

<https://doi.org/10.1038/s41746-025-02050-x>

Artificial intelligence prediction of age from echocardiography as a marker for cardiovascular disease



Meenal Rawlani¹, Hirotaka Ieki², Christina Binder¹, Victoria Yuan^{1,3}, I-Min Chiu¹, Ankeet Bhatt⁴, Joseph E. Ebinger¹, Yuki Sahashi¹, Andrew P. Ambrosy⁴, Hiroki Usuku⁵, Kenichi Tsujita⁵, Paul Cheng², Alan C. Kwan¹, Susan Cheng¹ & David Ouyang^{1,4} ✉

While chronological age is a universal risk predictor across most populations and diseases, distinguishing between biologically older from younger individuals may identify individuals with accelerated or delayed cardiovascular aging. This study presents a deep learning model to predict age from echocardiogram videos, leveraging 2,610,266 videos from 166,508 studies from 90,738 patients. Leveraging multi-view echocardiography, our model achieved a mean absolute error (MAE) of 6.76 (6.65–6.87) years and a coefficient of determination (R^2) of 0.732 (0.72–0.74) on Cedars-Sinai Medical Center test set. Consistent performance was observed across four external validation cohorts. Predictions were associated with increased risk of coronary artery disease, heart failure, and stroke, and captured discontinuities before and after a heart transplant. Guided back propagation highlighted the model's focus on the mitral valve, mitral apparatus, and basal inferior wall, underscoring the potential of computer vision-based assessment of echocardiography in enhancing cardiovascular risk assessment and aging.

Cardiovascular diseases remain the leading causes of mortality globally, making the accurate assessment of cardiac health essential for effective management and prevention strategies¹. Echocardiography, a widely used, non-invasive cardiac imaging modality, has played a pivotal role in this effort, offering detailed visualization of cardiac structure and function^{2–6}. Advancements in deep learning have opened new avenues for analyzing echocardiographic data, including the prediction of demographic features such as age^{7,8}. These predictions, beyond their immediate novelty, carry potential clinical significance by identifying markers of biological aging and indicators of cardiovascular risk.

A key challenge in the training of AI models in medicine is the heterogeneity, imprecision, and lack of labels for model training^{9,10}. Most medical imaging has few annotations and focused assessments on narrow diagnoses, even though there is much additional information in medical imaging not identifiable by clinicians^{11–13}. Studies are also often constrained by limited datasets, which not only restricts the generalizability of their findings but also acts as a constraint in learning relevant features for an unbiased prediction in deep learning models¹⁴. In contrast, precise age data is almost

always present for medical imaging. While models have largely focused on predicting outcomes like all-cause mortality, such outcomes are limited by few labels, potential loss-of-follow-up, and bias in how datasets are constructed. Less attention has been given to exploring other clinically relevant endpoints or demonstrating the broader predictive capabilities of these systems¹⁵.

In this study, we address these limitations by developing deep learning models to predict age from multiple echocardiographic views. By using parasternal long axis (PLAX), apical two-chamber (A2C), apical four-chamber (A4C), and subcostal (SC) echocardiographic views for a comprehensive assessment of cardiac form and function, as well as a large well-defined cohort of patients, we sought to evaluate whether AI models can predict age and evaluate whether discordance between predictions and chronological age can be used to understand accelerated and delayed biological aging. We also highlight the potential for the model predictions to predict outcomes beyond all-cause mortality demonstrating the versatility and clinical relevance of this approach.

¹Department of Cardiology, Smidt Heart Institute, Cedars-Sinai Medical Center, Los Angeles, CA, USA. ²Division of Cardiology, Department of Medicine, Stanford University, Palo Alto, CA, USA. ³David Geffen School of Medicine at University of California, Los Angeles, CA, USA. ⁴Department of Cardiology and Division of Research, Kaiser Permanente Northern California, Pleasanton, CA, USA. ⁵Department of Cardiovascular Medicine, Graduate School of Medical Sciences, Kumamoto University, Chuo-ku, Kumamoto, Japan. ✉e-mail: David.ouyang@kp.org

Results

Model development and performance

We trained a regression-based video ResNet architecture¹⁶ on each echocardiographic view (PLAX, A2C, A4C, SC) individually. A detailed dataset description is provided in Table 1. To reduce potential bias from advanced cardiovascular disease, our primary model was trained on a cohort that excluded individuals with prior cardiac surgery. Evaluation on a held-out test set from Cedars-Sinai Medical Center (CSMC) revealed that the individual view models achieved mean absolute errors (MAEs) ranging from 8.62 to 9.85 years (Table 2). These predictions from all four views were then combined using Histogram Gradient Boosting¹⁷ to create an ensemble model. Notably, the ensemble approach yielded an overall MAE of 6.76 (6.65–6.87) years and a coefficient of determination (R^2) of 0.732 (0.72–0.74) on the internal test set from CSMC. For comparison, a model trained on the full cohort (including surgical patients) resulted in higher

error in the test cohorts (MAE 7.03, R^2 0.711) (Supplementary Table 1). Imaging characteristics for each predicted groups are presented in Supplementary Table 2. An MAE of 7.20 (7.04–7.36) and R^2 of 0.683 (0.66–0.70) was observed for the external test set from Stanford Healthcare (SHC), MAE of 6.67 (6.64–6.7) and R^2 of 0.723 (0.72–0.73) was observed for the external test set from Kaiser Permanente (KP), MAE of 8.265 (8.06–8.47) and R^2 of 0.483 (0.46–0.51) was observed for the external test set from Chang Gung Memorial Hospital (CGMH) in Taiwan and MAE of 5.29 (4.33, 6.34) and R^2 of 0.35 (–0.19–0.60) was observed for the external test set from Kumamoto University Hospital (KU) in Japan (Table 2). To further evaluate generalizability, we conducted subgroup analyses by race and sex in both internal and external cohorts. The overall MAE and R^2 were robustly consistent, with variations attributable to chance alone. (Supplementary Table 3). To also examine whether echocardiographic age prediction truly reflects biological aging, we integrated creatinine, hemoglobin, and B-type natriuretic peptide

