

Aortic Stenosis Severity Classification: A Diagnostic Quandary with an Artificial Intelligence Solution

Hema Krishna^{1,2}; Wouter Ouwerkerk³, Josh H Arnold^{1,2}, Siddharth Bhayani¹, Matthew Frost⁴, Zhubo Jiang⁴, Cyril Equilbec⁴, Carolyn SP Lam^{4,5,6}, Patricia A. Pellikka⁷, Sanjiv Shah⁸, Dawood Darbar^{1,2}, Mayank Kansal^{1,2}

¹University of Illinois Chicago, Chicago, IL, USA, ²Jesse Brown VA Medical Center, Chicago, IL, USA ³Dept of Dermatology, Amsterdam UMC, Amsterdam, Netherlands ⁴Us2.ai, Singapore, ⁵National Heart Centre Singapore, Singapore, ⁶Duke-NUS Medical School, Singapore, ⁷Mayo Clinic, Rochester, MN, USA, ⁸Bluhm Cardiovascular Institute, Northwestern University, Chicago, IL, USA

BACKGROUND / OBJECTIVES

- Aortic stenosis (AS) is a life-threatening valvular heart condition affecting 7.6 million patients in the US and Europe¹.
- Severity grading by transthoracic echocardiography (TTE) is technically complex and time consuming, particularly in low stroke volume states. Maximally transparent artificial intelligence may offer a solution.
- Here, we demonstrate the capacity of two convolutional neural network (CNN)-augmented AS classifier methodologies to grade AS severity and predict clinical outcomes trained and compared against cardiologist classification.

METHODS

- Study Population: University of Illinois at Chicago AS Registry-Retrospective cohort of 954 adults (≥ 18 years) with any grade of AS 2005-2023. CNN automated 2D/Doppler measurements made on unlabeled complete TTEs (Us2.ai, Singapore). Measurements then used to classify severity by 2 methods:

 - Deterministic Algorithm:** Guideline-based rules determine classification (Table 1).
 - Gradient Boosted Model (GBM):** GBM trained/tested (70/30 split) with CNN measurements, against cardiologist interpretation, to output severity classification.

RESULTS

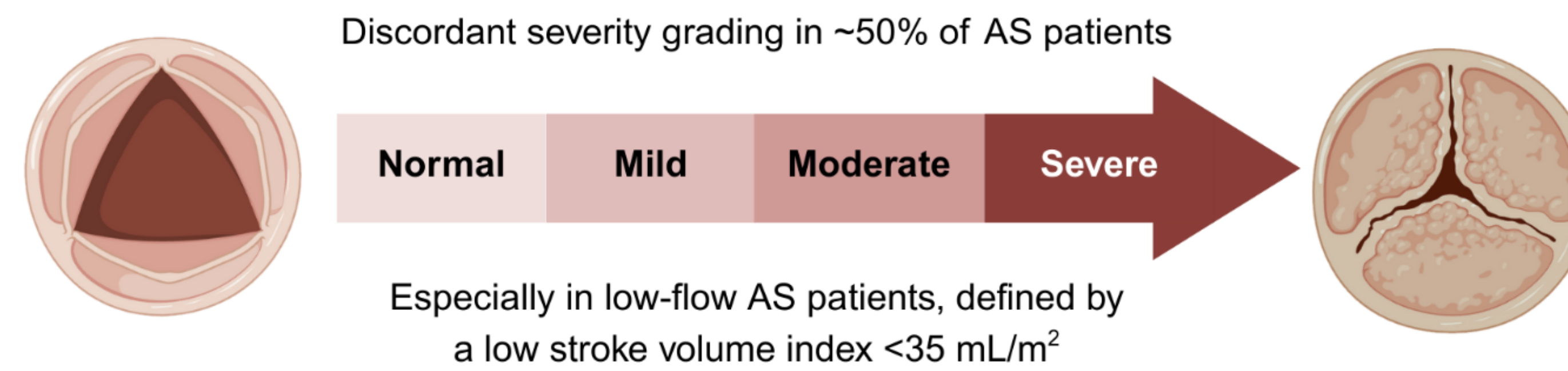
- Deterministic Algorithm n=682 with all 4 parameters
- GBM: n = 537 train/215 test (stats from test cohort)
- Low-flow, or stroke volume index < 35 mL/m² n=250 (26%).

