UCSF Health

# The future has already happened

James Pirruccello, MD

ACC Heart House

October 30, 2025

# Title attribution: William F Gibson (Author)

His second formulation:

“*The future is already here; it's just not evenly distributed.*”

# Atherosclerosis and beyond

We still want to address risk stratification and prevention for non-atherosclerotic cardiovascular disease:

- HF & cardiomyopathy (PREVENT is a start)
- Aortopathy
- Arrhythmia
- Valvular heart disease
- Stroke

… and to improve our ability to identify long-term atherosclerosis risk in all individuals, not just near-term risk in older men.

# Emerging technology & how we can use it

- Omics
  - Common variant genetics
  - Proteomics
  - Metabolomics
  - CRP, Lp(a), etc
- Digital data
  - ECG
  - Imaging
  - Text
- Machine learning / artificial intelligence
- Widespread availability of sand that can think

# THE AI ALIGNMENT CHART

| | ABILITY PURIST (EXCEEDS HUMAN ABILITY) | ABILITY NEUTRAL (MAKES TASK EASIER) | ABILITY REBEL (USEFULNESS QUESTIONABLE) |
|---|---|---|---|
| ALGORITHM PURIST (MIMICS HUMAN COGNITION) | "TERMINATOR IS AI" | "C3PO IS AI" | "WALL-E IS AI" |
| ALGORITHM NEUTRAL (LEARNS AND GENERALIZES) | "ALPHA GO IS AI" | "XGBOOST IS AI" | "TINDER IS AI" |
| ALGORITHM REBEL (METHOD IRRELEVANT) | "A METAL DETECTOR IS AI" | "BUBBLE SORT IS AI" | "MAGIC 8 BALL IS AI" |

https://x.com/rlmcelreath/status/1729106605621535022

# THE AI ALIGNMENT CHART

| | ABILITY PURIST (EXCEEDS HUMAN ABILITY) | ABILITY NEUTRAL (MAKES TASK EASIER) | ABILITY REBEL (USEFULNESS QUESTIONABLE) |
|---|---|---|---|
| ALGORITHM PURIST (MIMICS HUMAN COGNITION) | "TERMINATOR IS AI" | "C3PO IS AI" | "WALL-E IS AI" |
| ALGORITHM NEUTRAL (LEARNS AND GENERALIZES) | "ALPHA GO IS AI" | "XGBOOST IS AI" | "TINDER IS AI" |
| ALGORITHM REBEL (METHOD IRRELEVANT) | "A METAL DETECTOR IS AI" | "BUBBLE SORT IS AI" | "MAGIC 8 BALL IS AI" |

https://x.com/rlmcelreath/status/1729106605621535022

# THE AI ALIGNMENT CHART

| | ABILITY PURIST (EXCEEDS HUMAN ABILITY) | ABILITY NEUTRAL (MAKES TASK EASIER) | ABILITY REBEL (USEFULNESS QUESTIONABLE) |
|---|---|---|---|
| ALGORITHM PURIST (MIMICS HUMAN COGNITION) | "TERMINATOR IS AI" | "C3PO IS AI" | "WALL-E IS AI" |
| ALGORITHM NEUTRAL (LEARNS AND GENERALIZES) | "ALPHA GO IS AI" | "XGBOOST IS AI" | "TINDER IS AI" |
| ALGORITHM REBEL (METHOD IRRELEVANT) | "A METAL DETECTOR IS AI" | "BUBBLE SORT IS AI" | "MAGIC 8 BALL IS AI" |

https://x.com/rlmcelreath/status/1729106605621535022

# Common variant genetics

# Thoracic aortopathy as a motivating example

- 60% of the variance in ascending aortic diameter is attributable to common-variant genetics (heritability is 60%)

- Can conceptualize heritability as: “If you had a perfect polygenic score…”

