.76	59.5±12.8	69.3±5.8	58.5±11.7	66.6±7.8
Left ventricular end-diastolic volume index (mL/m^2)	43.6±21.3	36.5±18.7	43.9±19.2	40.6±4.3	47.0±18.0	46.4±16.2
Left ventricular end-systolic volume index (mL/m^2)	19.9±15.5	11.6±6.95	19.6±13.2	12.5±2.7	20.5±13.1	15.5±6.9
Left atrial dilation (LAV_i > 34 mL/m^2)	182 (10.0%)	143 (9.6%)	275 (76.0%)	3 (75.0%)	88 (14.9%)	70 (23.6%)
Diastolic interventricular septal thickness (mm)	1.34±0.39	1.43±0.44	1.76±0.43	1.31±0.37	1.44±0.40	1.53±0.41

Implanted cardioverter-defibrillator	291 (6.0%)	63 (2.6%)	77 (21.3%)	0 (0.0%)	31 (5.3%)	26 (8.8%)
Mitral regurgitation	404 (8.3%)	153 (6.4%)	240 (66.3%)	2 (50.0%)	20 (3.4%)	44 (15.1%)
Coronary artery disease	1,288 (26.5%)	434 (18.1%)	115 (31.8%)	2 (50.0%)	204 (51.5%)	139 (46.8%)
Atrial fibrillation	1,362 (28.0%)	428 (17.9%)	164 (45.3%)	3 (75.0%)	229 (38.8%)	105 (35.4%)
Hypertension	1,131 (23.3%)	430 (18.1%)	205 (56.6%)	3 (75.0%)	525 (89.0%)	236 (79.5%)
Diabetes mellitus	1,319 (27.1%)	482 (20.1%)	119 (32.9%)	1 (25.0%)	28 (4.7%)	6 (2.0%)

Table 2. Demographic information and comorbidities of train, validation, and test sets

	Train (n=30,565)	**Validation (n=3,730)**	**Test (n=3,848)**
Cases	13,247 (42.7%)	1,610 (42.2%)	1,639 (42.1%)
Controls	17,782 (57.3%)	2,209 (57.8%)	2,257 (57.9%)
Age	68.615.1	68.716.2	68.014.8
Sex			
Male	13,738 (44.3%)	1,634 (42.8%)	1,789 (45.9%)
Female	17,102 (55.1%)	2,217 (55.7%)	2,092 (53.7%)
Race			
American Indian or Alaska Native	67 (0.2%)	0 (0.0%)	4 (0.1%)
Asian	2,162 (7.0%)	276 (7.2%)	329 (8.4%)
Black or African American	5,401(17.4%)	560 (14.7%)	538 (13.8%)
Native Hawaiian/Pacific Islander	105 (0.3%)	8 (0.2%)	0 (0%)
White	20,380 (65.7%)	2,599 (68.1%)	2,713 (69.6%)
Other	2,205 (7.1%)	236 (6.2%)	212 (5.4%)
Unknown	709 (2.3%)	130 (3.4%)	100 (2.6%)

Left ventricular ejection fraction (%)	62.8%±13.4%	64.0%12.3%	62.613.3%
Left ventricular end-diastolic volume index (mL/m^2)	40.6±19.3	41.3±20.8	43.3±30.1
Left ventricular end-systolic volume index (mL/m^2)	16.8±13.6	16.3±12.3	17.9±14.5
Left atrial dilation (LAV_i > 34 mL/m^2)	1,152 (9.7%)	132 (8.2%)	147 (9.4%)
Diastolic interventricular septal thickness (mm)	1.40±0.47	1.40.43	1.360.41
Implanted cardioverter-defibrillator	1,798 (6.2%)	235 (6.6%)	183 (5.0%)
Mitral regurgitation	2,749 (9.5%)	267 (7.4%)	403 (11.0%)
Coronary artery disease	7,440 (25.6%)	947 (26.4%)	860 (23.4%)
Atrial fibrillation	7,854 (27.0%)	912 (25.4%)	1022 (27.8%)
Hypertension	7.212 (24.8%)	812 (22.7%)	782 (21.3%)
Diabetes mellitus	7,468 (25.7%)	839 (23.4%)	979 (26.7%)

Table 3. AUC, sensitivity, specificity, PPV, NPV, and positive and negative likelihood ratios of deep learning obstruction in the CSMC, SHC, and KPNC patient cohorts with 95% confidence intervals in parentheses

	CSMC	**SHC**	**KPNC**
AUC	0.858 [0.847, 0.870]	0.836 [0.740, 0.922]	0.817 [0.807, 0.827]
Sensitivity	63.0% [60.6%, 65.3%]	100.0% [100.0%, 100.0%]	86.8% [85.4%, 88.1%]
Specificity	88.5% [87.2%, 89.8%]	57.0% [51.9%, 62.0%]	55.2% [53.7%, 56.6%]
PPV	0.80 [0.776, 0.820]	0.03 [0.006, 0.052]	0.630 [0.612, 0.647]
NPV	0.77 [0.752, 0.783]	1.0 [1.0, 1.0]	0.825 [0.813, 0.836]
LR+	5.47 [4.87, 6.21]	2.32 [2.07, 2.63]	3.03 [2.86, 3.22]

LR-	0.42 [0.161, 0.205]	0.0 [0.0, 0.0]	0.378 [0.311, 0.350]