ESC Congress
Late-Breaking Science
Munich 2026

Automated Detection of Pulmonary Hypertension from a Single Apical Four-chamber Echocardiographic Clip

Development on the largest RHC-confirmed cohort to date and multinational external validation

Fawaz Alenezi, MD, FAHA, FACC, FASE
Associate Professor of Medicine
Program Director of Structural Heart Disease
Duke Heart Center
Duke University Medical Center

DukeHealth

Disclosures

PRESENTER DISCLOSURES

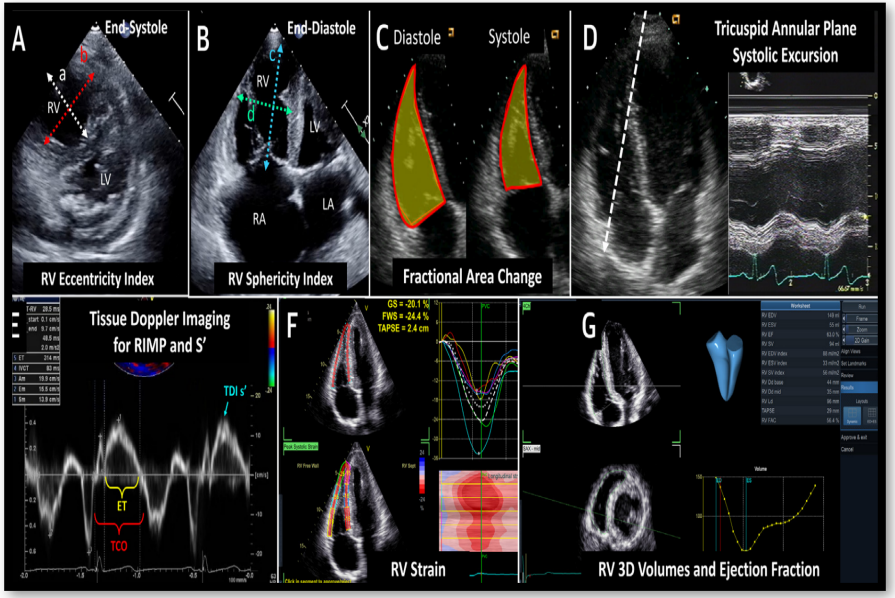
- **Research funding (unrelated to this project)**
  National Institutes of Health (NIH); American Heart Association (AHA) Career Development Award; United Therapeutics.

- **Relationship to this study**
  None of the above sources funded or supported this specific project.

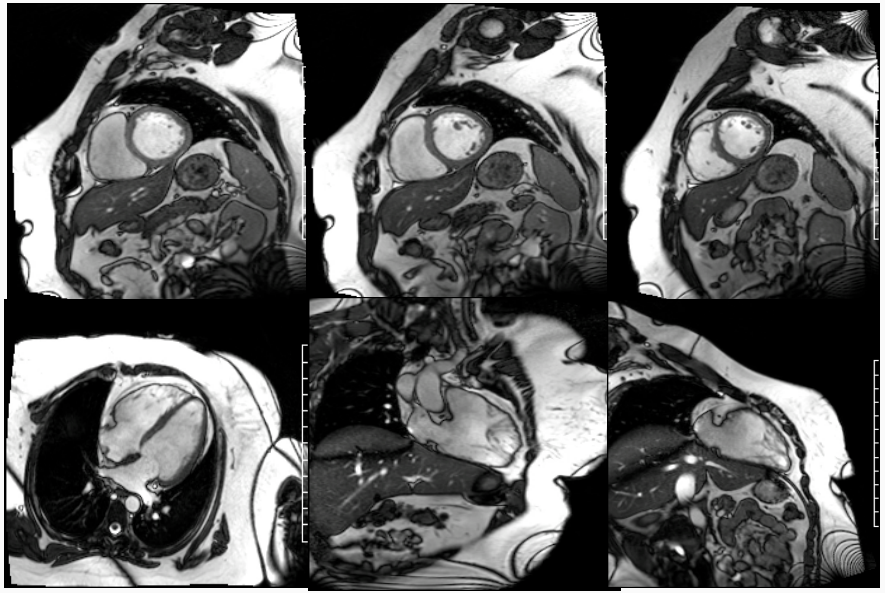
- **Study-specific funding and conflicts**
  This study received no external funding. Technology partner Us2.ai (Singapore) contributed model development and preprocessing infrastructure. No other relevant financial relationships to disclose.

