Published in final edited form as:

Circulation. 2024 September 17; 150(12): 923–933. doi:10.1161/CIRCULATIONAHA.124.069047.

High Throughput Deep Learning Detection of Mitral Regurgitation

Amey Vrudhula, M.D.[a,b], **Grant Duffy, B.S.**[a], **Milos Vukadinovic, B.S.**[a,c], **David Liang, M.D., Ph.D.**[d], **Susan Cheng, M.D., M.M.Sc., M.P.H.**[a], **David Ouyang, M.D.**[a,e]

[a)]Department of Cardiology, Smidt Heart Institute, Cedars-Sinai Medical Center, Los Angeles, CA

[b)]Icahn School of Medicine at Mt Sinai, New York, NY

[c)]Department of Bioengineering, University of California Los Angeles, Los Angeles, CA

[d)]Department of Medicine, Division of Cardiology, Stanford University, Palo Alto, CA

[e)]Division of Artificial Intelligence in Medicine, Cedars-Sinai Medical Center, Los Angeles, CA

Abstract

Background: Diagnosis of mitral regurgitation (MR) requires careful evaluation by echocardiography with Doppler imaging. This study presents the development and validation of a fully automated deep learning pipeline for identifying apical-4-chamber view videos with color Doppler echocardiography and detection of clinically significant (moderate or severe) MR from transthoracic echocardiograms.

Methods: A total of 58,614 transthoracic echocardiograms (2,587,538 videos) from Cedars-Sinai Medical Center (CSMC) were used to develop and test an automated pipeline to identify apical-4-chamber view videos with color Doppler across the mitral valve and then assess MR severity. The model was tested internally on a test set of 1,800 studies (80,833 videos) from CSMC and externally evaluated in a geographically distinct cohort of 915 studies (46,890 videos) from Stanford Healthcare (SHC).

Results: In the held-out CSMC test set, the view classifier demonstrated an area under the curve (AUC) of 0.998 (0.998 – 0.999) and correctly identified 3,452 of 3,539 echocardiography videos as having color Doppler information across the mitral valve (sensitivity of 0.975 (0.968–0.982) and specificity of 0.999 (0.999–0.999) compared with manually curated videos). In the external test cohort from SHC, the view classifier correctly identified 1,051 of 1,055 manually curated videos with color Doppler information across the mitral valve (sensitivity of 0.996 (0.990 – 1.000) and specificity of 0.999 (0.999 – 0.999)). In the CSMC test cohort, MR moderate or greater in

Correspondence: David Ouyang M.D., Assistant Professor, Smidt Heart Institute, Cedars-Sinai Medical Center, 127 S. San Vicente Blvd., AHSP Pavilion Suite A3100, Los Angeles, CA 90048, Fax: 310.423.0127, David.Ouyang@cshs.org, Twitter Handle: @David_Ouyang.

Supplemental Materials List
Expanded Methods
Tables S1–7
Figures S1–2
Videos S1–4
CONSORT AI Checklist

severity was detected with an AUC of 0.916 (0.899 – 0.932) and severe MR was detected with an AUC of 0.934 (0.913 – 0.953). In the SHC test cohort, the model detected MR moderate or greater in severity with an AUC of 0.951 (0.924 – 0.973) and severe MR with an AUC of 0.969 (0.946 – 0.987).

Conclusions: In this study, a novel automated pipeline for identifying clinically significant MR from full transthoracic echocardiography studies demonstrated excellent performance across high numbers of studies and across multiple institutions. Such an approach has the potential for automated screening and surveillence of MR.

Keywords

Mitral Regurgitation; Artificial Intelligence; Deep Learning; Echocardiography

Introduction

Mitral regurgitation (MR) is one of the most common forms of valvular heart disease, affecting more than 4 million Americans.[1,2,3,4] Often progressing insidiously and frequently underrecognized[4], both primary MR as well as secondary MR can be initially asymptomatic but lead to worsening heart failure and mortality.[5–8,2,9] There has been an increased focus on early MR diagnosis given advances in surgical and transcatheter treatment options[5,10,11]. Echocardiography with color Doppler is the most common method of initial evaluation of MR, with a holistic assessment combining left atrial size, effective regurgitant orifice area, regurgitant fraction, regurgitant volume, as well as other key clinical factors to accurately assess disease severity.[12,13] Despite ultrasound technology becoming more widely available, accurate assessment of MR still requires experienced expert image acquisition and evaluation.

Recent advances in machine learning offer opportunities to automate time-consuming steps in the interpretation of medical imaging. Artificial intelligence (AI) has the ability to precisely phenotype subtle cardiac physiology as well as identify imaging features of disease severity not recognized by clinicians.[14–17] Deep learning has been applied to echocardiography to improve the precision of common measurements, such as left ventricular ejection fraction[14] and wall thickness[16,18], as well as streamlining assessment of aortic stenosis[19], hypertrophic cardiomyopathy[16], and cardiac amyloidosis.[16,20] With increased ultrasound availability, AI guidance has been developed for both image acquisition and interpretation[14,21]. Considering the increasing prevalence of MR in an aging population with co-morbid heart failure, AI could aid in MR screening and surveillance.[22–26]

In this study, we developed and evaluated a deep learning pipeline's performance in automating identification of MR from standard transthoracic echocardiogram studies. We hypothesized that a deep learning approach can identify and assess MR severity on color Doppler apical-4-chamber (A4C) view videos with high-throughput automation. The automated pipeline was evaluated in two geographically distinct cohorts (Figure 1). Such an approach can be used for serial surveillance and screening of MR with a focus on changes over time.

Methods

Study Population and Data Source

Cedars-Sinai Medical Center (CSMC) Cohort: A total of 58,614 transthoracic echocardiograms from 38,461 patients receiving care at Cedars-Sinai Medical Center (CSMC) between October 11, 2011 and June 04, 2022 were used to train and evaluate the deep learning models. A total of 2,587,538 videos (an average of 44 videos per study after excluding still images) were initially sourced from Digital Imaging and Communications in Medicine (DICOM) files and underwent de-identification, view classification, and pre-processing into AVI videos (as described in Supplementary Material). A total of 354,117 videos were classified as A4C view videos using an automated view classifier, and then manually curated to identify 34,714 videos with color Doppler across the mitral valve.[27]

Following view selection, the CSMC cohort included 34,714 videos across 30,453 unique echocardiograms from 22,661 patients, and a subset enriched for moderate and severe MR were used for training. All examples of moderate and severe MR were used for training, while mild MR and control videos were sampled to be approximately the same number as moderate MR videos to mitigate class imbalance when training the MR severity model. A total of 20,604 videos from 18,133 studies in the dataset were randomly split on a patient level into train (80%), validation (10%), and test (10%) cohorts to train a deep neural network for MR severity classification (Figure 2). MR severity for each study was determined based on the clinical echocardiogram report determined in a high volume echocardiography laboratory in accordance with American Society of Echocardiography guidelines.[28] When MR was characterized as an intermediate category ("trace to mild" or "mild to moderate" or "moderate to severe"), videos were placed in the more severe categories. For example, "mild to moderate MR" was considered moderate MR and "moderate to severe MR" was considered severe MR for model training. Both primary and secondary MR were included. Studies with concomitant mitral stenosis, prosthetic valves, and reduced as well as normal left ventricular ejection fraction were also included in both training and validation datasets. For subgroup analysis of characteristics not frequently present in the test cohort, additional validation cohorts enriched for prior mitral valve intervention, concomitant moderate/severe mitral stenosis, cases of eccentric MR and cases with clinician quantification were obtained and used for additional testing. Patients from the MR severity model training and validation cohort were excluded from all test cohorts.

Stanford Healthcare (SHC) Cohort: The pipeline was evaluated on 915 studies (containing a total of 46,890 videos) from SHC's high-volume academic echocardiography laboratory. The automated view classification pipeline was compared with manual curation of videos within those studies to evaluate specificity. All videos identified by the view classifier were used for downstream MR severity model validation. Model output was compared with MR severity determined by expert cardiologists from the clinical reports. This study was approved by the Institutional Review Boards (IRBs) at Cedars-Sinai Medical Center and Stanford Healthcare. The need for informed consent was waived as the study involved secondary analysis of existing data.