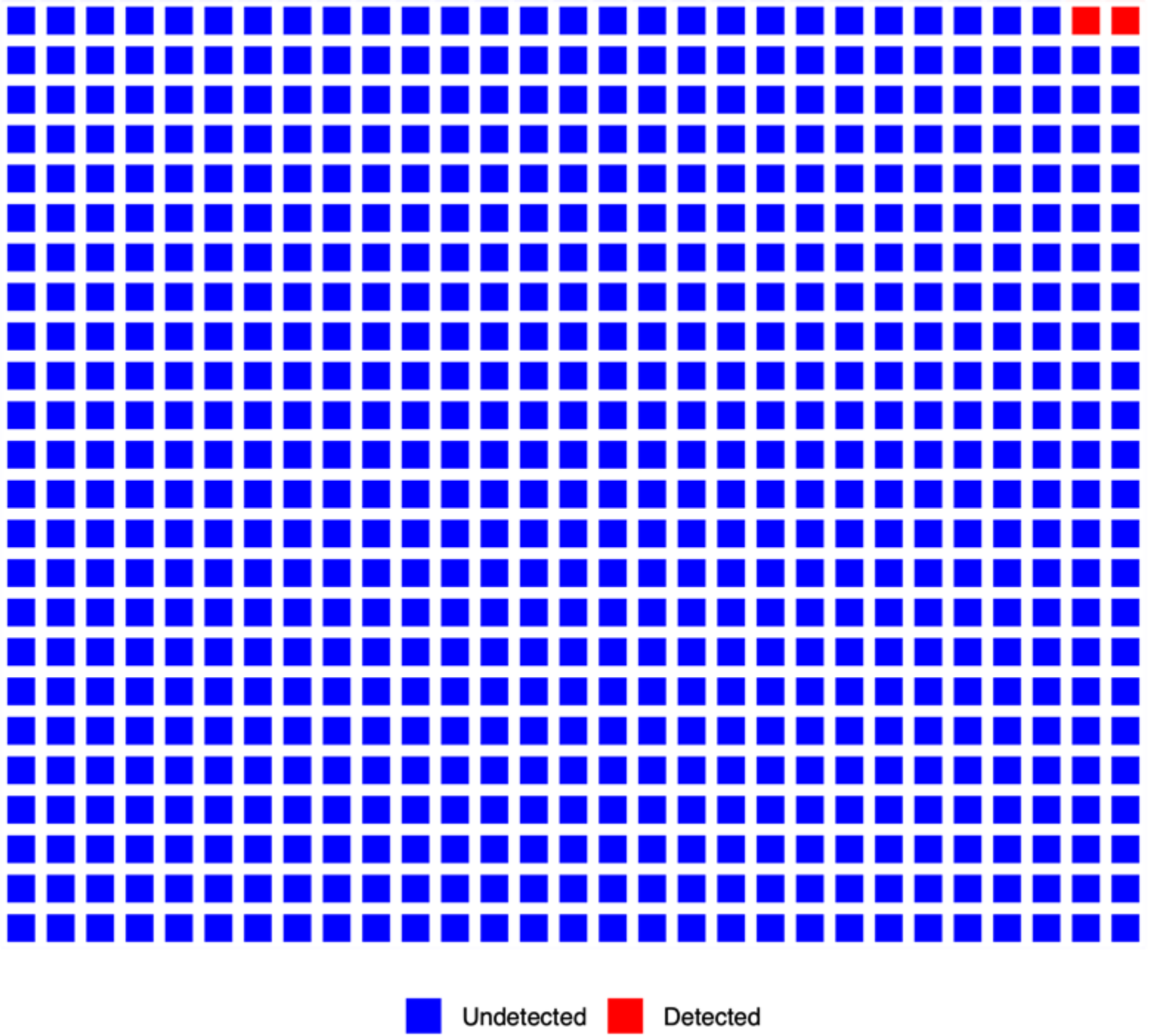
# Enlarged aortas: not widely documented