Table 1: CNN Deterministic Algorithm Rules of Classification

Aortic Valve Area	Mean Pressure Gradient	Peak Velocity	Dimensionless Index	AS Severity
< 1	≥ 40	≥ 4		Severe
≥ 1	≥ 40	≥ 4		Severe
< 1	< 40	≥ 4		Severe
< 1	≥ 40	< 4		Severe
< 1	< 40	< 4	≤ 0.25	Severe
< 1	< 40	< 4	> 0.25	Moderate
> 1	< 40	≥ 4		Moderate
> 1	> 40	< 4		Moderate
1 - 1.5	20 - 40	3 - 4		Moderate
> 1.5	< 20	< 3		Mild
		< 2.5		None

RESULTS

A. Problem: Guideline-Based Criteria Produces Discordance in AS Severity Grading

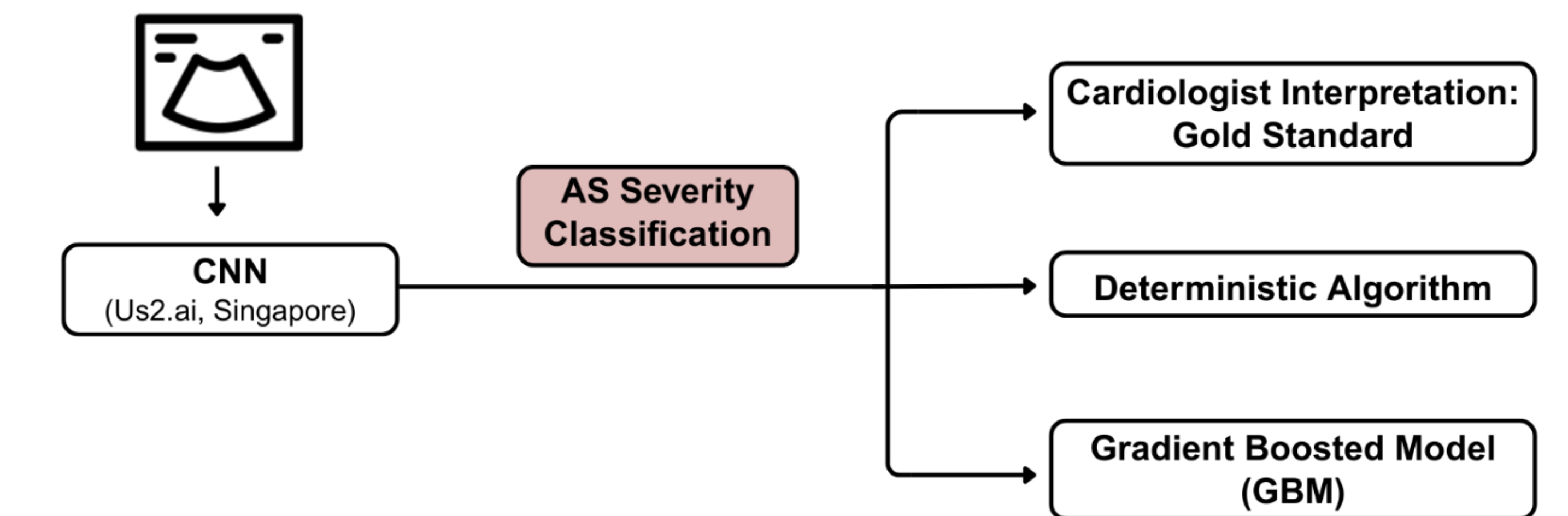


C. Model Performance

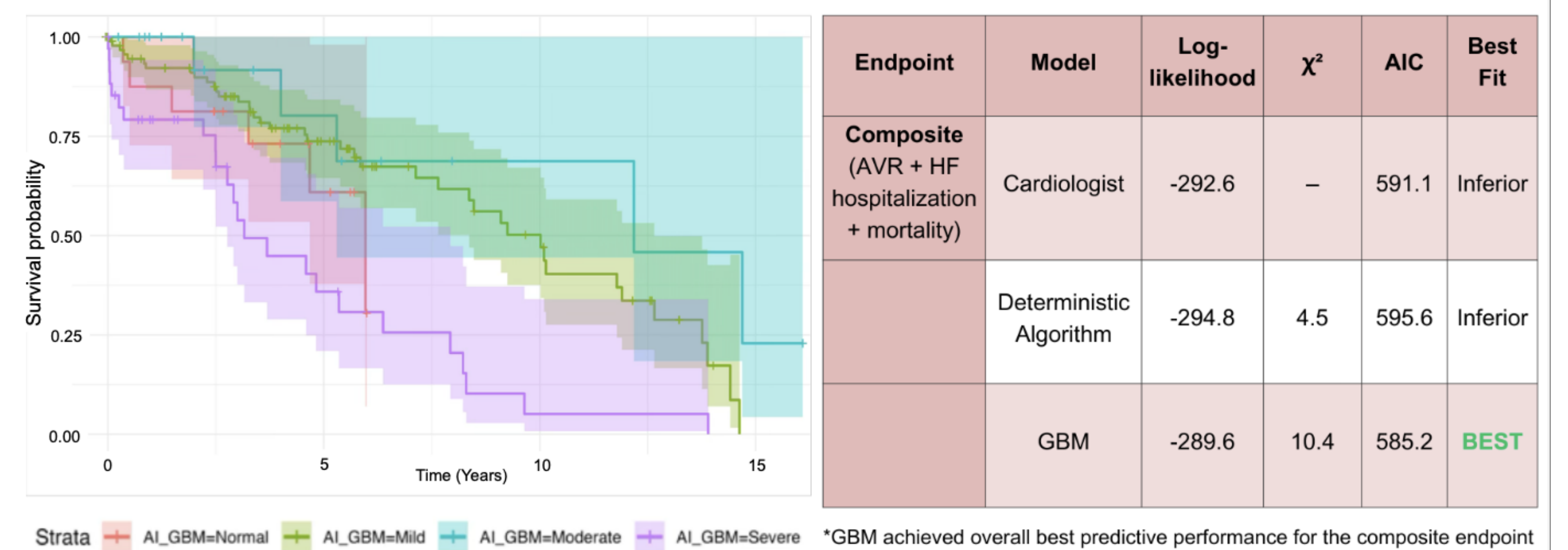
Performance of CNN Models Against Cardiologist Ground Truth AS Classification

Model	Severity	Sensitivity (%)	Specificity (%)	Accuracy (%)
Deterministic Algorithm in All AS (n = 682)	GT Mod	75	60	48 (44-52)
	GT Severe	63	96	
Deterministic Algorithm in Low-Flow AS (n= 250)	GT Mod	78	47	44 (38-50)
	GT Severe	74	91	
GBM in All AS (n = 215)	GT Mod	41	93	74 (67-79)
	GT Severe	82	96	
GBM in Low-Flow AS (n= 65)	GT Mod	50	94	75 (63-85)
	GT Severe	87	93	

B. Solution: CNN-augmented Classifier Methodologies to Grade AS Severity



D. Model Performance Across Clinical Endpoints



1. GBM AS Severity Classification is Accurate in Normal and Low-Flow AS

2. GBM Classification Outperforms Cardiologist in Prediction of Composite Endpoints

SUMMARY / CONCLUSION

- Building upon our prior work², we show here that a GBM utilizing CNN TTE measurements is accurate in grading any-flow (74%) and low-flow (75%) AS, against cardiologist grade.
- GBM classification was superior in predicting a composite of relevant clinical endpoints, in a retrospective data set.
- The GBM may reduce time and improve reproducibility in AS severity classification. Model training against outcomes is warranted as the next step.