Table 1 | Overview of all datasets

Characteristic	CSMC Cohort ^a			SHC Cohort (External Test) ^a	KP Cohort (External Test) ^a	CGMH Cohort (External Test) ^a	KU Cohort (External Test) ^a
	Training	Validation	Test				
Patients	85,755	897	4,086	4786	85,430	4411	129
Studies	155,243	1685	9,527	5528	96,380	4463	129
Videos	2,436,040	26,239	147,987	76,632	1,309,657	9858	1041
PLAX	873,198 (35.81%)	9479 (36.08%)	53,633 (36.20%)	27,114 (35.38%)	432,171 (33.0%)	4155 (42.15%)	403 (38.7%)
A4C	834,047 (34.20%)	8889 (33.84%)	50,512 (34.09%)	28,432 (37.10%)	466,216 (35.6%)	3875 (39.31%)	374 (35.9%)
SC	401,154 (16.45%)	4392 (16.72%)	23,888 (16.12%)	13,218 (17.25%)	222,461 (17.0%)	705 (7.15%)	78 (7.5%)
A2C	329,998 (13.53%)	3506 (13.34%)	20,135 (13.60%)	7868 (10.27%)	188,809 (14.4%)	1123 (11.39%)	186 (17.9%)
Age (years)	65.97 ± 17.23	65.46 ± 16.42	66.78 ± 16.96	60.96 ± 16.69	66.7 ± 16.1	63.25 ± 15.46	76.4 ± 10.2
Female Count (n, %)	1,155,284 (47.38%)	12,221 (46.53%)	62,416 (42.13%)	34,447 (44.95%)	636,699 (48.6%)	4402 (44.65%)	308 (29.59%)
Race							
White (n, %)	1,651,213 (67.78%)	17,392 (66.33%)	101,974 (68.84%)	41,551 (54.22%)	737,259 (56.3%)	-	-
Asian (n, %)	174,632 (7.17%)	1928 (7.36%)	9838 (6.65%)	12,527 (16.34%)	216,427 (16.5%)	9858 (100.0%)	1041 (100.0%)
Black or African American (n, %)	379,596 (15.59%)	4268 (16.28%)	22,170 (14.98%)	4284 (5.59%)	120,152 (9.2%)	-	-
Other (n, %)	232,956 (9.56%)	2678 (10.21%)	14,186 (9.59%)	18,270 (23.84%)	235,819 (18.0%)	-	-
LVEF < 41 (n, %)	366,729 (15.04%)	3313 (12.61%)	23,781 (16.05%)	9410 (12.28%)	127,564 (10.2%)	780 (8.20%)	97 (9.3%)
LVEF (%)	56.79 ± 15.03	58.07 ± 13.94	56.10 ± 15.41	54.33 ± 12.92	57.3 ± 22.2	63.54 ± 15.81	56.3 ± 10.1
Aortic Regurgitation (n, %)	507,170 (20.80%)	4797 (18.26%)	31,415 (21.20%)	14,657 (19.13%)	351,900 (26.9%)	317 (3.30%)	-
Mitral Regurgitation (n, %)	1,019,600 (41.81%)	10,607 (40.38%)	65,408 (44.14%)	31,734 (41.41%)	653,330 (49.9%)	640 (6.65%)	-
Tricuspid Regurgitation (n, %)	1,149,778 (47.15%)	11,716 (44.61%)	74,395 (50.21%)	37,806 (49.33%)	765,560 (58.5%)	663 (6.89%)	-
Aortic Stenosis (n, %)	169,753 (6.96%)	1570 (5.98%)	12,883 (8.69%)	3899 (5.10%)	123,627 (9.4%)	91 (0.95%)	-
BMI	27.25 ± 8.85	27.65 ± 8.01	27.10 ± 7.18	26.94 ± 6.42	29.0 ± 7.8	26.75 ± 11.98	23.1 ± 3.4
LVIDd 2D (cm)	4.53 ± 1.82	4.50 ± 1.16	4.55 ± 1.08	4.75 ± 0.86	4.66 ± 4.10	4.84 ± 0.72	-
LVIDs 2D (cm)	3.14 ± 1.06	3.09 ± 0.88	3.18 ± 1.06	3.34 ± 0.96	3.18 ± 0.81	3.15 ± 0.83	-
LVPWd 2D (cm)	1.08 ± 1.20	1.09 ± 0.29	1.09 ± 0.31	1 ± 0.21	1.06 ± 0.48	0.94 ± 0.17	-
IVSd 2D (cm)	1.10 ± 0.54	1.10 ± 0.30	1.13 ± 0.34	1.03 ± 0.24	1.09 ± 1.44	1.13 ± 0.23	13.3 ± 3.1
TR Peak Gradient (mmHg)	26.75 ± 12.91	24.91 ± 11.08	27.23 ± 12.66	29.54 ± 13.46	26.7 ± 11.9	24.35 ± 10.15	-

^amean ± std.dev.; n (%).

BMI Body Mass Index, LVIDd 2D Left Ventricular Internal Dimension in Diastole, 2D, LVIDs 2D Left Ventricular Internal Dimension in Systole, 2D, LVPWd 2D Left Ventricular Posterior Wall Dimension in Diastole, 2D, IVSd 2D Interventricular Septum Dimension in Diastole, 2D, TR Peak Gradient Tricuspid Regurgitation Peak Gradient.

Table 2 | Performance metrics for each echocardiographic view evaluated on all cohorts

Echocardiography View	CSMC Cohort		SHC Cohort		KP Cohort		CGMH Cohort		KU Cohort	
	MAE	R ²	MAE	R ²	MAE	R ²	MAE	R ²	MAE	R ²
PLAX	8.662 (8.6–8.72)	0.562 (0.55–0.59)	8.342 (7.72–8.14)	0.575 (0.57–0.58)	7.931 (7.91–7.95)	0.604 (0.6–0.61)	8.228 (8.02–8.43)	0.54 (0.51–0.57)	6.565 (6.05–7.08)	0.05 (–0.25–0.25)
A4C	8.62 (8.56–8.68)	0.569 (0.56–0.58)	9.46 (9.36–9.56)	0.434 (0.42–0.45)	8.639 (8.62–8.66)	0.515 (0.51–0.52)	9.045 (8.8–9.29)	0.376 (0.34–0.40)	6.421 (5.96–6.88)	0.369 (–0.07–0.345)
A2C	9.016 (8.91–9.12)	0.521 (0.51–0.53)	8.479 (8.32–8.64)	0.544 (0.53–0.56)	8.489 (8.46–8.52)	0.524 (0.52–0.53)	9.402 (9.07–9.73)	0.32 (0.28–0.36)	8.326 (7.30–9.37)	–0.39 (–1.14–0.022)
SC	9.851 (9.75–9.94)	0.456 (0.44–0.46)	10.399 (10.26–10.54)	0.351 (0.34–0.37)	9.965 (9.93–10.0)	0.392 (0.39–0.40)	10.142 (9.61–10.68)	0.377 (0.31–0.43)	7.574 (6.316–8.85)	–0.707 (–2.12–(–0.07))
Ensemble	6.76 (6.65–6.87)	0.732 (0.72–0.74)	7.20 (7.04–7.36)	0.683 (0.66–0.70)	6.67 (6.64–6.7)	0.723 (0.72–0.73)	8.265 (8.06–8.47)	0.483 (0.46–0.51)	5.29 (4.33–6.34)	0.35 (–0.19–0.60)

The model was trained on the CSMC cohort, excluding patients who had undergone major cardiac procedures (CABG, heart transplant, or valve replacement).