Analysis is expertise-dependent

Echo: many measurements, many views



CMR: gold standard for RV, limited access



Diagnostic delay in pulmonary hypertension

≥ 2 yrs

MEDIAN DIAGNOSTIC DELAY

from symptom onset to a formal PH diagnosis

1. **Symptoms are non-specific**
   Dyspnea and fatigue overlap with common cardiac and pulmonary disease; referral is often late.

2. **Echocardiographic screening is variable**
   Depends on an adequate TR jet, multi-view acquisition, and experienced interpretation.

3. **Cardiac MRI is not widely available**
   Right ventricular remodeling appears late in disease; limited availability and specialist expertise required.

4. **Right-heart catheterization is invasive**
   Not suitable as a first-line screening test.

Clinical need: a simple, scalable way to flag patients who need expert review and confirmatory RHC.

Study objectives

Can a video-based transformer detect PH from one complete heartbeat on a single apical four-chamber view, without Doppler, structured measurements, or a second view?

Four pre-specified objectives:

1 Detect

Develop and internally validate the model against RHC-confirmed PH.

2 Validate

Across three external cohorts spanning continents, vendors, and workflows.

3 Sensitivity to PH severity

Assess whether performance scales with disease severity (mild to severe).

4 Differentiate subtypes

Assess whether performance differs across hemodynamic PH subtypes.

METHODS · COHORTS

Development and external validation cohorts

Reference standard: 2022 ESC/ERS PH definition, RHC-confirmed resting mPAP > 20 mmHg

In collaboration with Us2.ai (Singapore), whose platform underlies the deep learning model and automated echo preprocessing

DEVELOPMENT

Duke

17,969

patients

8,644 PH · 9,325 controls

80 % train · 10 % val · 10 % locked test

EXTERNAL VALIDATION · NO RETRAINING

France — n = 576

Kremlin-Bicêtre · Tertiary PH referral center

Stanford — n = 830

Independent US academic center

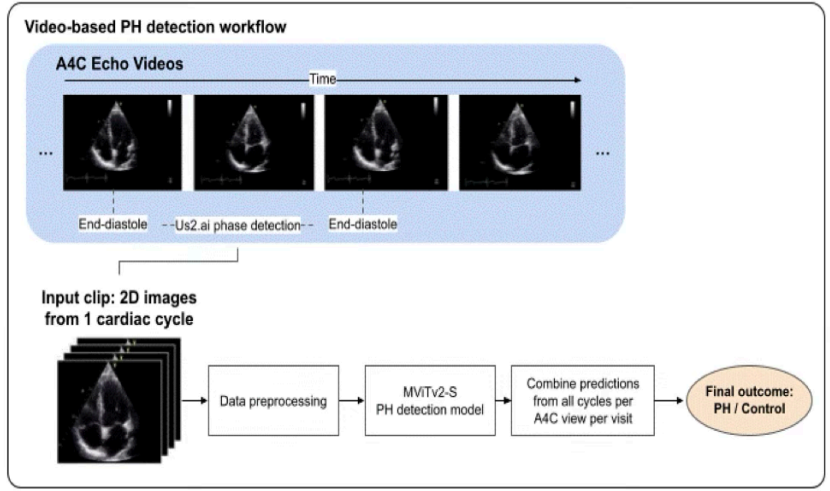
China — n = 469

CardiacNet public dataset · Hitachi vendor absent from training

Each external cohort differed from Duke on at least one axis: geography, patient mix, vendor, or acquisition protocol.