REFERENCES

- Osnabrugge RLJ, Mylotte D, Head SJ, et al. Aortic stenosis in the elderly: Disease prevalence and number of candidates for transcatheter aortic valve replacement: A meta-analysis and modeling study. *J Am Coll Cardiol.* 2013;62(11):1002-1012. doi: 10.1016/j.jacc.2013.05.015.
- Krishna H, Desai K, Slostad B, et al. Fully automated artificial intelligence assessment of aortic stenosis by echocardiography. *J Am Soc Echocardiogr.* 2023;36(7):769-777. doi: 10.1016/j.echo.2023.03.008.



Uncharted Waters: Examining Prevalence and Prognosis of Low-Gradient Aortic Stenosis in Women and Minority Populations

J. H. Arnold¹, D. Smith¹, S. Bhayani¹, M. Frost², D. Tiu¹, S. Hsu¹, J. Sweeney¹, S. Sandhu¹, V. Alluri¹, P. A. Pellikka³, D. Darbar¹, Bavishi A¹, M. Kansal¹, H. Krishna¹

¹University of Illinois Chicago, Chicago, IL, USA, ²Us2.ai, Singapore, ³Mayo Clinic, Rochester, MN, USA

PURPOSE / OBJECTIVES

- Low-gradient aortic stenosis (LGAS)—defined by reduced aortic valve area (AVA) with low mean gradients—includes classical and paradoxical low-flow subtypes.
- While Black and Hispanic/Latinx patients undergo aortic valve replacement (AVR) at lower rates than White patients, LGAS patterns and outcomes in these populations remain poorly described.

MATERIAL & METHODS

- Study Population: Adults (>=18 years) with aortic stenosis (AS) undergoing transthoracic echocardiography (TTE) at the University of Illinois Chicago.