(BNP) into gradient boosting model alongside echocardiographic predictions from different views, resulting in MAE of 6.86 and R² of 0.721.

Survival analysis based on predicted age groups

We stratified patients based on their predicted age to explore associations with clinical outcomes. Kaplan-Meier survival curves¹⁸ for 5-year follow-up period revealed a clear separation between groups, with studies predicted to be in the older age groups showing significantly worse survival. This separation was also pronounced for coronary artery disease, heart failure, and stroke (Fig. 1)¹⁹. Statistical analysis confirmed that these differences were highly significant across all outcomes ($p < 0.001$). To compare the predictive ability of chronological and biological age, we also performed stratification using chronological age (Supplementary Fig. 1).

We also analyzed the association between age, predicted age, and key cardiovascular outcomes, adjusting for relevant clinical predictors including smoking status, diabetes mellitus, systolic blood pressure, high-density lipoprotein cholesterol, total cholesterol, left ventricular ejection fraction (LVEF), valvular heart disease and prior cardiovascular interventions. Hazard ratios (HRs) were computed per 10-year increase in both chronological age and predicted age (Fig. 2b). For all-cause mortality, while chronological age demonstrated an association (HR = 1.33, 95% CI: 1.28–1.37, $p < 0.005$), predicted age exhibited a stronger effect (HR = 1.42, 95% CI: 1.37–1.48, $p < 0.005$). Similar trends were observed for coronary artery disease (CAD) (chronological age: HR = 1.15, 95% CI: 1.09–1.21, $p < 0.005$; predicted age: HR = 1.25, 95% CI: 1.17–1.34, $p < 0.005$), heart failure (chronological age: HR = 1.16, 95% CI: 1.10–1.21; predicted age: HR = 1.21, 95% CI: 1.14–1.28, $p < 0.005$), and stroke (chronological age: HR = 1.07, 95% CI: 1.01–1.15, $p = 0.03$; predicted age: HR = 1.16, 95% CI: 1.06–1.26, $p < 0.005$). Across all outcomes, predicted age demonstrated a higher HR compared to chronological age.

Predicted cardiac aging phenotypes and their association with valvular conditions

Based on the ensemble model predictions, we classified our study cohort into three cardiac aging phenotypes: “Accelerated”, “Expected”, and “Delayed” based on the difference between chronological age and predicted age (Fig. 2a). Accelerated cardiac aging was defined as having a residual (Predicted Age – Chronological Age) below the lower 95% confidence interval limit, where the predicted age was significantly higher than the chronological age. Delayed cardiac aging was defined as having a residual above the upper limit, where the predicted age was much lower than the chronological age. Expected cardiac aging was characterized by residuals within the confidence interval, where the predicted age closely matched the chronological age.

We found a clear trend of increasing prevalence of valvular heart disease as the cohort’s predicted age group progresses (Table 3). Specifically, individuals classified as having accelerated cardiac aging exhibited higher rates of aortic stenosis as well as aortic, mitral and tricuspid regurgitation compared to those in the expected and delayed aging groups²⁰. Conversely,

individuals in the delayed aging category generally showed significantly lower rates of these conditions across most predicted age groups.

Age prediction in heart transplant patients

We further investigated age-related patterns in heart transplant patients by analyzing the relationship between chronological age and model predicted age. We calculated the mean age of heart transplant patients ($n = 53$) within 1 year before and after the transplant. For the pre-transplant period, the mean age was 62.57 ± 7.57 years, while for the post-transplant period, the mean age decreased to 57.32 ± 7.78 years (t-statistic = 4.81, $p < 0.001$), indicating a substantial reduction in mean predicted age following the transplant. Importantly, 37 of the 50 patients who had studies both before and after the transplant in the one-year period were predicted to be younger following the transplant. Our analysis revealed that the model frequently predicted patients to be younger following heart transplantation, a trend reflected in Supplementary Fig. 2.

Atherosclerotic cardiovascular disease risk score across predicted age groups

To further evaluate the clinical relevance of our model’s age predictions, we grouped studies according to their predicted age (< 40, 40–50, 50–60, 60–70, 70–80 and > 80) and compared their atherosclerotic cardiovascular disease (ASCVD) risk scores²¹ (Fig. 2c). Out of the 9580 studies in the test set, we were able to calculate the ASCVD score for 4773 studies, as traditional ASCVD risk scores typically exclude individuals under 40, over 80, and those with pre-existing ASCVD. Additionally, certain parameters were unavailable for some studies, further limiting score calculation. We observed an incremental rise in mean 10-year ASCVD risk with each predicted age group. For individuals predicted to be under 40 years (54 studies), the mean 10-year ASCVD risk was 4.6%. In the predicted 40–50 age group (427 studies), the mean risk increased to 8.3%. The predicted 50–60 age group (1140 studies) had a further increase to 9.6%, while those predicted 60–70 years (1726 studies) had a mean 10-year ASCVD risk of 15.2%. In the predicted 70–80 age group (1154 studies), the mean risk rose significantly to 22.6%. Finally, for individuals predicted to be over 80 (272 studies), the mean 10-year ASCVD risk was 27.8%. We also compared the studies grouped by predicted age with Framingham risk score²² (Supplementary Fig. 3).

Interpretability

To evaluate the most relevant cardiac structures, we employed guided backpropagation²³ to visualize regions of interest within each echocardiographic view that contributed most significantly to the model’s predictions (Fig. 3, Supplementary Fig. 4). The saliency maps indicated that the model concentrated predominantly on the regions of the crux cordis, mitral valve, and the mitral annulus across multiple views, suggesting consistent regions of interest that highlight important cardiac structures^{24,25}. To further quantify the contribution of specific anatomical structures, SHAP²⁶ was implemented (Supplementary Fig. 5). A quantitative analysis