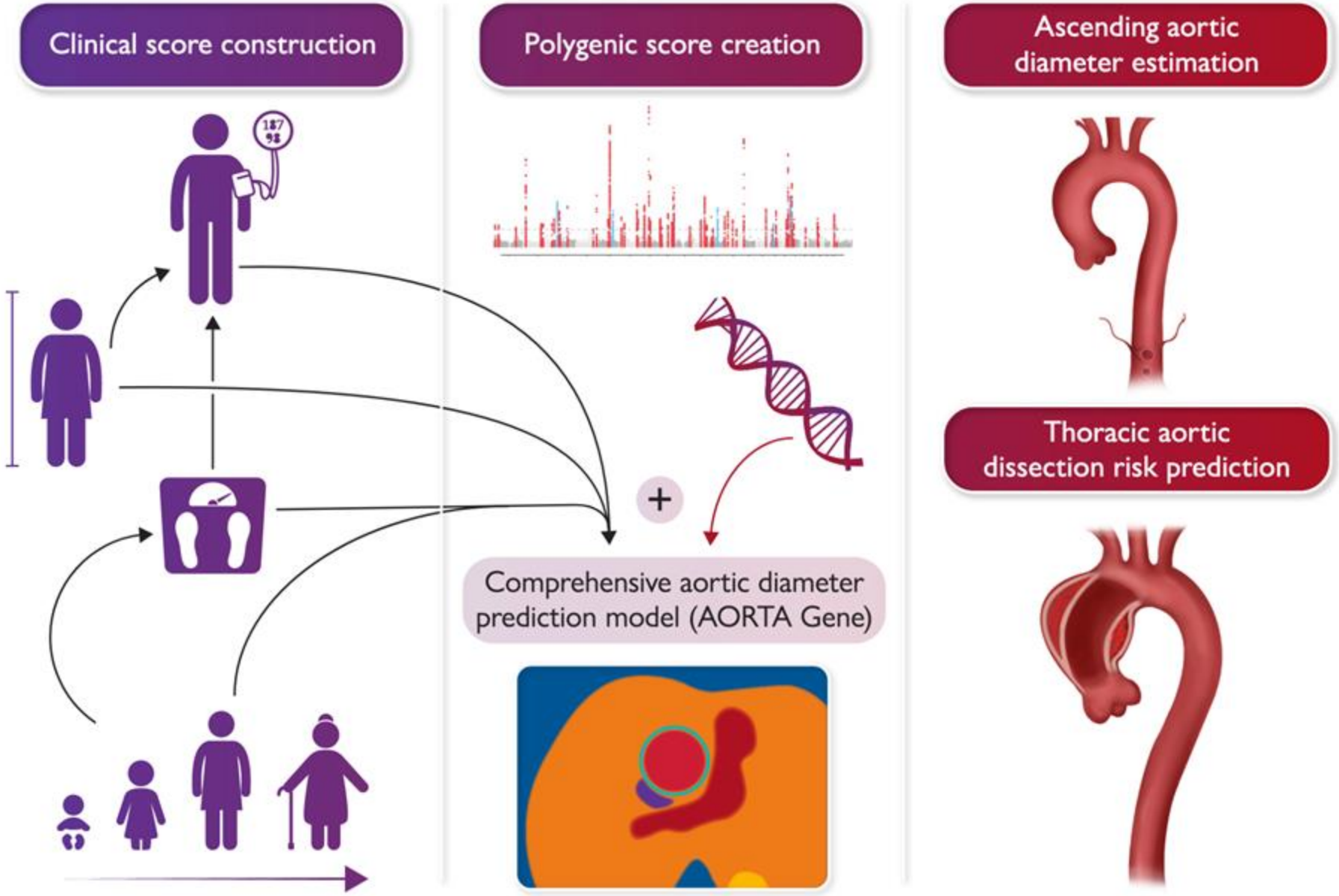
30,018 participants

696 (2.3%) w/ ascending aortic diameter ≥ 4cm

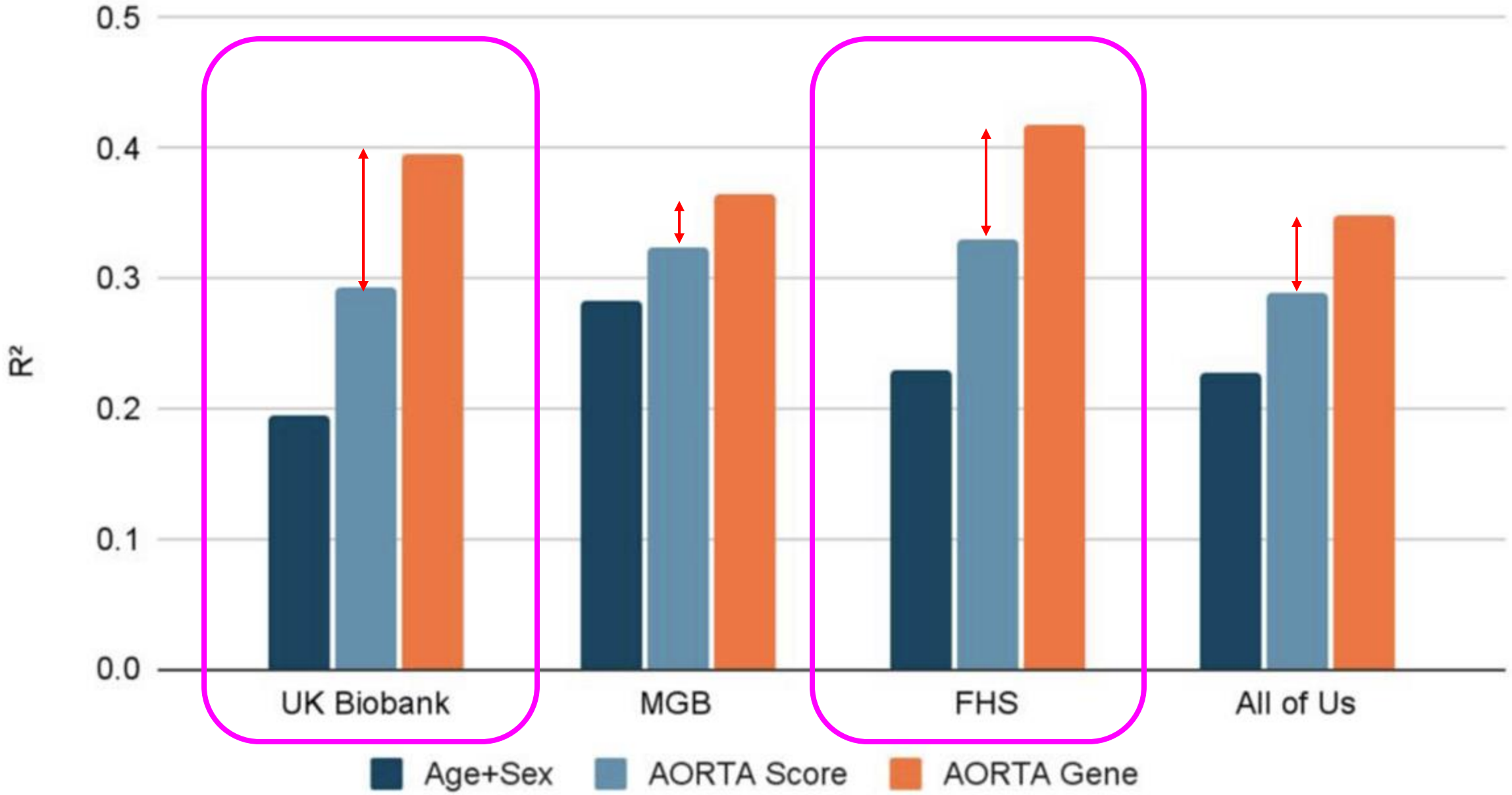
**2 (0.3%)** had diagnostic codes for thoracic aortic aneurysm prior to MRI

Aneurysm