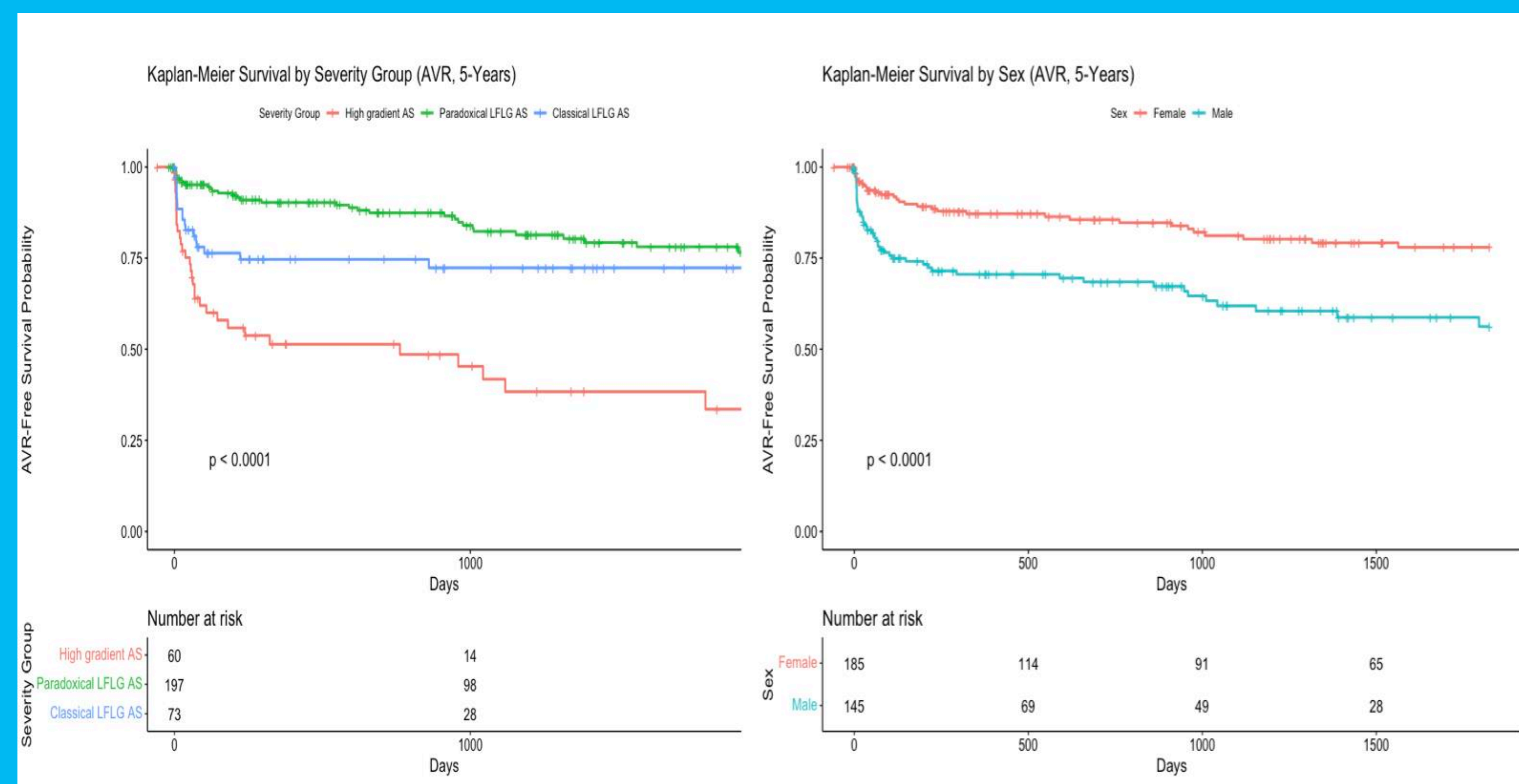
Imaging Analysis: CNN (US2.ai, Singapore) was used to derive AS severity and flow states (Table 1).

- Definitions: Moderate AS = AVA 1–1.5 cm²; Severe AS = AVA <1 cm². LGAS, MG (Severe <40 mmHg; Moderate <20 mmHg), and stratified by SVi <35 ml/m² and LVEF < 50.
- Outcomes: AVR, HF hospitalization, and all-cause mortality determined via blinded retrospective chart review.

RESULTS

- Among 968 patients with AS, 334 (37.8%) had moderate (AVA 1-1.5 cm², 49.1%) and severe (AVA <1 cm², 51.9%) disease.
- AVR, heart failure hospitalization, mortality, were similar across racial groups. Women disproportionately exhibited paradoxical LFLG AS (p=0.001) and were significantly less likely to receive AVR (p=0.001).
- LG AS patients were significantly less likely to undergo AVR compared to high-gradient AS patients (p<0.001). Heart failure hospitalizations were highest in classical LFLG AS.

CNN-derived low-gradient aortic stenosis was equally prevalent across races, more common in women—who received fewer valve replacements—despite similar outcomes to high-gradient disease.