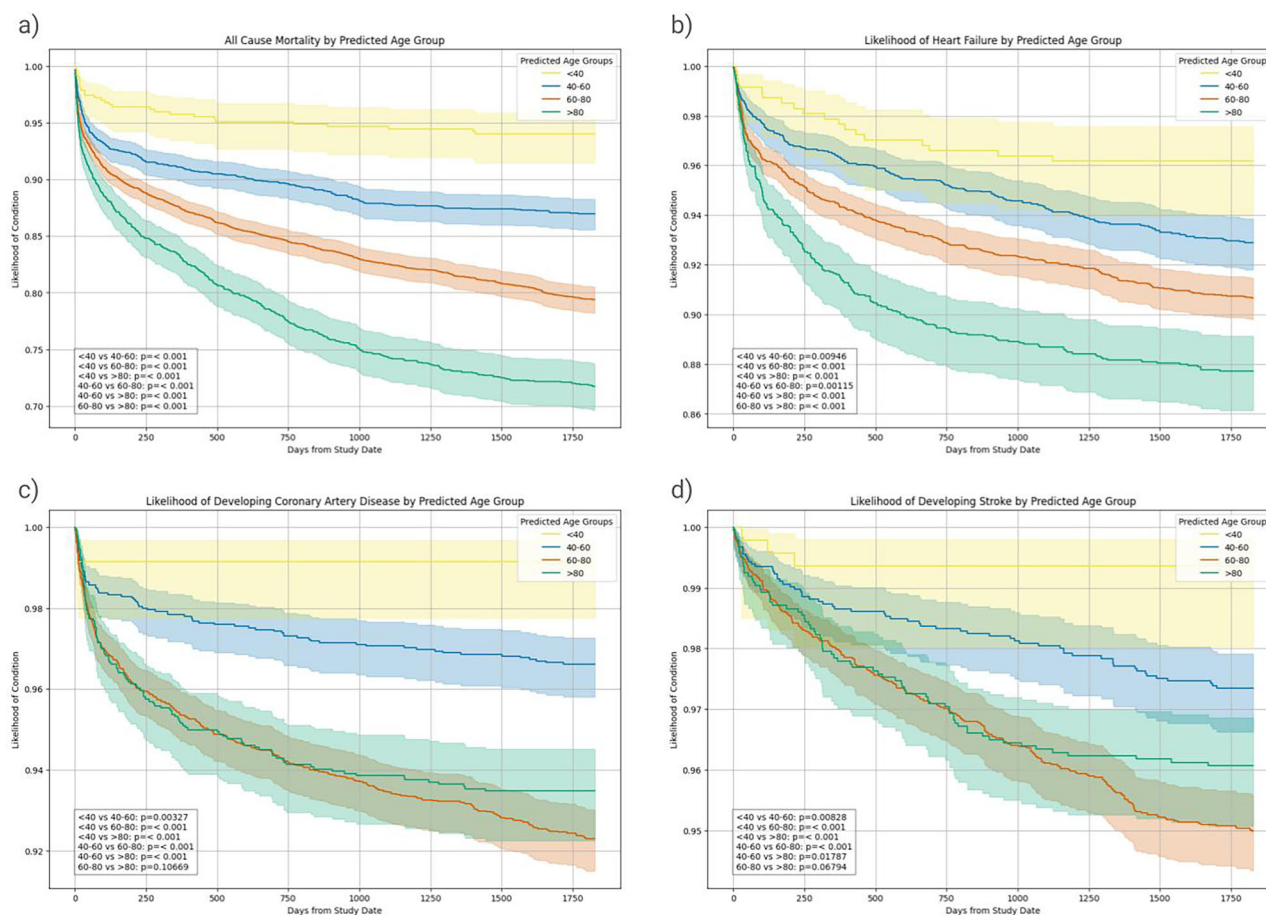


Fig. 1 | Kaplan-Meier curves depicting the association between predicted age group and clinical endpoints. The survival curves illustrate the association between predicted age group and four clinical endpoints: **a** all-cause mortality, **b** heart failure,

c coronary artery disease, and **d** stroke. Each panel displays the survival probability over 5 years for groups stratified by predicted age, with steeper declines indicating higher event rates.

demonstrating a monotonic increase in predicted biological age with increasing mitral annular calcification severity^{25,27} was also done (Supplementary Fig. 6).

Discussion

Our study demonstrates that deep learning models can accurately predict age from echocardiogram videos, potentially serving as a proxy for biological cardiac aging. A key aspect of our analysis is the interpretation of residuals and discrepancies between predicted and chronological age, as ability to identify individuals whose predicted age deviates from their chronological age could offer valuable insights into cardiovascular health. Our model demonstrated robust generalizability, with consistent performance across diverse sex and race subgroups as well as across multiple institutions, including CSMC, SHC, KP, CGMH, and KU. We note that the KU cohort from Japan had a smaller sample size and represented a demographically distinct population which may have contributed to the lower performance observed in that cohort. Overall, the cross-site and subgroup consistency supports the broad applicability of predicted cardiac age as a clinically meaningful biomarker. Furthermore, by comparing saliency maps across multiple cardiac views and images showing different anatomic structures, we identify key cardiac structures that change over time and with accelerated aging. Our findings suggest that these imaging features of aging may capture clinically significant aging processes.

The distinct separation observed in Kaplan–Meier survival curves for overall survival, coronary artery disease, heart failure, and stroke when classifying patients into accelerated or delayed aging underscores the clinical relevance of our model's predictions. This provides strong evidence that predicted age serves as a valuable prognostic tool for cardiovascular risk,

outperforming stratification by chronological age as an alternative. Across all outcomes, predicted age consistently demonstrated a higher hazard ratio, suggesting that predicted age estimates capture additional risk information beyond traditional age-based assessments. Additionally, the incremental rise in mean ASCVD risk scores across predicted age groups further validates our model's clinical relevance²¹ and underscores the potential of integrating our model for more accurate risk stratification and prognosis.

An intriguing aspect of our findings is the “rejuvenation” effect observed in post-heart transplant patients, where the model predicted younger ages following transplant. This observation aligns with the model's focus on age-sensitive cardiac features, suggesting that post-transplant cardiac morphology and function resemble that of a younger heart, despite the recipient's chronological age²⁸. The observed reduction in predicted age reflects a transformation in cardiac imaging phenotype from a structurally diseased to a healthy donor heart rather than a reversal of biological age. Predicted age in this context should be understood as a surrogate for cardiac age, not systemic aging.

The guided backpropagation saliency maps revealed that our model focuses on clinically relevant cardiac structures, primarily the crux cordis, mitral annulus, and mitral valve across all views. The mitral annulus is a prominent region where age-related structural and functional changes, such as mitral annular calcification, commonly occur²⁵. Mitral annular calcification has been identified as a potential marker of biological aging that is associated with cardiovascular as well as non-cardiovascular morbidity and mortality^{29,30}. The mitral valve is also prone to degenerative changes like sclerosis or calcification with age^{25,27,31}. This alignment lends biological plausibility to our model's predictions and enhances its interpretability for clinicians. This was also confirmed by SHAP interpretability method.