diagnosis in **0** of the N=45 w/diameter ≥4.5cm





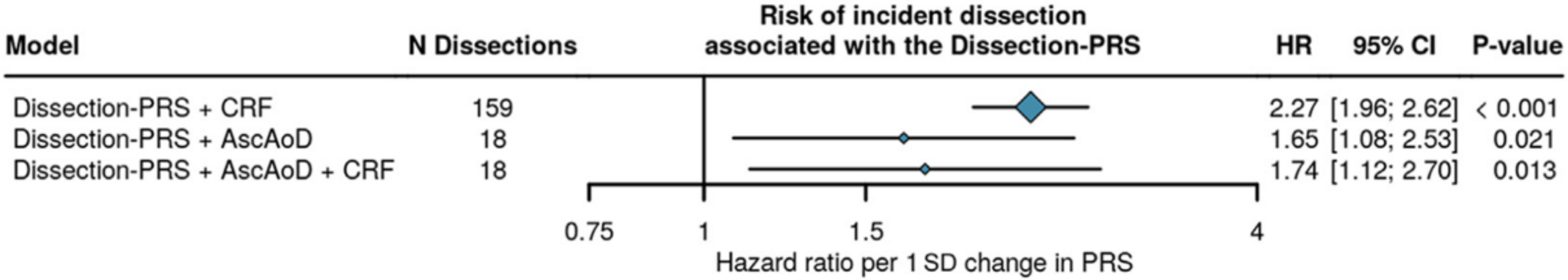

Pirruccello JP, *et al*. The AORTA Gene score for detection and risk stratification of ascending aortic dilation. *European Heart Journal*. 2024.