AI Model Training

Deep learning models were trained using the PyTorch Lightning deep learning framework. When patients had multiple echocardiograms, each video was considered an independent example during training, with care not to have patient overlap across training, validation, and test cohorts. Video-based convolutional neural networks (R2+1D) were used for view classification and MR severity assessment.[29] This model architecture was previously used for other echocardiography tasks and shown to be effective.[18] The models were initialized with random weights and trained using a binary cross entropy loss function for up to 100 epochs, using an ADAM optimizer, an initial learning rate of 1e-2, and a batch size of 24 on two NVIDIA RTX 3090 GPUs. Early stopping was performed based on the validation loss.

The view classifier was trained using the 34,714 manually curated videos with color Doppler across the mitral valve as cases and 49,263 other apical-4-chamber videos as controls. Controls were a combination of videos that did not have color Doppler or had color Doppler window not focused on the mitral valve (videos focused on the tricuspid valve, intra-atrial septum, or ventricular septum). The MR severity model was trained on 6,206 videos without MR, 6,128 videos with mild MR, 6,174 videos with moderate MR and 2,042 videos with severe MR. This process is summarized in Figure 2.

Statistical Analysis

Model performance was evaluated using area under the receiver operating characteristic curve (AUROC) and confusion matrices. F1-score, recall (sensitivity), positive predictive value (PPV), and negative predictive value (NPV) were also evaluated for both greater than moderate MR and severe MR. During external validation, the view classifier and MR classifier were evaluated serially as an automated pipeline. Statistical analysis was performed in Python (version 3.8.0) and R (version 4.2.2). Confidence intervals were computed via bootstrapping with 10,000 samples. Reporting of study results is consistent with guidelines put forth by CONSORT-AI (Supplementary Material).[30,31]

Subgroup analysis was conducted to assess model performance in patients with different etiologies of MR, different ranges of left ventricular ejection fraction, study characteristics, associated co-morbidities and other clinical characteristics. Primary MR was identified by study reports with findings predominantly highlighting mitral valve prolapse, rheumatic changes of the mitral valve, flail leaflet, or eccentric MR jet. Secondary MR was defined as patients with moderately or severely reduced left ventricular ejection fraction and no identifiers of primary MR were present. All other videos were considered to be of unclear or mixed MR etiology. Clinical characteristics were obtained from the electronic health record or associated echocardiography report. Echocardiogram study quality was determined by clinicians and extracted from the clinical report. Studies where clinicians commented on technical difficulty, poor study quality, or where one or more major cardiac structures (left ventricle, right ventricle, pulmonary artery, etc.) were not well visualized were classified as technically difficult.

Model Explainability

The key imaging features identified by the MR severity model were evaluated using saliency mapping generated using the Integrated Gradients method.[32] This method generated a heatmap for every frame of the video, summarized as a final 2-dimensional heatmap generated by using the maximum value along the temporal axis for each pixel location in the video. Pixels brighter in intensity and closer to yellow were more salient to model predictions, while those darker in color were less important to the model's final prediction. When assessing videos with no MR, heatmaps were obtained by taking the maximum of saliency maps for the moderate and severe class output neurons for each pixel location (Figure 4). Finally, correlation between model predictions and quantitative metrics of MR severity was examined in cases with moderate/severe MR showing model prediction severity correlated with more extreme metrics (Table S6)."

Serial Monitoring of MR Severity

The performance of the MR severity model in monitoring patients longitudinally and detecting changes in MR across serial studies was evaluated. All patients in the test cohort with more than one echo with a mitral A4C Doppler video were included in the analysis. This cohort consisted of 3,538 echocardiograms across 768 patients, with patients having anywhere from 2 to 25 serial echocardiogram studies. Predicted MR severity across serial studies was compared to cardiologist determined MR severity across serial studies. Sankey plots showcasing MR severity across time in cardiologist determined severity and model predicted severity across serial studies were compared (Figure 5).

Data and Code Availability

The dataset of videos and reports used to train EchoNet-MR is not publicly available due to its potentially identifiable nature; it is available for use with approval by the CSMC IRB. Our code and model weights are available at https://github.com/echonet/MR

Results:

Study Population

A total of 58,614 transthoracic echocardiograms from 38,461 patients were used to train the deep learning pipeline. From 2,587,538 initial videos, a total of 354,117 videos were identified as apical-4-chamber view and subsequently manually curated to identify videos that had color Doppler across the mitral valve. The manually curated color Doppler videos were used to train a view-classification model and linked with clinician reports to train the MR severity model. Patient characteristics are presented in Tables 1 & 2 and are representative of the general CSMC and SHC patient populations that received echocardiograms. The data were split on patient level for training and validation and had similar patient age, left ventricular ejection fraction, left atrial volume index and proportions of male sex, coronary artery disease, and atrial fibrillation.

View Classifier Performance Across Two Institutions

On a test set of 3,109 transthoracic echocardiograms (132,767 videos) from CSMC not seen during model training, the view classifier identified 3,452 videos (97.5% of manually identified cases). This corresponds to an AUC of 0.998 (0.998 – 0.999), and at the Youden Index, with a sensitivity of 0.975 (0.968–0.982) and specificity of 0.999 (0.999–0.999). To evaluate generalization of the view classification model at a geographically distinct site, we evaluated its performance on 915 studies from SHC. The view classifier isolated 1,091 videos from a total of 46,890 videos, while manual review identified 1,055 videos with color Doppler across the mitral valve. The view classifier correctly identified 1,051 (99.6%) of manually curated videos, with 4 videos not found by the AI pipeline and 40 false positives. This corresponds to a sensitivity of 0.996 (0.990 – 1.000) and specificity of 0.999 (0.999 – 0.999).

Mitral Regurgitation Severity Performance Across Two Institutions

The MR severity model showed strong performance in distinguishing MR severity and identifying clinically significant MR (Figure 3). In the internal CSMC test set not used during model training, the model demonstrated an AUC of 0.916 (0.899 – 0.932) in detecting at least moderate MR and an AUC of 0.934 (0.913 – 0.953) for severe MR. The AI model had an NPV of 0.954 (0.940 – 0.967) for severe MR and an NPV of 0.863 (0.835 – 0.890) for at least moderate MR. Further information on MR model performance is presented in Table 3. The MR severity model performance was similar across institutions. In the SHC cohort, the model identified severe MR with an AUC of 0.969 (0.946 – 0.987) and moderate/severe MR (defined as at least moderate MR) with an AUC of 0.951 (0.924 – 0.973). In this cohort, the model had an NPV of 0.977 (0.962 – 0.990) for severe MR and an NPV of 0.986 (0.974 – 0.995) for ≥ moderate MR. The potential for serial surviellence and evaluation of change over time is shown in the CSMC studies among patients with four or more echocardiogram studies, showing progression of MR over time in a select population identified by the AI model similar to rate of progression identified by cardiologists (Figure 5).

Subgroup Analysis

The MR severity model showed strong performance across test set subgroups (Table 4). Moderate or more severe primary MR was detected with an AUC of 0.913 (0.877 – 0.943), and severe primary MR was detected with an AUC of 0.982 (0.970 – 0.992). Moderate or more severe secondary MR was detected with AUC of 0.952 (0.930– 0.970), and severe secondary MR was detected with AUC of 0.975 (0.956 – 0.990). Similar performance was observed for moderate or more severe MR of unclear/mixed etiology (AUC of 0.903 (0.881–0.923)) and severe MR of unclear/mixed etiology (AUC of 0.939 (0.909–0.965)) of unclear etiology. The model performed well across the range of LVEF, pre-existing atrial fibrillation, mitral annular calcification, left atrial dilation, prior mitral valve intervention (surgical repair or surgical replacement), or concomitant mitral stenosis or aortic valve disease. AUCs for moderate/severe MR ranged from 0.850 (0.807 – 0.973) to 0.899 (0.816 – 0.964) in these groups, while AUCs for severe MR ranged from 0.868 (0.808 – 0.920) to 0.914 (0.871

– 0.949). AUCs were similar irrespective of study quality and acquisition date as well as eccentric MR jets.