Table 3 | Chi-square test results for the association between predicted and chronological age groups across various valvular abnormalities (stenosis and regurgitation) in different predicted age categories

Predicted Age Group	Valvular Disorders	Predicted age significantly lower than chronological age	Predicted age within expected range	Predicted age significantly higher than chronological age	Chi² Statistic	Degrees of Freedom	p-value	Odds Ratio ^a
<40	Aortic Stenosis	—	3 (0.50%)	5 (3.97%)	8.59	1	0.003	8.26 (1.95, 35.04)
	Mitral Regurgitation	—	96 (15.92%)	45 (35.71%)	24.92	1	<0.001	2.93 (1.92, 4.49)
	Tricuspid Regurgitation	—	185 (30.68%)	62 (49.21%)	15.15	1	<0.001	2.19 (1.48, 3.23)
	Aortic Regurgitation	—	37 (6.14%)	26 (20.63%)	25.94	1	<0.001	3.98 (2.31, 6.86)
40–60	Aortic Stenosis	0 (0.00%)	40 (1.84%)	17 (11.26%)	52.75	2	<0.001	6.77 (3.74, 12.26)
	Mitral Regurgitation	1 (12.50%)	615 (28.29%)	67 (44.37%)	18.73	2	<0.001	2.02 (1.45, 2.82)
	Tricuspid Regurgitation	3 (37.50%)	799 (36.75%)	88 (58.28%)	27.73	2	<0.001	2.40 (1.72, 3.36)
	Aortic Regurgitation	0 (0.00%)	220 (10.12%)	43 (28.48%)	48.59	2	<0.001	3.54 (2.42, 5.17)
60–80	Aortic Stenosis	0 (0.00%)	260 (6.40%)	3 (15.00%)	9.37	2	0.009	2.58 (0.75, 8.86)
	Mitral Regurgitation	7 (7.00%)	1590 (39.15%)	12 (60.00%)	46.55	2	<0.001	2.33 (0.95, 5.72)
	Tricuspid Regurgitation	31 (31.00%)	1882 (46.34%)	14 (70.00%)	13.87	2	<0.001	2.70 (1.04, 7.04)
	Aortic Regurgitation	3 (3.00%)	721 (17.75%)	9 (45.00%)	25.18	2	<0.001	3.79 (1.56, 9.18)
>80	Aortic Stenosis	14 (13.59%)	451 (20.58%)	—	2.56	1	0.11	—
	Mitral Regurgitation	46 (44.66%)	1303 (59.47%)	—	8.31	1	0.004	—
	Tricuspid Regurgitation	63 (61.17%)	1507 (68.78%)	—	2.3	1	0.129	—
	Aortic Regurgitation	30 (29.13%)	810 (36.97%)	—	2.28	1	0.131	—

^aOdds ratio of event rate comparing predicted age significantly higher than chronological age vs. predicted age within expected range.

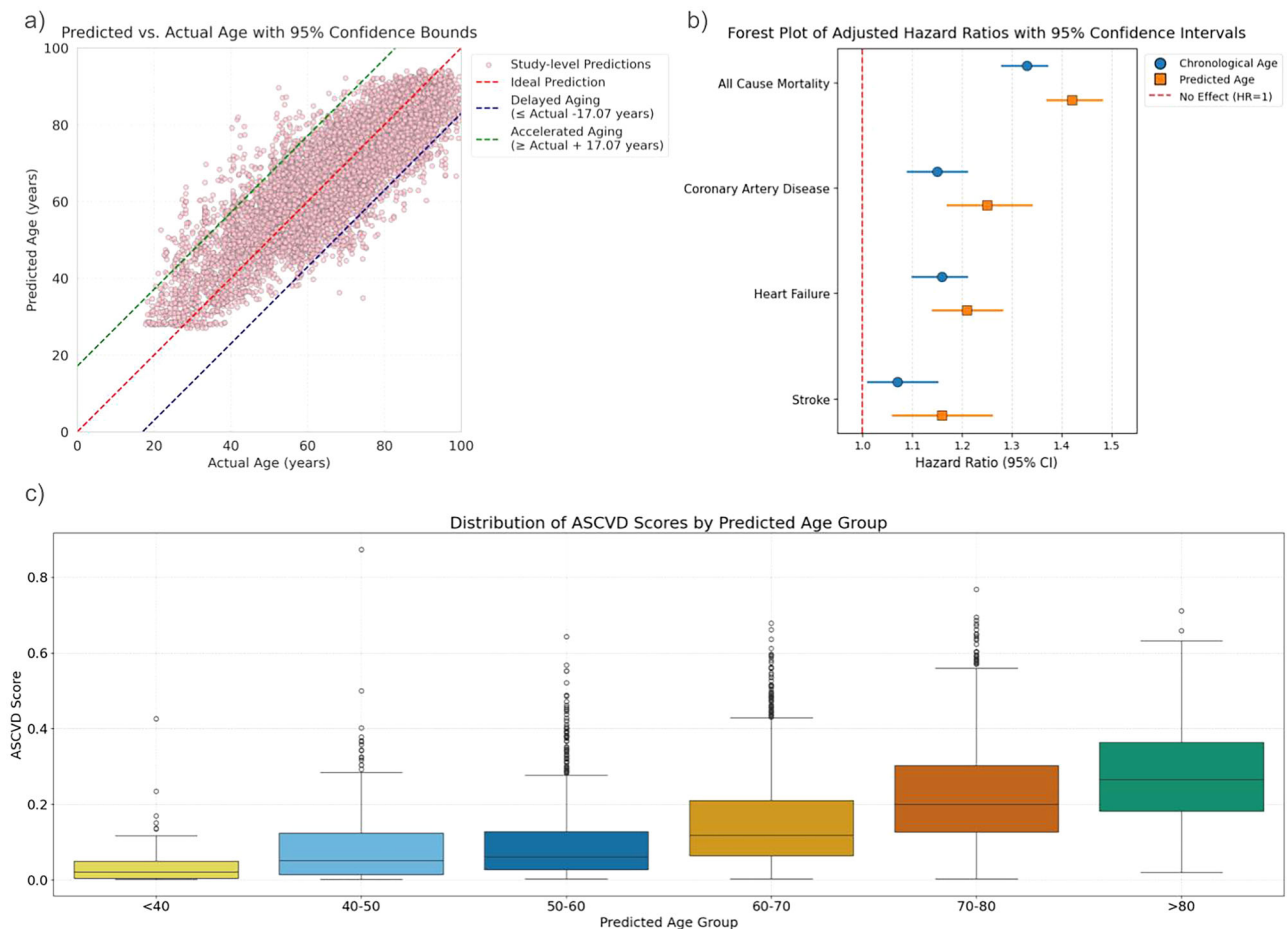


Fig. 2 | Quantifying accelerated and delayed aging and the use of predicted age for cardiovascular risk stratification. a Scatter plot of predicted vs. chronological age. The red diagonal line indicates normal aging. Dashed lines show accelerated (above) and delayed (below) aging, defined using the 95% confidence interval (CI) of the residuals, where the residual is defined as (predicted age – chronological age). Thresholds are calculated by adding and subtracting 1.96 times the standard

deviation of the residuals to the identity line. Each dot represents a study. **b** Forest plot of adjusted hazard ratios (HRs) with 95% CI for age and predicted age across cardiovascular outcomes. **c** Box plots of ASCVD risk scores by six predicted age groups. Boxes show IQRs with medians; whiskers represent $1.5 \times \text{IQR}$, and outliers are shown as individual points.