Model architecture and inference pipeline

1. A4C clip identified.
2. Segmented into individual cardiac cycles.
3. 16 frames sampled per cycle.
4. Cycle predictions averaged into one PH probability.



Predictions from every cycle across every A4C loop are averaged into one patient-level PH probability

✕ No Doppler

✕ No structured measurements

✕ No second view

Study objectives

Can a video-based transformer detect PH from one complete heartbeat on a single apical four-chamber view, without Doppler, structured measurements, or a second view?

Four pre-specified objectives:

1 Detect

Develop and internally validate the model against RHC-confirmed PH.

2 Validate

Validate across three external cohorts spanning continents, vendors, and workflows.

3 Sensitivity to PH severity

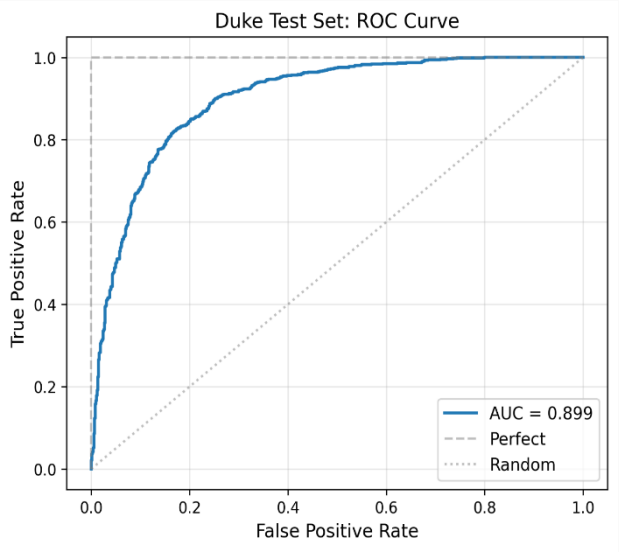
Assess whether performance scales with disease severity (mild to severe).

4 Differentiate subtypes

Assess whether performance differs across hemodynamic PH subtypes.

Can one A4C clip detect RHC-confirmed PH?

Locked Duke internal test set (n = 1,808 · 862 PH, 946 controls)



Duke internal test set · ROC curve

AUROC

0.899

Sensitivity
82%
707 / 862 PH

Specificity
81.7%
711 / 946 controls

PPV
80.2%
in flagged studies

NPV
83.0%
in non-flagged studies

Study objectives:

Can a video-based transformer detect PH from one complete heartbeat on a single apical four-chamber view, without Doppler, structured measurements, or a second view?

Four pre-specified objectives:

1 **Detect**

Develop and internally validate the model against RHC-confirmed PH.

2 **Validate**

Across three external cohorts spanning continents, vendors, and workflows.

3 **Track severity**

Assess whether performance scales with disease severity (mild to severe).

4 **Track phenotype**

Assess whether performance differs across hemodynamic PH subtypes.

Does the model generalize across health systems and vendors?

France	Stanford	China
Kremlin-Bicêtre	*Independent US academic center*	*CardiacNet, unseen vendor*
n = 576	**n = 830**	**n = 469**
AUROC	AUROC	AUROC
0.917	**0.732**	**0.871**
Tertiary PH referral cohort in Europe; higher prevalence and severity than the general echo lab	*Tertiary PH referral cohort in the United States; Relatively lower AI performance*	*Hitachi ultrasound vendor absent from training; a strictly out-of-training-distribution test.*

Interpretation: Discrimination is meaningful in all three settings including unseen ultrasound vendor.

Study objectives:

Can a video-based transformer detect PH from one complete heartbeat on a single apical four-chamber view, without Doppler, structured measurements, or a second view?

Four pre-specified objectives:

1 **Detect**

Develop and internally validate the model against RHC-confirmed PH.

2 **Validate**

Validate across three external cohorts spanning continents, vendors, and workflows.

3 **Sensitivity to PH severity**

Assess whether performance scales with disease severity (mild to severe).