Comparison with Cardiac Magnetic Resonance Imaging

We identified an additional held-out test population of 1,085 videos from 831 transthoracic echocardiograms across 569 CSMC patients who underwent cardiac magnetic resonance imaging (MRI) and echocardiography within 180 days. The MRI assessment of MR severity was then used to evaluate our deep learning model. CMR was performed in patients with varying degrees of MR, and studies where patients had moderate or severe MR based on echocardiographic imaging comprised 146 studies (17.5%) in the cohort (Table S2). Of these echocardiograms with moderate or severe MR, 23.2% exhibited primary MR, 33.1 % had secondary MR, and 43.7% had MR of unclear etiology, showing no clear bias to a given etiology (Table S3). In this cohort, our AI model had an AUC of 0.805 (0.717–0.884) in detecting at least moderate MR and an AUC of 0.828 (0.688–0.940) for severe MR (Table S4). When comparing the cardiologist's echocardiography-based assessment of MR severity with MRI-based assessment, cardiologists had an AUC of 0.833 (0.586–0.937) in detecting at least moderate MR and an AUC of 0.791 (0.753–0.902) for severe MR (Table S5). This difference in AUC between AI model and cardiologist when compared with MRI was not significant for at least moderate MR (DeLong test, $p = 0.295$) or severe MR (DeLong test, $p = 0.556$).

Model Interpretation

Notably, saliency maps for our model demonstrate that the model focuses on the clinically relevant imaging features of MR. Saliency maps from Integrated Gradients were used to identify regions of interest in each video contributing the most to detection of MR severity (Figure 4).[32] These interpretability techniques demonstrated localization of the activation signal in the color Doppler window and primarily highlighting the MR jet, indicating that the model used appropriate, physiologic features of MR to make predictions. Frame-by-frame saliency visualizations are shown in Supplementary Videos S1–S4. When comparing of model predictions with quantitative MR metrics, predicted severe MR cases had more extreme measurements than moderate MR cases across all quantitative metrics. (Table S6)

Discussion

We developed and validated an automated pipeline for detection and assessment of MR severity with echocardiography. From a full transthoracic echocardiogram, the algorithm automatically identifies appropriate A4C videos with color Doppler on the mitral valve and then assesses MR severity. For both severe MR as well as at least moderate MR, the model demonstrated strong performance (> 0.916 AUC and > 0.863 NPV). This automated workflow worked in unselected external validation studies without preselection or exclusion of other concomitant comorbidities and performed similarly for both primary and secondary MR.

Our algorithm learns features of MR that generalize across variability in imaging practices in two geographically distinct sites. Many prior echocardiography AI models have

focused on standard 2-dimensional black-and-white B-mode images and videos[14,16,19,33]. Meanwhile, our study focuses on the AI assessment of color Doppler videos and utilized a video-based model for the incorporation of rich temporal and Doppler information, both crucial for accurate MR assessment. Incorporation of color Doppler information greatly expands the opportunities for AI in echo, particularly in valve disease. Previous works have applied deep learning to color Doppler echocardiography videos to assess valvular heart disease and rheumatic heart disease.[34,35] Similar to image based B-mode models mentioned previously, these approaches rely on image-based deep learning models for frame selection, localization of the left atrium, and jet segmentation to yield a probability of mitral regurgitation from either A4C video frames or a combination of A4C and parasternal long-axis videos. Meanwhile, the current approach relies on a state-of-the-art video based architecture to directly detect and stratify MR from A4C videos, which are automatically isolated from entire echocardiogram studies. This novel approach allows for the use of information across multiple frames while also reducing the need for multiple deep learning models for intermediate steps (frame selection, jet segmentation, etc.). The present work also includes more representative training data (including both mild and moderate cases as well as severe MR cases during training) and utilized more training data (more than ten thousand videos of MR used for training in our model compared to 777 and 282 training cases respectively in prior work).

For comprehensive expert clinical evaluation of MR, a variety of views, including color Doppler, as well as Doppler metrics and calculations are synthesized together for a holistic assessment of MR severity. In this work, our current AI model focuses on the color Doppler videos in the A4C view, which might be insufficient to visualize eccentric or off axis MR jets. Despite this limitation, our AI algorithm results in concordant interpretations with the comprehensive clinical approach, suggesting there is overlapping information or ancillary information such as left atrial size or mitral valve motion that can assist an AI model in evaluating MR severity even when there is a limited MR jet. Similarly, in clinical practice, cardiologists often integrate conflicting calculations of MR severity, and a provider's clinical gestalt for a particular severity of MR might correlate with hidden findings in the apical-4-chamber view as well as jet area on color Doppler. Regardless, this clinical context suggests that future work synthesizing information across multiple echocardiographic views will only improve the performance of future MR models.

The present work builds upon prior work in the space of echocardiography and AI. Several recent works have reported strides in computer vision and echocardiography, including automated view classification[36,37], phenotyping of left ventricular hypertrophy[16], assessment of LV systolic function[14], aortic stenosis risk stratification, and detection of complex congenital heart defects.[38] Prior work in machine learning applied to MR has primarily focused on structured data and non-deep learning approaches. The combination of our algorithm with previously published works using AI to guide novices in acquiring imaging could potentially increase access to screening of MR.[21,39] Future work could focus on automatically quantifying parameters like valve leaflet thickness, effective regurgitant orifice area, regurgitant volume and fraction.

In summary, we introduce a model to screen for and stratify MR severity from transthoracic echocardiogram videos. To do so, we provide a workflow for isolating mitral valve color Doppler videos and automation of MR severity assessment. The models were evaluated to have good performance in internal and external test cohorts. The use of such a deep learning pipeline, with a high AUC, NPV, and generalizability across sites, could aid in the preliminary assessment of MR, facilitate review of institutional databases, or expand access for screening in primary care or low-resource settings.

Study Limitations

While promising, the present work carries limitations. Echocardiographic assessment of MR depends on appropriate images being obtained, with different views potentially maximizing the visualized MR jet. This algorithm would not overcome incomplete input information and insufficient images that would result in underestimation of MR. The model was trained on clinical assessment of MR, which did not always have comprehensive quantitative assessment (Table S7), although model output correlated with quantitative assessments when available. When compared with MRI assessment of MR severity, our AI model reduced in performance compared to the comparison with echocardiographic assessment of MR severity. In parallel, the cardiologists' assessment of echocardiographic MR severity was no better in concordance with MRI findings. Given the recognized discordance between clinician assessment of MR severity across modalities[40] and the fact that our AI model was trained on labels from cardiologists' echocardiographic assessments, the modest reduction in performance when compared across modalities should be expected.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements:

A.V. acknowledges support from the Sarnoff Cardiovascular Research Award. S.C. acknowledges support from the Erika J Glazer Family Foundation.

Funding Sources:

At the time of this work, A.V. was a research fellow supported by the Sarnoff Cardiovascular Research Award. S.C. acknowledges support from the Erika J Glazer Family Foundation. D.O. reports research grants R00 HL157421 and R01HL173526 and support from AstraZeneca Alexion.

Disclosures:

A.V., G.D., M.V., D.L., and D.S.B. report no disclosures. A.C.K. reports consulting fees from InVision. S.C. reports consulting fees from UCB and Viz.ai and research grants NIH R01-HL131532 and NIH R01-HL142983. D.O. reports research grants R00 HL157421 and R01HL173526 and support from AstraZeneca Alexion, as well as consulting from EchoIQ, Ultromics, Pfizer, and InVision.