Despite these promising findings, several limitations warrant consideration. Although echocardiography is commonly used as a screening tool in both symptomatic and asymptomatic individuals, the dataset is derived from patients undergoing clinically indicated imaging and therefore may still be enriched for individuals with underlying health concerns due to the clinical setting. While we excluded patients with major cardiovascular interventions (e.g., CABG, valve replacement, transplantation) to minimize disease related bias, the cohort may not fully reflect the distribution of cardiac aging in an asymptomatic or community-based population. Future studies validating the model in asymptomatic screening or population-based cohorts will be important to assess generalizability to lower risk groups and to support broader use in preventive cardiology. Given the accessibility of echocardiography in many care settings, including those with limited resources, AI-predicted age may offer a scalable tool for cardiovascular risk stratification. Prospective studies are needed to assess how such tools affect clinical decision-making. Furthermore, the performance of creatinine, hemoglobin, and BNP integrated model was comparable to models using echocardiographic views alone, suggesting that the age predictions reflect underlying biological processes associated with cardiovascular aging. Further validation with multi-omics markers, such as DNA methylation-based clocks or proteomic aging indices, is warranted to understand the biological underpinnings of predicted cardiac age³². We acknowledge that this study represents preliminary discovery science, and further clinical validation is essential. Prospective, randomized trials will be

necessary to establish whether discrepancies between predicted and chronological age can inform clinical decision-making or guide therapeutic interventions.

While the prediction of biological age could improve cardiovascular risk stratification and informing preventive care, it also raises important ethical considerations. There is potential for misuse if these predictions are used by institutions, specifically insurance companies, to discriminate based on model predicted biological age. Similar issues have recently come into light in the context of epigenetic age estimators too and hence calls for regulatory oversight emerged to prevent misuse³³. As biological age estimators, including those derived from echocardiography, become more accessible, it is imperative that ethical guidelines and protections are developed to govern their appropriate use and ensure they benefit individuals without reinforcing bias against the already vulnerable populations.

In conclusion, this study highlights the potential of deep learning models applied to echocardiographic data for predicting biological age and assessing cardiovascular risk. By providing interpretable insights and demonstrating clinical relevance, this approach paves the way for more personalized and proactive cardiovascular care.

Methods

Cohort selection

Echocardiography studies spanning from 2012 to 2022 were obtained from the CSMC. The videos were originally acquired using GE or Philips

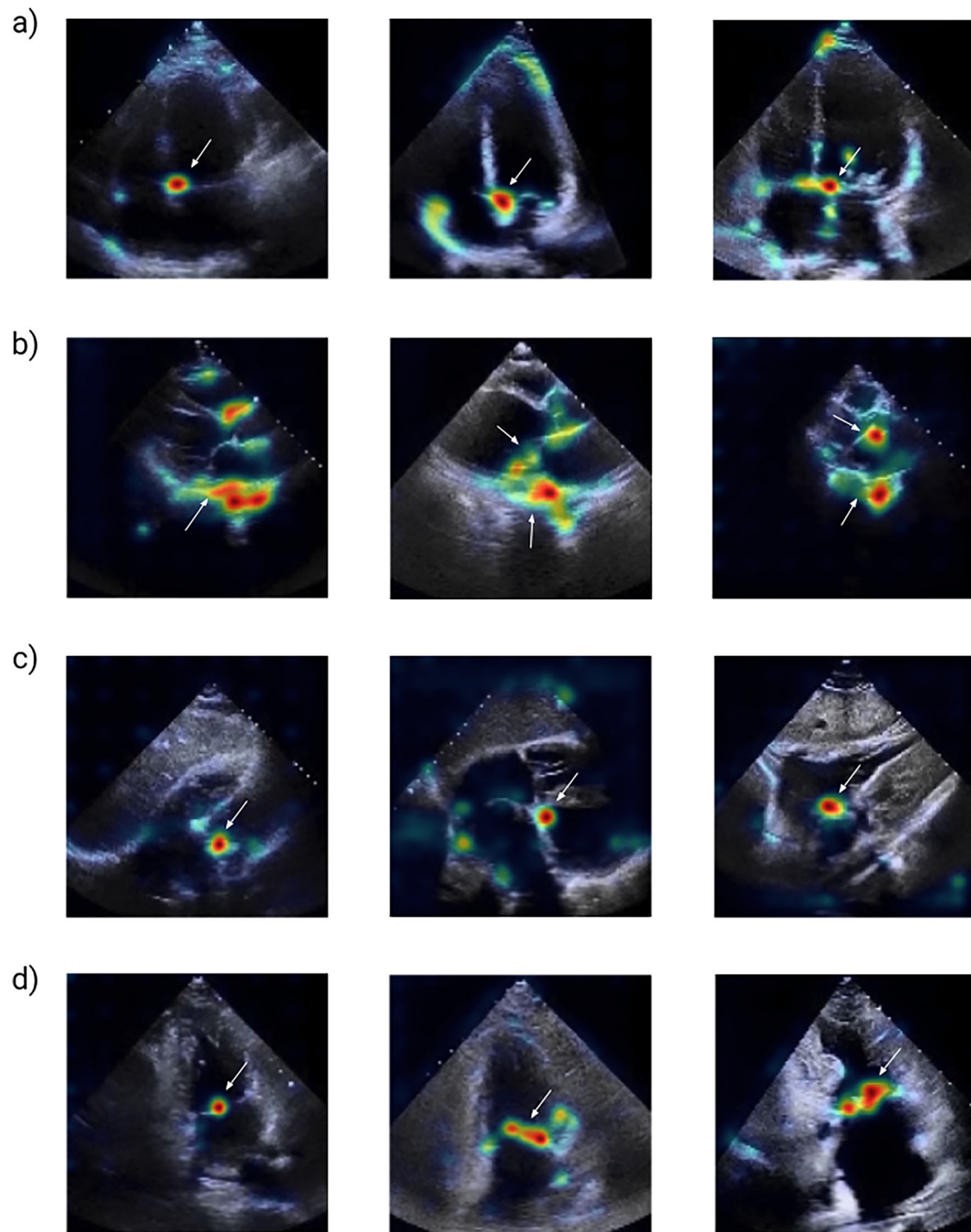


Fig. 3 | Saliency maps for all echocardiographic views. Guided backpropagation interpretability method highlighting the most important regions of interest for the model to make determination of age for each echocardiographic view: **a** A4C, **b** PLAX, **c** SC, and **d** A2C. As highlighted by the heatmap and the white arrows, the regions of the crux cordis, mitral valve, and the mitral annulus are highlighted across views.