RESULTS

Table 1

		AVA <1 cm ²			p-value	AVA 1-1.5 cm ²			p-value
		High Gradient AS: AVA <1-1.5 and MG >20-40 N = 46	Paradoxical LFLG AS: AVA <1-1.5, MG <40, SVi <35, LVEF >50% N = 87	Classical LFLG AS: AVA <1-1.5, MG <40, SVi <35, LVEF <50% N = 37		High Gradient AS: AVA <1-1.5 and MG >20-40 N = 16	Paradoxical LFLG AS: AVA <1-1.5, MG <40, SVi <35, LVEF >50% N = 116	Classical LFLG AS: AVA <1-1.5, MG <40, SVi <35, LVEF <50% N = 32	
Sex	Female	26 (56.5)	58 (66.7)	13 (35.1)	0.005	4 (25.0)	71 (61.2)	15 (46.9)	0.014
	Male	20 (43.5)	29 (33.3)	24 (64.9)		12 (75.0)	45 (38.8)	17 (53.1)	
Race/Ethnicity	HL	13 (28.3)	32 (36.8)	15 (40.5)	0.453	5 (31.3)	27 (23.3)	5 (15.6)	0.484
	Non-HL Black or AA	14 (30.4)	24 (27.6)	14 (37.8)		4 (25.0)	47 (40.5)	16 (50.0)	
	Non-HL White	12 (26.1)	16 (18.4)	6 (16.2)		6 (37.5)	27 (23.3)	9 (28.1)	
	Other	7 (15.2)	15 (17.2)	2 (5.4)		1 (6.3)	15 (12.9)	2 (6.3)	
Atrial Fibrillation/Flutter		6 (13.0)	11 (12.8)	10 (27.0)	0.116	1 (6.3)	23 (19.8)	9 (28.1)	0.202
Hypertension		18 (39.1)	30 (34.9)	18 (48.6)	0.357	8 (50.0)	56 (48.3)	17 (53.1)	0.888
Heart Failure		9 (19.6)	6 (7.0)	11 (29.7)	0.004	2 (12.5)	10 (8.6)	5 (15.6)	0.494
AVR		19 (41.3)	18 (20.9)	11 (29.7)	0.046	10 (62.5)	15 (12.9)	2 (6.3)	<.001
HF Hospitalization		12 (27.9)	16 (19.5)	15 (40.5)	0.054	3 (21.4)	25 (21.6)	8 (25.0)	0.915
Mortality	Alive	37 (80.4)	64 (73.6)	28 (75.7)	0.678	13 (81.3)	89 (76.7)	25 (78.1)	0.916
	Deceased	9 (19.6)	23 (26.4)	9 (24.3)		3 (18.8)	27 (23.3)	7 (21.9)	

AA: African American, AVA: Aortic Valve Area (cm²), AVR: Aortic Valve Replacement, HF: Heart Failure, HL: Hispanic/Latinx, LF: Low Flow, LG: Low Gradient, LVEF: Left Ventricular Ejection Fraction MG: Mean Gradient (mmHg), SVi: Stroke Volume Index (ml/m²).

SUMMARY / CONCLUSION

- Racial disparities in LGAS do not account for AVR gaps in underserved populations
- Significantly higher rates of LGAS exist in women, potentially contributing to reduced AVR rates
- LGAS patients experience similar rates of heart failure hospitalization and death across subtypes compared to high-gradient AS patients, despite lower AVR utilization.



Convolutional Neural Network-Derived Left Atrial Reservoir Strain, Left Ventricular Global Longitudinal Strain, & Diastolic Function Predict Heart Failure Hospitalization and Time to Aortic Valve Replacement in Patients with Aortic Stenosis

S. Bhayani¹, J.H Arnold¹, M. Frost², D. Tiu¹, S. Hsu¹, J. Sweeney¹, S. Sandhu¹, V. Alluri¹, P. Pellikka³, D. Darbar¹, A. Bavishi¹, M. Kansal¹, H. Krishna¹

¹University of Illinois Chicago, Chicago, IL, USA, ²Us2.ai, Singapore, ³Mayo Clinic, Rochester, MN, USA

PURPOSE / OBJECTIVES

- Left ventricular GLS and diastolic function are key prognostic indicators in aortic stenosis due to pressure-induced remodeling.
- Left atrial reservoir strain (LARS) is a promising but underused marker for predicting adverse cardiovascular events.
- This study evaluates CNN-derived diastolic and strain metrics in relation to heart failure hospitalization and time to AVR across all severities of aortic stenosis.