Non-standard Abbreviations:

MR	mitral regurgitation
TTE	trans-thoracic echocardiogram

A4C	apical-4-chamber
SHC	(Stanford Healthcare)
CSMC	Cedars-Sinai Medical Center
AUC	area under the curve
AUROC	area under the receiver operating characteristic curve

References

1. Yadgir S, Johnson CO, Aboyans V, Adebayo OM, Adedoyin RA, Afarideh M, Alahdab F, Alashi A, Alipour V, Arabloo J, Azari S, Barthelemy CM, Benziger CP, Berman AE, Bijani A, Carrero JJ, Carvalho F, Daryani A, Durães AR, Esteghamati A, Farid TA, Farzadfar F, Fernandes E, Filip I, Gad MM, Hamidi S, Hay SI, Ilesanmi OS, Naghibi Irvani SS, Jürisson M, Kasaeian A, Kengne AP, Khan AR, Kisa A, Kisa S, Kolte D, Manafi N, Manafi A, Mensah GA, Mirrakhimov EM, Mohammad Y, Mokdad AH, Negoi RI, Thi Nguyen HL, Nguyen TH, Nixon MR, Otto CM, Patel S, Pilgrim T, Radfar A, Rawaf DL, Rawaf S, Rawasia WF, Rezapour A, Roever L, Saad AM, Saadatagah S, Senthilkumaran S, Sliwa K, Tesfay BE, Tran BX, Ullah I, Vaduganathan M, Vasankari TJ, Wolfe CDA, Yonemoto N, Roth GA, Global Burden of Disease Study 2017 Nonrheumatic Valve Disease Collaborators. Global, regional, and national burden of calcific aortic valve and degenerative mitral valve diseases, 1990–2017. Circulation. 2020;141:1670–1680. [PubMed: 32223336]
2. Messika-Zeitoun D, Candolfi P, Vahanian A, Chan V, Burwash IG, Philippon J-F, Toussaint J-M, Verta P, Feldman TE, Iung B, Glineur D, Mesana T, Enriquez-Sarano M. Dismal Outcomes and High Societal Burden of Mitral Valve Regurgitation in France in the Recent Era: A Nationwide Perspective. J Am Heart Assoc. 2020;9:e016086. [PubMed: 32696692]
3. Harb SC, Griffin BP. Mitral Valve Disease: a Comprehensive Review. Curr Cardiol Rep. 2017;19:73. [PubMed: 28688022]
4. Dziadzko V, Clavel M-A, Dziadzko M, Medina-Inojosa JR, Michelena H, Maalouf J, Nkomo V, Thapa P, Enriquez-Sarano M. Outcome and undertreatment of mitral regurgitation: a community cohort study. Lancet. 2018;391:960–969. [PubMed: 29536860]
5. Enriquez-Sarano M, Akins CW, Vahanian A. Mitral regurgitation. Lancet. 2009;373:1382–1394. [PubMed: 19356795]
6. Otto CM, Verrier ED. Mitral regurgitation--what is best for my patient? N. Engl. J. Med 2011;364:1462–1463. [PubMed: 21463152]
7. Del Forno B, De Bonis M, Agricola E, Melillo F, Schiavi D, Castiglioni A, Montorfano M, Alfieri O. Mitral valve regurgitation: a disease with a wide spectrum of therapeutic options. Nat Rev Cardiol. 2020;17:807–827. [PubMed: 32601465]
8. Ocher R, May M, Labin J, Shah J, Horwich T, Watson KE, Yang EH, Press Calfon, Marcella A. Mitral Regurgitation in Female Patients: Sex Differences and Disparities. Catheter Cardiovasc Interv. 2023;2:101032.
9. Simpson TF, Kumar K, Samhan A, Khan O, Khan K, Strehler K, Fishbein S, Wagner L, Sotelo M, Chadderdon S, Golwala H, Zahr F. Clinical Predictors of Mortality in Patients with Moderate to Severe Mitral Regurgitation. Am J Med. 2022;135:380–385.e3. [PubMed: 34648779]
10. Tribouilloy CM, Enriquez-Sarano M, Schaff HV, Orszulak TA, Bailey KR, Tajik AJ, Frye RL. Impact of preoperative symptoms on survival after surgical correction of organic mitral regurgitation: rationale for optimizing surgical indications. Circulation. 1999;99:400–405. [PubMed: 9918527]
11. David TE, Ivanov J, Armstrong S, Rakowski H. Late outcomes of mitral valve repair for floppy valves: Implications for asymptomatic patients. J Thorac Cardiovasc Surg. 2003;125:1143–1152. [PubMed: 12771888]
12. Fadel BM, Bakarman H, Dahdouh Z, Di Salvo G, Mohty D. Spectral Doppler interrogation of mitral regurgitation -spot diagnosis. Echocardiography. 2015;32:1179–1183. [PubMed: 25611451]

13. Hagendorff A, Knebel F, Helfen A, Stöbe S, Haghi D, Ruf T, Lavall D, Knierim J, Altiok E, Brandt R, Merke N, Ewen S. Echocardiographic assessment of mitral regurgitation: discussion of practical and methodologic aspects of severity quantification to improve diagnostic conclusiveness. Clin Res Cardiol. 2021;110:1704–1733. [PubMed: 33839933]

14. Ouyang D, He B, Ghorbani A, Yuan N, Ebinger J, Langlotz CP, Heidenreich PA, Harrington RA, Liang DH, Ashley EA, Zou JY. Video-based AI for beat-to-beat assessment of cardiac function. Nature. 2020;580:252–256. [PubMed: 32269341]

15. Elias P, Poterucha TJ, Rajaram V, Moller LM, Rodriguez V, Bhave S, Hahn RT, Tison G, Abreau SA, Barrios J, Torres JN, Hughes JW, Perez MV, Finer J, Kodali S, Khalique O, Hamid N, Schwartz A, Homma S, Kumaraiah D, Cohen DJ, Maurer MS, Einstein AJ, Nazif T, Leon MB, Perotte AJ. Deep Learning Electrocardiographic Analysis for Detection of Left-Sided Valvular Heart Disease. J Am Coll Cardiol. 2022;80:613–626. [PubMed: 35926935]

16. Duffy G, Cheng PP, Yuan N, He B, Kwan AC, Shun-Shin MJ, Alexander KM, Ebinger J, Lungren MP, Rader F, Liang DH, Schnittger I, Ashley EA, Zou JY, Patel J, Witteles R, Cheng S, Ouyang D. High-Throughput Precision Phenotyping of Left Ventricular Hypertrophy With Cardiovascular Deep Learning. JAMA Cardiol. 2022;7:386–395. [PubMed: 35195663]

17. He B, Kwan AC, Cho JH, Yuan N, Pollick C, Shiota T, Ebinger J, Bello NA, Wei J, Josan K, Duffy G, Jujjavarapu M, Siegel R, Cheng S, Zou JY, Ouyang D. Blinded, randomized trial of sonographer versus AI cardiac function assessment. Nature. 2023;616:520–524. [PubMed: 37020027]

18. Soto JT, Weston Hughes J, Sanchez PA, Perez M, Ouyang D, Ashley EA. Multimodal deep learning enhances diagnostic precision in left ventricular hypertrophy. Eur Heart J Digit Health. 2022;3:380–389. [PubMed: 36712167]

19. Holste G, Oikonomou EK, Mortazavi BJ, Coppi A, Faridi KF, Miller EJ, Forrest JK, McNamara RL, Ohno-Machado L, Yuan N, Gupta A, Ouyang D, Krumholz HM, Wang Z, Khera R. Severe aortic stenosis detection by deep learning applied to echocardiography. Eur Heart J. 2023;44:4592–4604 [PubMed: 37611002]

20. Vrudhula A, Stern L, Cheng PC, Ricchiuto P, Daluwatte C, Witteles R, Patel J, Ouyang D. Impact of case and control selection on training artificial intelligence screening of cardiac amyloidosis. JACC Adv. 2024;100998.

21. Narang A, Bae R, Hong H, Thomas Y, Surette S, Cadieu C, Chaudhry A, Martin RP, McCarthy PM, Rubenson DS, Goldstein S, Little SH, Lang RM, Weissman NJ, Thomas JD. Utility of a Deep-Learning Algorithm to Guide Novices to Acquire Echocardiograms for Limited Diagnostic Use. JAMA Cardiol. 2021;6:624–632. [PubMed: 33599681]

22. Nkomo VT, Gardin JM, Skelton TN, Gottdiener JS, Scott CG, Enriquez-Sarano M. Burden of valvular heart diseases: a population-based study. Lancet. 2006;368:1005–1011. [PubMed: 16980116]

23. Martin A-C, Bories M-C, Tence N, Baudinaud P, Pechmajou L, Puscas T, Marijon E, Achouh P, Karam N. Epidemiology, Pathophysiology, and Management of Native Atrioventricular Valve Regurgitation in Heart Failure Patients. Front Cardiovasc Med. 2021;8:713658. [PubMed: 34760937]