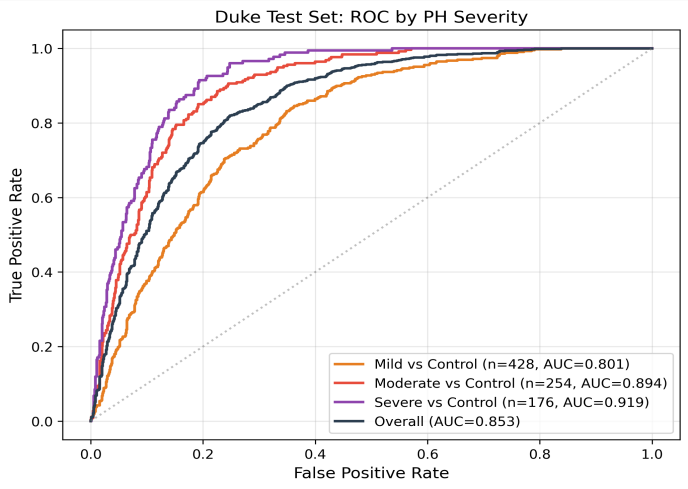
4 **Track phenotype**

Assess whether performance differs across hemodynamic PH subtypes.

Discrimination scales with hemodynamic severity

Classified as mild, moderate, or severe based on mPAP (mild: < 35 mmHg, moderate: 35-45 mmHg, severe: > 45 mmHg) and PVR (mild: 2-5 WU, moderate: 5-8 WU, severe: > 8 WU), with the higher-severity category assigned when both criteria were available and discordant.

DUKE

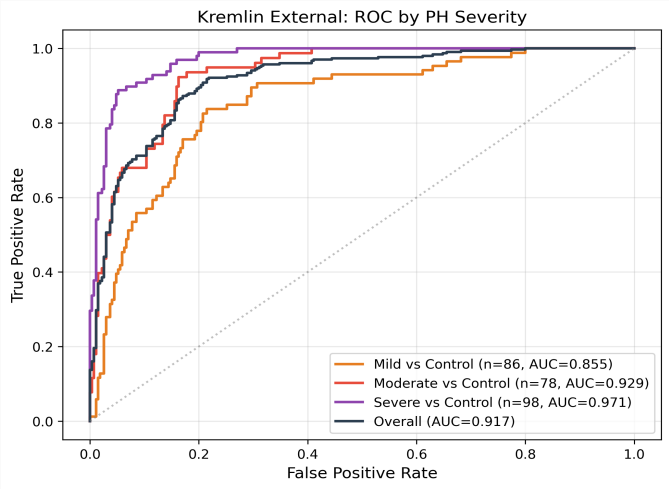


AUROC 0.801 — 0.894 — 0.919

Mild — *Moderate* — *Severe*

Discrimination increases with disease severity

KREMLIN-BICÊTRE (France)



AUROC 0.855 — 0.929 — 0.971

Mild — *Moderate* — *Severe*

Discrimination increases with disease severity

Study objectives

Can a video-based transformer detect PH from one complete heartbeat on a single apical four-chamber view, without Doppler, structured measurements, or a second view?

Four pre-specified objectives:

1 **Detect**

Develop and internally validate the model against RHC-confirmed PH.

2 **Validate**

Validate across three external cohorts spanning continents, vendors, and workflows.

3 **Track severity**

Assess whether performance scales with disease severity (mild to severe).

4 **Differentiate subtypes**

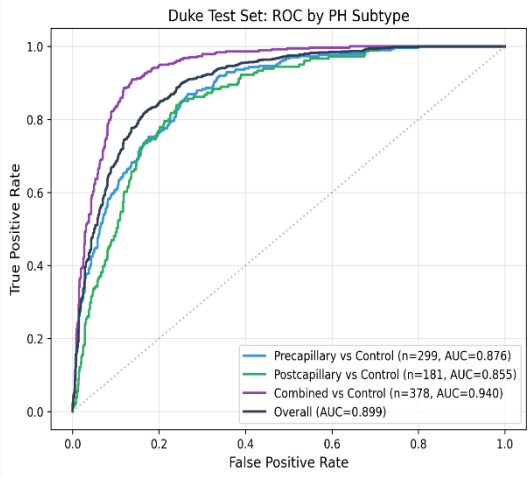
Assess whether performance differs across hemodynamic PH subtypes.