ultrasound machines, stored as DICOM files, and converted into AVI format³⁴. These videos files underwent pre-processing to ensure consistency and extraneous details beyond the ultrasound section were removed, metadata was de-identified, and the videos were resized to resolution of 112×112 to facilitate computational efficiency.

The dataset included videos from four standard echocardiographic views: A4C, A2C, PLAX, and SC. Doppler videos were excluded from the dataset. Each video was filtered to include at least 32 frames. To capture temporal diversity while ensuring computational efficiency, we selected 16

frames from a continuous 32-frame window using a uniform stride of 2, i.e., every second frame was sampled until a total of 16 frames was obtained³⁵. The studies also underwent additional filtering to exclude patients with prior histories of coronary artery bypass grafting (CABG), cardiac valve replacement, or heart transplant from the training and the validation splits. The demographics, diagnoses, and procedural history was extracted from the electronic health record, where it is stored by international classification of disease (ICD) codes (Supplementary Table 4). This exclusion was done to ensure model learned features without confounding effects from such

interventions. In the CSMC cohort, 25,940 individuals (28.6%) had no documented history of heart diseases. The dataset was divided into training, validation, and test splits based on unique patients to prevent data leakage. A detailed overview of the included studies is presented in Table 1. Separately, a cohort of 5,528 studies from SHC, 96,380 studies from KP, 4,463 studies from CGMH and 129 studies from KU were used as an external validation set (Table 1). This study was approved by the Cedars-Sinai Medical Center (Reference Number: STUDY00001049), Stanford Healthcare (Reference Number: 43721), Kaiser Permanente (Reference Number: 2276927), Chang Gung Memorial Hospital and Kumamoto University Hospital Institutional Review Boards, and the requirement for informed consent was waived for retrospective data analysis without patient contact.

Model description and implementation details

The spatiotemporal relationships within echocardiographic videos were captured using a deep learning model based on the ResNet architecture (R2 + 1D)¹⁶. This model processes 3D video data using a combination of 2D spatial convolutions and 1D temporal convolutions at every convolutional layer. Each block in the network applies spatial filtering across the 2D image plane, followed by temporal filtering across frames, allowing the network to learn both anatomical structure and cardiac motion over time. The architecture was chosen to balance computational efficiency with the ability to capture dynamic cardiac features across both space and time, superior generalization with lower and more stable validation loss³⁵. Models were trained using stochastic gradient descent with the ADAM optimizer, with an initial learning rate of 0.01, momentum of 0.9, and a learning rate decay schedule. Training was conducted for up to 50 epochs, with early stopping applied if validation loss failed to improve for 10 consecutive epochs. A batch size of 32 was used.

Deep learning models were trained separately for each echocardiographic view (A4C, A2C, PLAX, and SC). To further improve performance, an ensemble model was constructed using a gradient boosting regressor¹⁷. The inputs to the ensemble model were predictions from the individual view-specific models, generated by running inference on the validation set. This method builds an ensemble of decision trees in a stage-wise manner by optimizing a squared error loss function, where each new tree attempts to correct the residuals from the previous iteration. This allowed the ensemble model to leverage complementary information from different echocardiographic views while making sure that the same test set was used for evaluating both individual and ensemble models, ensuring a fair comparison of performance.

Statistical analysis

Model performance was assessed through mean absolute error (MAE) and the coefficient of determination (R^2) for the age prediction task, using both the test and external validation datasets. Stratified analyses were conducted based on predicted age groups to explore clinical outcomes. The Z-score of 1.96 is used to calculate the 95% confidence interval (CI) by multiplying it with the standard deviation of the residuals (chronological age – predicted age). The upper bound (accelerated aging) is the mean residual plus this CI, and the lower bound (delayed aging) is the mean residual minus this CI. To validate the use of a global 95% CI for residual-based aging classification, we computed subgroup-specific residual distributions (Supplementary Fig. 7). Kaplan–Meier survival curves were employed to assess the separation between groups for overall survival, coronary artery disease, heart failure, and stroke. The significance of these differences was evaluated using log-rank tests for pairwise comparisons and a multivariate log-rank test to assess overall differences across all age groups. A multivariate Cox proportional hazards model to assess the association between chronological age and predicted age and the time-to-event outcome. Covariates included smoking status, diabetes mellitus, systolic blood pressure, high-density lipoprotein cholesterol (hdlc), total cholesterol, left ventricular ejection fraction (LVEF), valvular heart disease and prior

cardiovascular interventions and were selected based on clinical relevance. Hazard ratios (HR) and 95% confidence intervals (CI) were reported to quantify the effects of each variable on survival outcomes. Pooled Cohort Equations were used to calculate the ASCVD risk score. Statistical analysis was performed with Python (version 3.10.12), and 95% confidence intervals were computed based on 10,000 bootstrap samples.

Data availability

The patient data is not publicly available due to their potentially identifiable nature.

Code availability

The code and model weights are available at <https://github.com/echonet/aging>.