24. Avierinos J-F, Tribouilloy C, Grigioni F, Suri R, Barbieri A, Michelena HI, Ionico T, Rusinaru D, Ansaldi S, Habib G, Szymanski C, Giorgi R, Mahoney DW, Enriquez-Sarano M, Mitral regurgitation International DAtabase (MIDA) Investigators. Impact of ageing on presentation and outcome of mitral regurgitation due to flail leaflet: a multicentre international study. Eur Heart J. 2013;34:2600–2609. [PubMed: 23853072]

25. Grave C, Tribouilloy C, Tuppin P, Weill A, Gabet A, Juillière Y, Cinaud A, Olié V. Fourteen-Year Temporal Trends in Patients Hospitalized for Mitral Regurgitation: The Increasing Burden of Mitral Valve Prolapse in Men. J Clin Med Res [Internet]. 2022;11. Available from: 10.3390/jcm11123289

26. Mitchell E, Walker R. Global ageing: successes, challenges and opportunities. Br J Hosp Med. 2020;81:1–9.

27. Zhang J, Gajjala S, Agrawal P, Tison GH, Hallock LA, Beussink-Nelson L, Lassen MH, Fan E, Aras MA, Jordan C, Fleischmann KE, Melisko M, Qasim A, Shah SJ, Bajcsy R, Deo RC. Fully

automated echocardiogram interpretation in clinical practice. Circulation. 2018;138:1623–1635. [PubMed: 30354459]

28. Zoghbi WA, Adams D, Bonow RO, Enriquez-Sarano M, Foster E, Grayburn PA, Hahn RT, Han Y, Hung J, Lang RM, Little SH, Shah DJ, Shernan S, Thavendiranathan P, Thomas JD, Weissman NJ. Recommendations for noninvasive evaluation of native valvular regurgitation. J Am Soc Echocardiogr. 2017;30:303–371. [PubMed: 28314623]
29. Tran D, Wang H, Torresani L, Ray J, LeCun Y, Paluri M. A Closer Look at Spatiotemporal Convolutions for Action Recognition [Internet]. arXiv [cs.CV]. 2017 [cited 2023 Oct 19];Available from: http://arxiv.org/abs/1711.11248
30. Selvaraju RR, Cogswell M, Das A, Vedantam R, Parikh D, Batra D. Grad-CAM: Visual explanations from deep networks via gradient-based localization. Int J Comput Vis. 2020;128:336–359.
31. Saito T, Rehmsmeier M. The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets. PLoS One. 2015;10:e0118432. [PubMed: 25738806]
32. Sundararajan M, Taly A, Yan Q. Axiomatic attribution for deep networks [Internet]. arXiv [cs.LG]. 2017 [cited 2023 Nov 4];Available from: http://arxiv.org/abs/1703.01365
33. Yuan N, Jain I, Rattehalli N, He B, Pollick C, Liang D, Heidenreich P, Zou J, Cheng S, Ouyang D. Systematic Quantification of Sources of Variation in Ejection Fraction Calculation Using Deep Learning. JACC Cardiovasc Imaging. 2021;14:2260–2262. [PubMed: 34274282]
34. Yang F, Chen X, Lin X, Chen X, Wang W, Liu B, Li Y, Pu H, Zhang L, Huang D, Zhang M, Li X, Wang H, Wang Y, Guo H, Deng Y, Zhang L, Zhong Q, Li Z, Yu L, Duan Y, Zhang P, Wu Z, Burkhoff D, Wang Q, He K. Automated analysis of Doppler echocardiographic videos as a screening tool for valvular heart diseases. JACC Cardiovasc Imaging. 2022;15:551–563. [PubMed: 34801459]
35. Brown K, Roshanitabrizi P, Rwebembera J, Okello E, Beaton A, Linguraru MG, Sable CA. Using artificial intelligence for rheumatic heart disease detection by echocardiography: Focus on mitral regurgitation. J Am Heart Assoc. 2024;13:e031257. [PubMed: 38226515]
36. Madani A, Arnaout R, Mofrad M, Arnaout R. Fast and accurate view classification of echocardiograms using deep learning. NPJ Digit Med [Internet]. 2018;1. Available from: 10.1038/s41746-017-0013-1
37. Steffner KR, Christensen M, Gill G, Bowdish M, Rhee J, Kumaresan A, He B, Zou J, Ouyang D. Deep learning for transesophageal echocardiography view classification. Sci Rep. 2024;14:11. [PubMed: 38167849]
38. Arnaout R, Curran L, Zhao Y, Levine JC, Chinn E, Moon-Grady AJ. An ensemble of neural networks provides expert-level prenatal detection of complex congenital heart disease. Nat Med. 2021;27:882–891. [PubMed: 33990806]
39. Chiu I-M, Lin C-HR, Yau F-FF, Cheng F-J, Pan H-Y, Lin X-H, Cheng C-Y. Use of a Deep-Learning Algorithm to Guide Novices in Performing Focused Assessment With Sonography in Trauma. JAMA Netw Open. 2023;6:e235102. [PubMed: 36976564]
40. Uretsky S, Gillam L, Lang R, Chaudhry FA, Argulian E, Supariwala A, Gurram S, Jain K, Subero M, Jang JJ, Cohen R, Wolff SD. Discordance between echocardiography and MRI in the assessment of mitral regurgitation severity: a prospective multicenter trial. J Am Coll Cardiol. 2015;65:1078–1088. [PubMed: 25790878]

Clinical Perspective

What is new?

- We developed and validated an automated pipeline for detection of mitral regurgitation from transthoracic echocardiogram studies.
- This pipeline automatically screens for appropriate apical 4-chamber view videos with color Doppler on the mitral valve and then assesses MR severity.

What are the clinical implications?

- The deep learning model proposed in the present study could aid in the screening of MR.
- This deep learning model, and others like it, could facilitate review of institutional databases, or expand access for screening in low-resource settings.

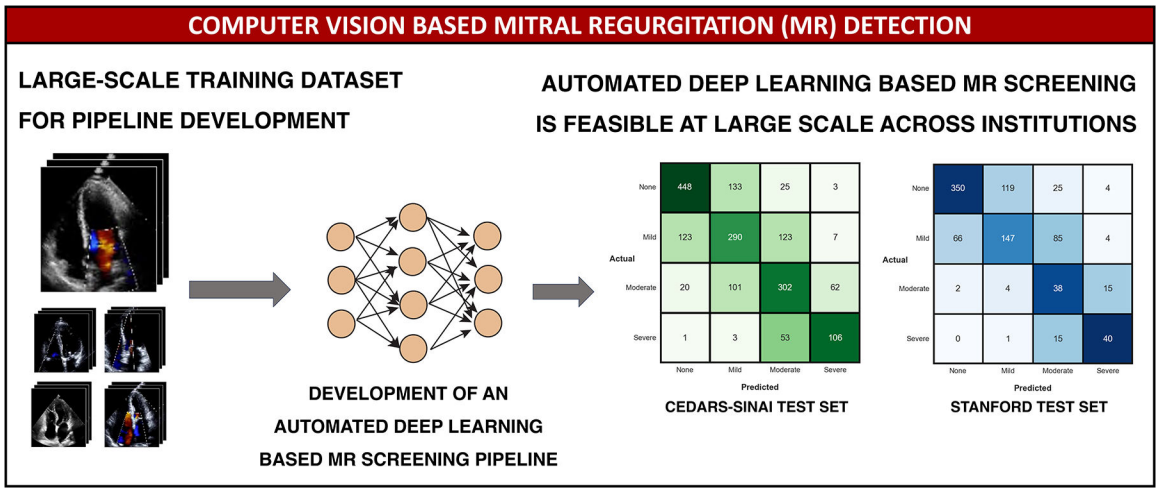

Figure 1: Computer Vision Based Mitral Regurgitation (MR) Detection:

An automated deep learning pipeline was trained to detect and stratify mitral regurgitation severity using large-scale data consisting of apical-4-Chamber (A4C) echocardiogram videos with Color Doppler across the mitral valve (CSMC). The automated pipeline showed strong and consistent performance in test sets at CSMC and SHC. These results show that deep learning can accurately detect clinically significant MR using single-view TTE videos with Doppler information. Deep learning-based MR detection tools could serve as a part of point-of-care ultrasound screening as part of clinic visits or in resource limited settings where imaging may be obtained by individuals with minimal training.