The model tracks right heart remodeling, not pressure elevation

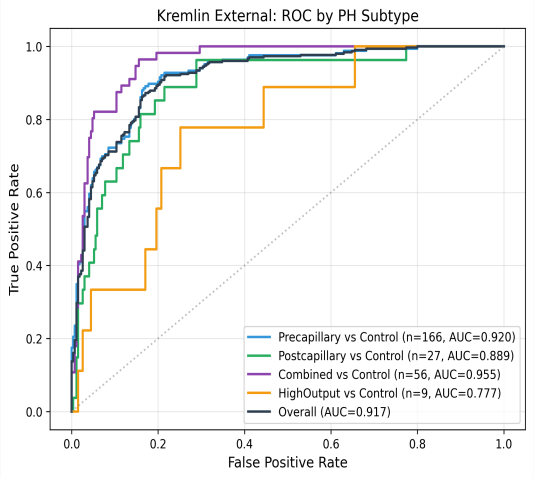
PH subtypes per 2022 ESC/ERS

AUC	Duke	France
Precapillary PH *PVR > 2 WU, PCWP ≤ 15 mmHg*	0.876	0.920
Isolated postcapillary PH *PVR ≤ 2 WU, PCWP > 15 mmHg*	0.855	0.889
Combined pre/ postcapillary PH *PVR > 2 WU, PCWP > 15 mmHg*	0.940	0.955
Unclassified PH *PVR ≤ 2 WU, elevated flow · n=0, Duke*	—	0.777

Duke



France



FINDING: Strong discrimination across phenotypic subtypes (AUC 0.86 – 0.96); reduced sensitivity of flow-driven high-output PH confirms the model is responding to imaging correlates of vascular pathology, not elevated pressure alone.

LIMITATIONS

1. Limited specificity.
34 % on CardiacNet, where the negative-class definition is unclear and qualitative review found abnormal RVs among nominal controls.

2. Controls not universally RHC-negative.
Mild PH may sit in the control group and inflate apparent false positives.

3. Binary PH training label.
We did not regress on mPAP directly, which would have enabled severity stratification at inference and may have captured clinically important subthreshold elevations.

4. Adult population only.
View selection is age-specific; pediatric PH favors PSAX rather than A4C.

KEY FINDINGS

1. A single A4C clip can screen for PH.
One routine clip, no Doppler, no measurements, reproducible across three continents. This could be of significant impact in resource-limited areas.

2. Validation across multiple cohorts.
Diverse cohorts from different regions and imaging platforms.

3. Differentiation of hemodynamic subtypes.
Clinically important, as not all patients with PH require RHC, such as those with isolated post-capillary PH.

4. Sensitivity to PH severity.
Discrimination increased with hemodynamic severity in both cohorts, supporting use as a severity aware triage signal.

AUTHORS

Fawaz Alenezi1, Cyril Equilbec2, Sreekanth Vemulapalli1, Matthew Frost2, Carolyn S. P. Lam3, Monica Mukherjee4, Francois Haddad5, Fabrice Bauer6, Sudarshan Rajagopal1

Underlined = presenting author

AFFILIATIONS

1 Division of Cardiology, Duke University School of Medicine, Durham, NC, USA
2 Us2.ai, Singapore
3 National Heart Centre Singapore; Duke–NUS Medical School, Singapore
4 Johns Hopkins University Division of Cardiology, Baltimore, MD, USA
5 Division of Cardiovascular Medicine, Stanford University, Palo Alto, CA, USA
6 Kremlin-Bicêtre University Hospital, AP-HP, Paris, France

Acknowledgement: **Technology partner –** Us2.ai (Singapore) – deep learning model development and automated echocardiographic preprocessing
Funding: None. **Ethics:** Duke IRB Pro00114294.

Questions?

Email: fawaz.enezi@dm.duke.edu