MATERIAL & METHODS

- Patients with aortic stenosis aged >18 years who underwent clinically indicated transthoracic echocardiography (TTE) at the University of Illinois Chicago were included.
- Retrospective chart review, blinded to other data, was conducted to determine clinical endpoints.
- Measurements were performed on unlabeled TTEs using the Us2.ai convolutional neural network (Singapore), and statistical analysis was conducted using R.

RESULTS

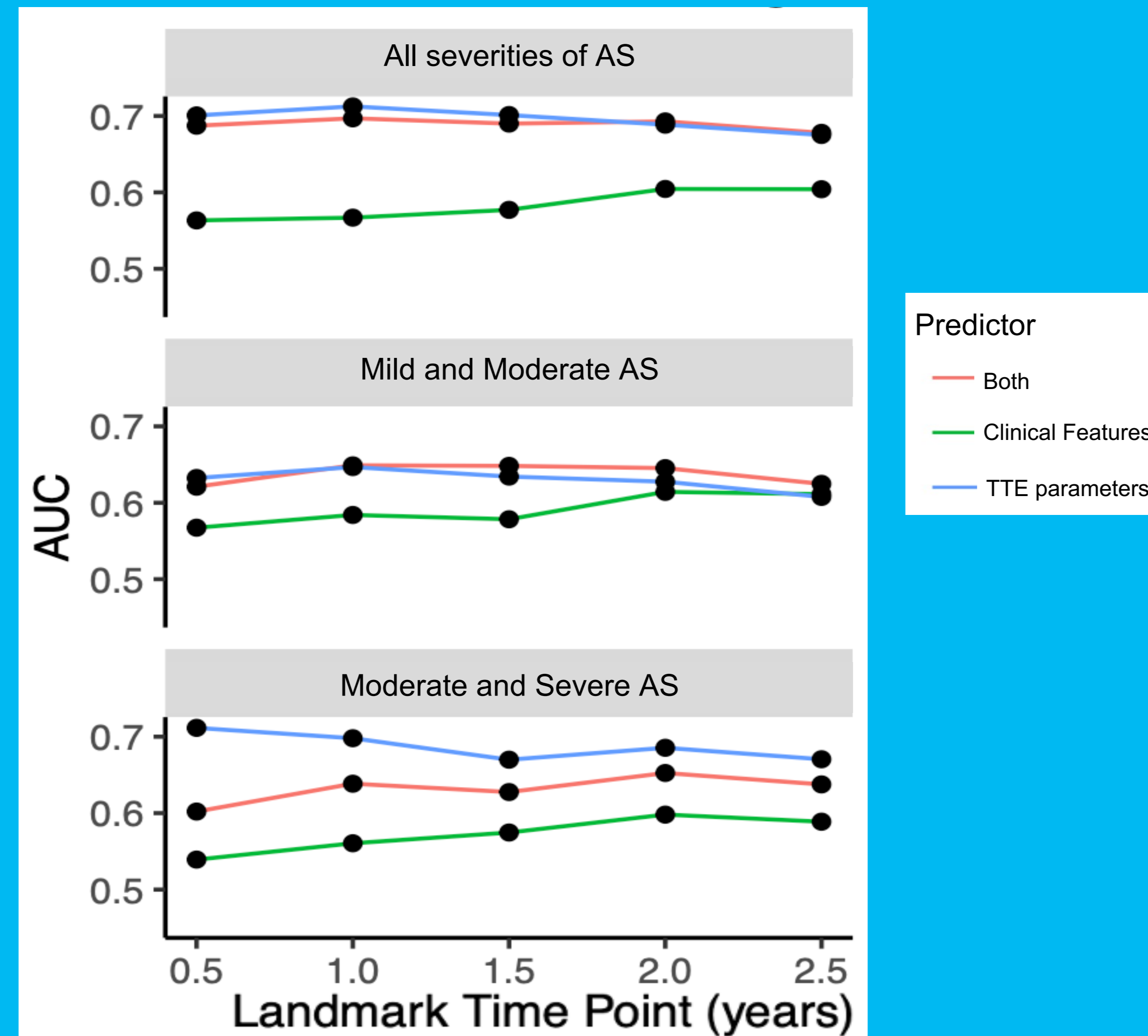
Variable	1 month followup			6 month followup			12 month followup		
	CHF Hospitalization N = 53	No CHF Hospitalization N = 824	p-value ²	CHF Hospitalization N = 99	No CHF Hospitalization N = 778	p-value ²	CHF Hospitalization N = 114	No CHF Hospitalization N = 763	p-value ²
LV GLS	-14.4 ± 4.5	-16.8 ± 4.4	0.004	-15.0 ± 4.7	-16.8 ± 4.4	0.003	-15.3 ± 4.6	-16.8 ± 4.4	0.006
E/e' lateral	15.4 ± 6.2	11.8 ± 5.7	0.001	14.2 ± 6.0	11.7 ± 5.7	0.002	14.1 ± 6.0	11.7 ± 5.7	<0.001
E/e' septal	20 ± 8	15 ± 7	<0.001	18 ± 8	15 ± 7	0.003	17 ± 7	15 ± 7	0.004
TR Vmax	3.14 ± 0.95	2.65 ± 0.70	0.002	2.91 ± 0.84	2.65 ± 0.71	0.010	2.92 ± 0.82	2.65 ± 0.71	0.003
LA Reservoir	19 ± 10	21 ± 9	0.13	19 ± 9	21 ± 9	0.076	19 ± 9	21 ± 9	0.076
E/A Ratio	1.34 ± 0.56	1.03 ± 0.39	0.002	1.25 ± 0.51	1.03 ± 0.39	<0.001	1.21 ± 0.49	1.03 ± 0.40	0.002
e' septal	6.17 ± 1.98	7.33 ± 2.71	<0.001	6.60 ± 2.33	7.34 ± 2.72	0.012	6.68 ± 2.44	7.34 ± 2.71	0.020
e' lateral	8.03 ± 2.56	9.30 ± 3.08	0.004	8.22 ± 2.77	9.35 ± 3.08	<0.001	8.19 ± 2.75	9.38 ± 3.08	<0.001
LAVI	39 ± 13	35 ± 16	0.13	38 ± 12	35 ± 16	0.069	39 ± 14	35 ± 16	0.027