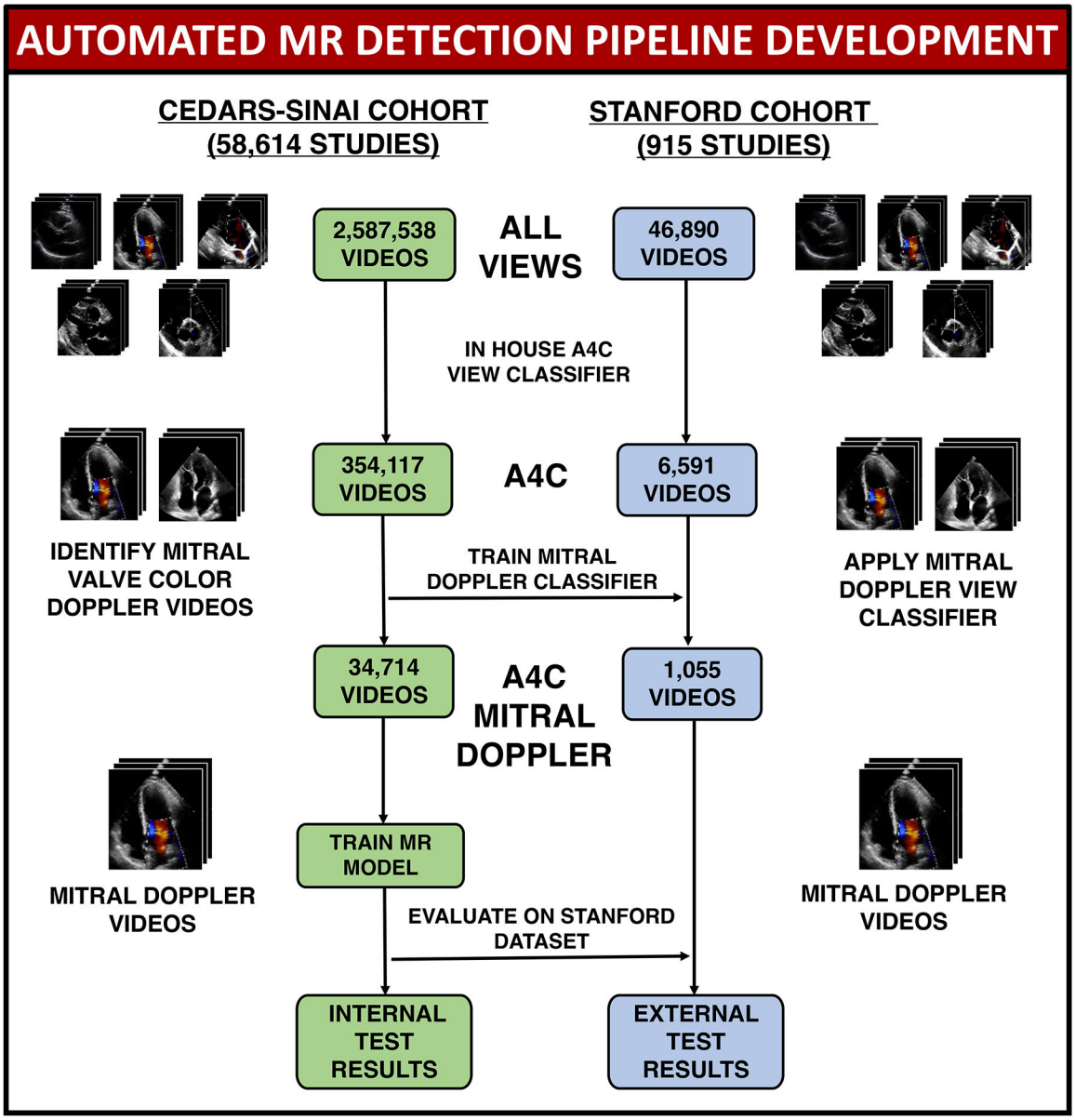

Figure 2: CSMC and Stanford Dataset Isolation.

34,714 Color Doppler A4C videos were isolated from a larger set of videos from CSMC. A view classifier was trained and used to isolate A4C Mitral Doppler videos from 915 studies containing 1,055 suitable videos from Stanford Healthcare. The MR classification model was then benchmarked on an internal test set from CSMC and an external test set from Stanford Healthcare.

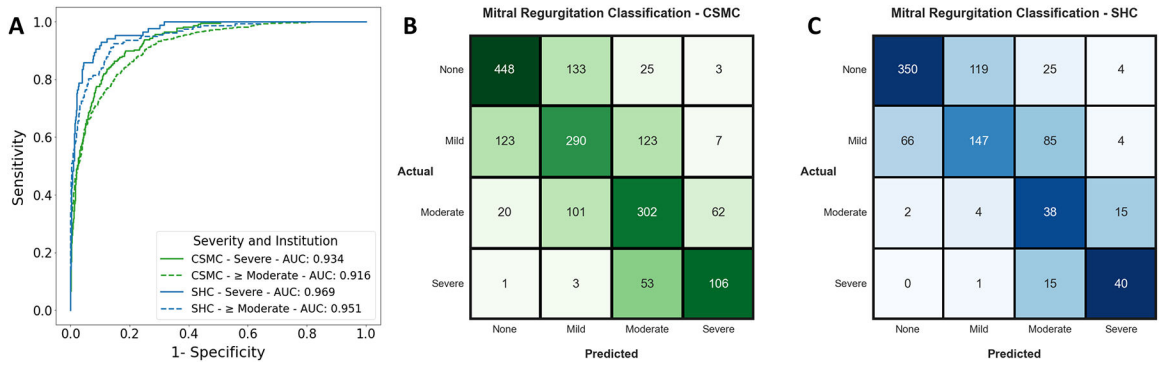

Figure 3: Model Performance Across Severity and Institution –

A. Receiver operating characteristic (ROC) curves for detection of Severe or ≥ Moderate MR at CSMC and Stanford. "≥ Moderate" included moderate, moderate to severe, and severe MR. **3B and 3C:** MR Classification on test set videos from CSMC and Stanford, respectively. Confusion matrix colormap values were scaled based on the proportion of actual disease cases in each class that were predicted in each possible disease category. This was done to allow for relative comparison of model performance across disease classes (None, Mild, Moderate, and Severe) given class imbalance.