Received: 8 April 2025; Accepted: 30 September 2025;

Published online: 18 November 2025

References

1. Roth, G. A. et al. Global, regional, and national age-sex-specific mortality for 282 causes of death in 195 countries and territories, 1980–2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet* **392**, 1736–1788 (2018).
2. Marwick, T. H. The role of echocardiography in heart failure. *J. Nucl. Med.* **56**(Suppl 4), 31S–38S (2015).
3. Feigenbaum, H. et al. Role of echocardiography in patients with coronary artery disease. *Am. J. Cardiol.* **37**, 775–786 (1976).
4. Patrianakos, A. P. et al. The growing role of echocardiography in interventional cardiology: the present and the future. *Hellenic J. Cardiol.* **58**, 17–31 (2017).
5. Villarraga, H. R., Herrmann, J. & Nkomo, V. T. Cardio-oncology: role of echocardiography. *Prog. Cardiovasc. Dis.* **57**, 10–18 (2014).
6. Lee, J.-H. & Park, J.-H. Role of echocardiography in clinical hypertension. *Clin. Hypertens.* **21**, 9 (2015).
7. Faierstein, K. et al. Artificial intelligence assessment of biological age from transthoracic echocardiography: Discrepancies with chronologic age predict significant excess mortality. *J. Am. Soc. Echocardiogr.* **37**, 725–735 (2024).
8. Duffy, G. et al. Confounders mediate AI prediction of demographics in medical imaging. *NPJ Digit. Med.* **5**, 188 (2022).
9. Carini, C. & Seyhan, A. A. Tribulations and future opportunities for artificial intelligence in precision medicine. *J. Transl. Med.* **22**, 411 (2024).
10. Ahmed, M. I. et al. A systematic review of the barriers to the implementation of artificial intelligence in healthcare. *Cureus* **15**, e46454 (2023).
11. Holmstrom, L. et al. Deep learning-based electrocardiographic screening for chronic kidney disease. *Commun. Med.* **3**, 73 (2023).
12. Hughes, J. W. et al. Deep learning evaluation of biomarkers from echocardiogram videos. *EBioMedicine* **73**, 103613 (2021).
13. Sahashi, Y. et al. Opportunistic screening of chronic liver disease with deep learning enhanced echocardiography. <https://doi.org/10.1101/2024.06.13.24308898> (2024).
14. Cui, Z. & Gong, G. The effect of machine learning regression algorithms and sample size on individualized behavioral prediction with functional connectivity features. *Neuroimage* **178**, 622–637 (2018).
15. Ghassemi, M. et al. A review of challenges and opportunities in machine learning for health. *AMIA Summits Transl. Sci. Proc.* **2020**, 191–200 (2020).
16. Tran, D. et al. A Closer Look at Spatiotemporal Convolutions for Action Recognition. *arXiv:1711.11248 [cs.CV]* (2017).
17. Preetenhofer, P. & Louppe, G. Gradient Boosted Regression Trees in Scikit-Learn. PyData 2014 <https://hdl.handle.net/2268/163521> (2014).

18. Bland, J. M. & Altman, D. G. Survival probabilities (the Kaplan-Meier method). *BMJ* **317**, 1572 (1998).
19. Wilson, P. W. F. Overview of established risk factors for cardiovascular disease. *UpToDate Updated* **13**, 1–2 (2018).
20. Singh, J. P. et al. Prevalence and clinical determinants of mitral, tricuspid, and aortic regurgitation (the Framingham Heart Study). *Am. J. Cardiol.* **83**, 897–902 (1999).
21. Goff, D. C. Jr et al. 2013 ACC/AHA guideline on the assessment of cardiovascular risk: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines: a report of the American college of cardiology/American heart association task force on practice guidelines. *Circulation* **129**, S49–S73 (2014).
22. D'Agostino, R. B. Sr, Pencina, M. J., Massaro, J. M. & Coady, S. Cardiovascular disease risk assessment: insights from Framingham. *Glob. Heart* **8**, 11–23 (2013).
23. Springenberg, J. T., Dosovitskiy, A., Brox, T. & Riedmiller, M. Striving for simplicity: The all convolutional net. *arXiv:1412.6806 [cs.LG]* (2014).
24. Otto, C. M., Lind, B. K., Kitzman, D. W., Gersh, B. J. & Siscovick, D. S. Association of aortic-valve sclerosis with cardiovascular mortality and morbidity in the elderly. *N. Engl. J. Med.* **341**, 142–147 (1999).
25. Eberhard, M. et al. Mitral annular calcification in the elderly - Quantitative assessment. *J. Cardiovasc. Comput. Tomogr.* **15**, 161–166 (2021).
26. Lundberg, S. & Lee, S.-I. A unified approach to interpreting model predictions. *arXiv:1705.07874 [cs.AI]* (2017).
27. Opris, C. E., Suci, H., Opris, C. I. & Gurzu, S. An update on mitral valve aging. *Life (Basel)* **14**, 950 (2024).
28. Stehlik, J. et al. The registry of the International Society for Heart and Lung Transplantation: 29th official adult heart transplant report-2012. *J. Heart Lung Transplant.* **31**, 1052–1064 (2012).
29. Morariu, P. C. et al. Rethinking mitral annular calcification and its clinical significance: From passive process to active pathology. *J. Pers. Med.* **14**, 900 (2024).
30. Oni, E. et al. The association of mitral annular calcification with cardiovascular and noncardiovascular outcomes: The Multi-Ethnic Study of Atherosclerosis. *Am. J. Cardiol.* **225**, 75–83 (2024).
31. Rostagno, C. Heart valve disease in elderly. *World J. Cardiol.* **11**, 71–83 (2019).
32. Field, A. E. et al. DNA methylation clocks in aging: categories, causes, and consequences. *Mol. Cell* **71**, 882–895 (2018).
33. Dupras, C. et al. Potential (mis)use of epigenetic age estimators by private companies and public agencies: human rights law should provide ethical guidance. *Environ. Epigenet.* **5**, dvz018 (2019).
34. Ouyang, D. ConvertDICOMToAVI.ipynb at master · echonet/dynamic. *GitHub*. <https://github.com/echonet/dynamic/blob/master/scripts/ConvertDICOMToAVI.ipynb>.
35. Ouyang, D. et al. Video-based AI for beat-to-beat assessment of cardiac function. *Nature* **580**, 252–256 (2020).

Acknowledgements

V.Y. acknowledges funding from the Sarnoff Cardiovascular Research Foundation. Y.S. reports support from the KAKENHI (Japan Society for the Promotion of Science: 24K10526) and honoraria for consulting from m3.com inc. D.O. reports support from the National Institute of Health (NIH; NHLBI R00HL157421, R01HL173487 and R01HL173526) and Alexion, and consulting or honoraria for lectures from EcholQ, Ultromics, Pfizer, InVision, the Korean Society of Echocardiography, and the Japanese Society of Echocardiography. CB reports support from the Max-Kade Foundation.

Author contributions

Concept and design: M.R., D.O. Code and programming: M.R., H.I. Data acquisition, analysis, or interpretation of data: M.R., I.C., V.Y., Y.S., C.B., K.T., H.U., D.O. Drafting and Critical revision of the manuscript: M.R., C.B., A.B., S.C., A.P.A., A.C.K., J.E.E., P.C., S.C., D.O. Statistical analysis: M.R., H.I., D.O. Obtained funding: and Supervision: D.O.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41746-025-02050-x>.

Correspondence and requests for materials should be addressed to David Ouyang.

Reprints and permissions information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025