¹ Mean ± SD

² Welch Two Sample t-test

CNN-derived diastolic parameters and LV GLS demonstrate prognostic value in identifying aortic stenosis patients at elevated risk for adverse outcomes.

Figure 1. Performance of CHF admission prediction models across AS severities



RESULTS

- A total of 877 patients were included (51% male, mean age 79 ± 15 years, 24% with LVEF <50%), comprising 443 mild, 226 moderate, and 208 severe AS cases.
- TTE markers indicating impaired myocardial relaxation, elevated left atrial pressure, elevated PASP, and reduced LV longitudinal function were associated with HF hospitalizations at 1, 6, and 12 months; LAVI was associated with HF hospitalizations at 12 months.
- A trend toward significance was observed for LARS in predicting HF hospitalizations.
- Time to AVR was shorter in patients with worse LV GLS, higher E/e' (lateral and septal), higher TR Vmax, and lower e' septal; conversely, higher LARS and higher e' lateral were associated with longer time to AVR.
- Figure 1 demonstrates the AUC of predictive models (utilizing clinical features and CNN-derived TTE parameters) for HF exacerbations across different severities of AS

SUMMARY / CONCLUSION

- This study in a cohort with varying degrees of AS is the first to demonstrate the prognostic value of CNN-derived diastology and LV GLS in identifying AS patients at increased risk of adverse outcomes.
- These techniques may allow for faster, more sophisticated, and more reproducible risk discrimination to identify AS patients deriving benefits from earlier AVR.