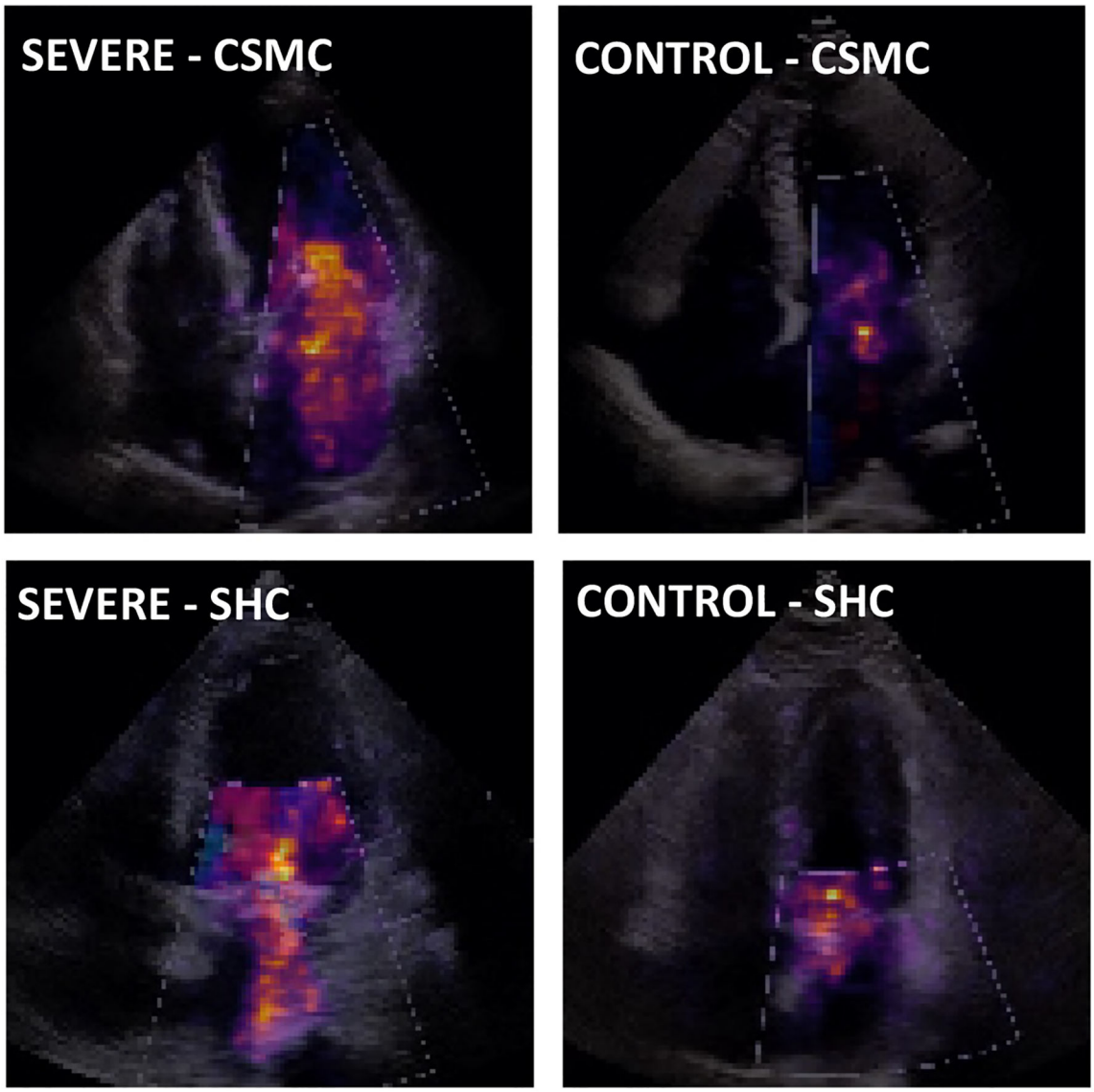

Figure 4: Saliency Map Visualization for MR classification models.

Echocardiogram videos with severe MR from CSMC (top left) and Stanford (bottom left) are shown on the left, while videos with no MR from CSMC (top right) and Stanford (bottom right) are shown on the right. Saliency maps were computed using the Integrated Gradients method. A final 2-dimensional heatmap was generated by using the maximum value along the temporal axis for each pixel location in the video. Pixels brighter in color and closer to yellow were more salient to model predictions, while those darker in color were less important to the model's final prediction. Severe MR was assessed by using the activation function for severe disease to generate a heatmap. When assessing controls, heatmaps were generated by stacking heatmaps for severe and moderate classes and taking the maximum between the two at each pixel location.

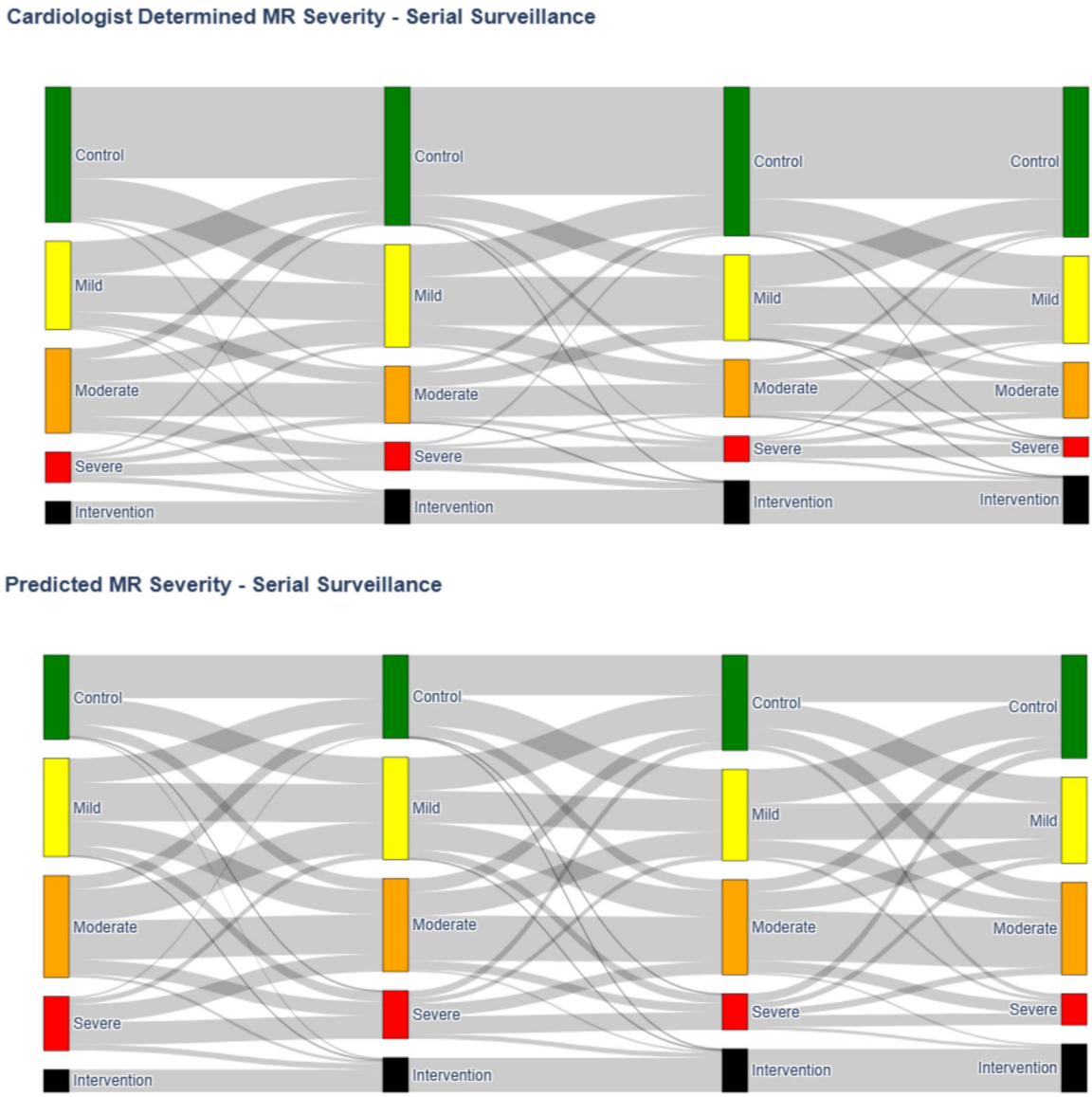

Figure 5: Tracking MR Severity Across Serial Studies.

Sankey diagram of patients with four or more echocardiogram studies with severity of MR tracked over time as evaluated by clinicians and by AI model. Progression and change demonstrated by flow diagrams moving across MR severity.

Table 1 –

View Classifier Demographics

Clinical and demographic characteristics represented in the training, validation, and internal test data sets for the 83,977 apical-4-chamber videos used to train, validate, and test the mitral doppler A4C view classifier. Values outside and inside parentheses represent number and percent, respectively, for categorical variables and mean and standard deviation for continuous variables.

	Train	Validation	Test
Videos, n (%)	67,179 (80.0)	8,344 (9.9)	8,454 (10.1)
Patients, n (%)	30,761 (80.0)	3,848 (10.0)	3,851 (10.0)
Studies, n (%)	46,882 (80.0)	5,790 (9.9)	5,942 (10.1)
Male, n (%)	38,367 (57.1)	4,907 (58.8)	4,854 (57.4)
Hypertension, n (%)	40,982 (61.0)	5,070 (60.8)	5,065 (59.9)
Coronary Artery Disease, n (%)	29,661 (44.2)	3,670 (44.0)	3,758 (44.5)
Atrial Fibrillation, n (%)	19,909 (29.6)	2,289 (27.4)	2,498 (29.5)
Ejection Fraction, mean (SD), %	56.56 (16.2)	57.11 (15.5)	55.85 (16.4)
Left Atrial Volume Index (LAVI), mean (SD), mL/m^2	34.71 (15.4)	34.49 (15.4)	34.05 (15.1)
Race, n (%)			
White	47,110 (70.1)	5,927 (71.0)	5,709 (67.5)
Black	8,282 (12.3)	972 (11.6)	1,228 (14.5)
Asian	5,232 (7.8)	579 (6.9)	640 (7.6)
Other	6,555 (9.8)	866 (10.4)	877 (10.4)
MR Severity, n (%)			
No MR	24,840 (37.0)	3,107 (37.2)	3,291 (38.9)
Mild	25,235 (37.6)	3,143 (37.7)	3,104 (36.7)
Moderate	12,477 (18.6)	1,519 (18.2)	1,518 (18.0)
Severe	4,627 (6.9)	575 (6.9)	541 (6.4)

Table 2 – Mitral Regurgitation Cohort Demographics

Clinical and demographic characteristics of the 20,604 apical-4-chamber videos used to train, validate, and test the MR model. Values outside and inside parentheses represent number and percent, respectively, for categorical variables and mean and standard deviation for continuous variables.

	Train	Validation	CSMC Test	SHC Test
Videos, n (%)	16,533 (80.2)	2,007 (9.7)	2,064 (10.0)	1,091
Patients, n (%)	11,623 (80.0)	1,450 (10.0)	1,450 (10.0)	889
Studies, n (%)	14,565 (80.3)	1,768 (9.8)	1,800 (9.9)	915
Male, n (%)	9,628 (58.2)	1,143 (57.0)	1,152 (55.8)	202 (56.9) *
Hypertension, n (%)	10,087 (61.0)	1,179 (58.7)	1,296 (62.8)	225 (63.4) *
Coronary Artery Disease, n (%)	7,482 (45.3)	888 (44.2)	963 (46.7)	141 (35.5) *
Atrial Fibrillation, n (%)	5,414 (32.7)	599 (29.8)	642 (31.1)	131 (36.9) *
Ejection Fraction, mean (SD), %	54.53 (17.6)	55.26 (17.5)	54.0 (17.9)	51.1 (15.5) **
Left Atrial Volume Index (LAVI), mean (SD), mL/m^2	36.85 (15.8)	36.35 (15.8)	37.44 (16.3)	N/A
Race, n (%)				
White	11,567 (70.0)	1,378 (68.7)	1,400 (67.8)	186 (52.9) *
Black	1,983 (12.0)	277 (13.8)	283 (13.7)	26 (7.3) *
Asian	1,428 (8.6)	147 (7.3)	192 (9.3)	52 (14.6) *
Other	1,555 (9.4)	205 (10.2)	189 (9.2)	62 (25.6) *
MR Severity, n (%)				
No MR	4,949 (29.9)	633 (31.5)	624 (30.2)	555 (50.9)
Mild	4,990 (30.2)	606 (30.2)	586 (28.4)	379 (34.7)
Moderate	4,952 (30.0)	595 (29.6)	627 (30.4)	72 (6.6)
Severe	1,642 (9.9)	173 (8.6)	227 (11.0)	85 (7.8)
MR Etiology				
Primary	1,300 (7.8)	139 (7.0)	157 (7.7)	N/A
Secondary	1,801 (10.9)	222 (11.2)	203 (9.9)	N/A
Unclear	3,500 (21.2)	412 (20.7)	494 (24.2)	N/A
LV Systolic Function				
LVEF ≥ 50	11,602 (70.2)	1,410 (70.8)	1,477 (72.3)	220 (61.9)
LVEF < 50	4,920 (29.8)	580 (29.1)	567 (27.7)	135 (38.0)
LVEF ≥ 35	13,520 (81.8)	1,622 (81.5)	1,708 (83.6)	284 (80.0)
LVEF < 35	3,002 (18.2)	368 (18.5)	336 (16.4)	71 (20.0)
Annular Calcification ┐	1475 (8.9)	148 (7.3)	207 (10.0)	N/A
Aortic Stenosis ❦	981 (5.9)	102 (5.1)	132 (6.4)	N/A
Aortic Regurgitation ❦	1,406 (8.5)	155 (7.7)	212 (10.3)	N/A
Left Atrial Dilation ↖	2,967 (17.9)	363 (18.1)	353 (17.1)	N/A

	Train	Validation	CSMC Test	SHC Test
Difficult Study Quality	2,468 (14.9)	321 (16.0)	321 (15.5)	N/A
Study Year				
< 2017	3,497 (21.2)	413 (20.6)	456 (22.1)	203 (18.6%)
≥ 2017	13,036 (78.8)	1,594 (79.4)	1,608 (77.9%)	888 (81.4%)

*Demographic and comorbidity information was not available for 736 videos in the SHC test set.

**EF information was available for 355 videos from 337 studies in the SHC test set.

τLAVI information was not available for patients in the SHC test set.

Table 3 -

Model performance across institutions -

AUC, PPV, NPV, Recall and F1-score for MR on an internal test set from CSMC and an external validation set at Stanford. 95% confidence intervals were obtained by bootstrapping 10,000 samples. "≥ moderate" includes moderate and severe MR.

Site	Class	AUC	PPV	NPV	Recall	F1-Score
CSMC	≥ Moderate MR	0.916 (0.899 – 0.932)	0.805 (0.766 – 0.842)	0.863 (0.835 – 0.890)	0.807 (0.769 – 0.843)	0.806 (0.776 – 0.834)
	Severe MR	0.934 (0.913 – 0.953)	0.626 (0.533 – 0.713)	0.954 (0.940 – 0.967)	0.626 (0.535 – 0.713)	0.626 (0.546 – 0.695)
Stanford	≥ Moderate MR	0.951 (0.924 – 0.973)	0.509 (0.427– 0.591)	0.986 (0.974 – 0.995)	0.930 (0.867 – 0.983)	0.658 (0.581 – 0.727)
	Severe MR	0.969 (0.946 – 0.987)	0.674 (0.537 – 0.809)	0.977 (0.962 – 0.990)	0.729 (0.580 – 0.860)	0.701 (0.581 – 0.800)

Table 4 –

MR Model Subset Analysis

Subset	Number of Cases	MR Severity Model Performance
Etiology		
Primary	157	Moderate/Severe: 0.913 (0.877 – 0.943) Severe: 0.982 (0.970 – 0.992)
Secondary	203	Moderate/Severe: 0.952 (0.930– 0.970) Severe: 0.975 (0.956 – 0.990)
Unclear	494	Moderate/Severe: 0.903 (0.881–0.923) Severe: 0.939 (0.909–0.965)
Systolic Function		
LVEF ≥ 50	1,477	Moderate/Severe: 0.908 (0.885 – 0.929) Severe: 0.942 (0.915 – 0.965)
LVEF < 50	567	Moderate/Severe: 0.893 (0.854 – 0.928) Severe: 0.901 (0.858 – 0.939)
LVEF ≥ 35	1,708	Moderate/Severe: 0.913 (0.892 – 0.931) Severe: 0.938 (0.914 – 0.959)
LVEF < 35	336	Moderate/Severe: 0.871 (0.812 – 0.923) Severe: 0.893 (0.836 – 0.942)
Atrial fibrillation	642	Moderate/Severe: 0.890 (0.853 – 0.923) Severe: 0.914 (0.871 – 0.949)
Annular Calcification$^{\eta}$	207	Moderate/Severe: 0.892 (0.827 – 0.948) Severe: 0.876 (0.797 – 0.941)
Aortic Stenosis$^{\varphi}$	132	Moderate/Severe: 0.899 (0.816 – 0.964) Severe: 0.873 (0.772 – 0.957)
Aortic Regurgitation$^{\varphi}$	212	Moderate/Severe: 0.857 (0.782 – 0.922) Severe: 0.906 (0.819 – 0.972)
Left Atrial Dilation$^{\kappa}$	353	Moderate/Severe: 0.876 (0.822 – 0.923) Severe: 0.868 (0.808 – 0.920)
Difficult Study Quality	321	Moderate/Severe: 0.926 (0.872 – 0.968) Severe: 0.964 (0.905 – 0.998)
Study Year		
< 2017	456	Moderate/Severe: 0.934 (0.901 – 0.961) Severe: 0.941 (0.905 – 0.970)
≥ 2017	1,608	Moderate/Severe: 0.910 (0.890 – 0.929) Severe: 0.928 (0.901 – 0.952)

$^{\eta}$Patients with moderate/severe annular calcification.

$^{\varphi}$Patients with concomitant moderate/severe valve disease.

$^{\kappa}$Patients with moderate/severe left atrial dilation.