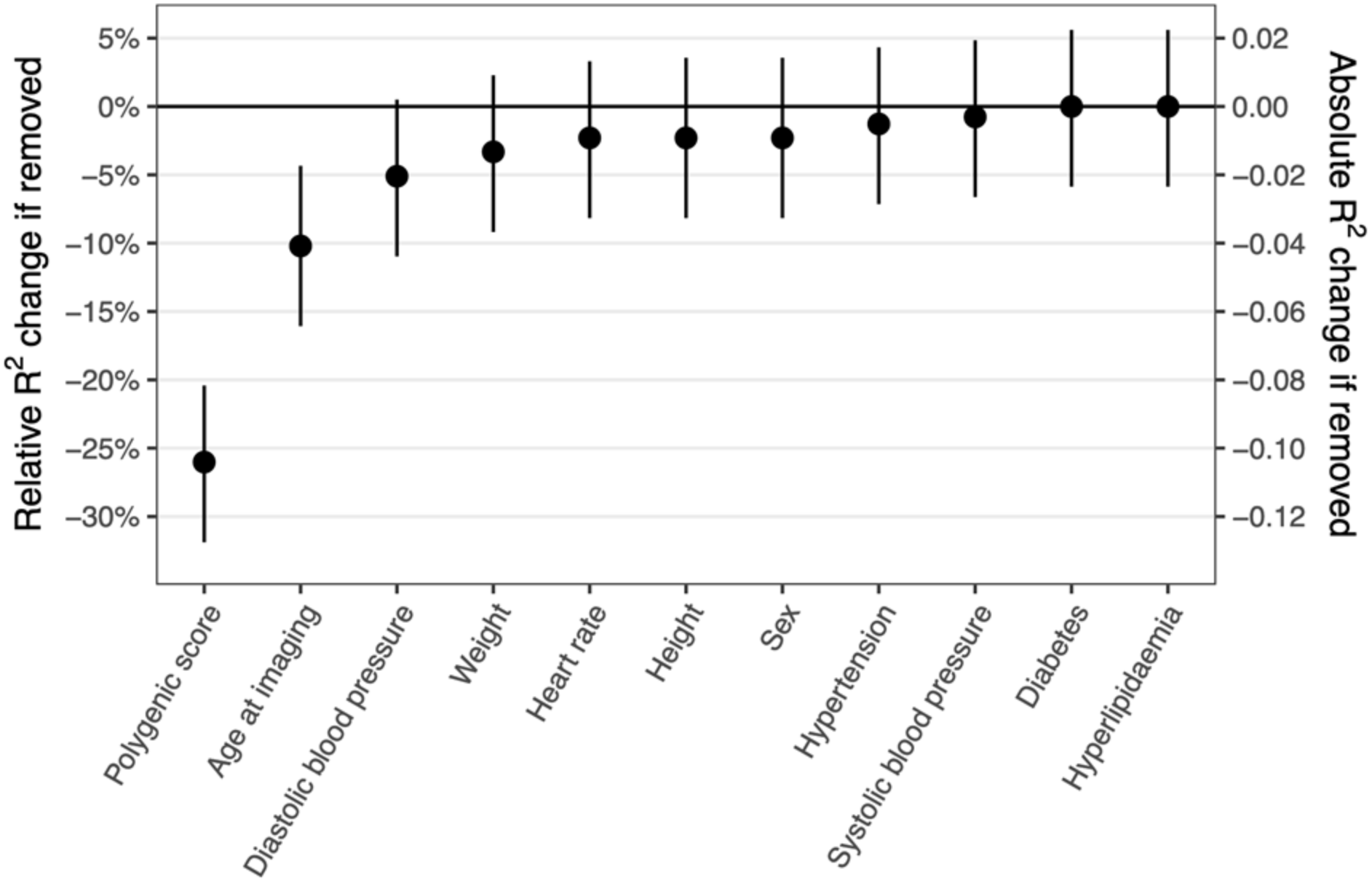

Pirruccello JP, *et al*. The AORTA Gene score for detection and risk stratification of ascending aortic dilation. *European Heart Journal*. 2024.

# Diameter-independent dissection risk: PRS HR 1.7 per SD *even after you have measured the aortic diameter*

**B**




DePaolo J, et al. Predicting Thoracic Aortic Dissection in a Diverse Biobank Using a Polygenic Risk Score. *JACC Adv*. 2025;4:101743.




Pirruccello JP, *et al*. The AORTA Gene score for detection and risk stratification of ascending aortic dilation. *European Heart Journal*. 2024.

# PRS for CAD is not well explained by observable risk factors

PRS must exert its influence through the physical world… but not necessarily through traditional risk factors.

Opportunity to figure out what these factors are so that we can observe them.

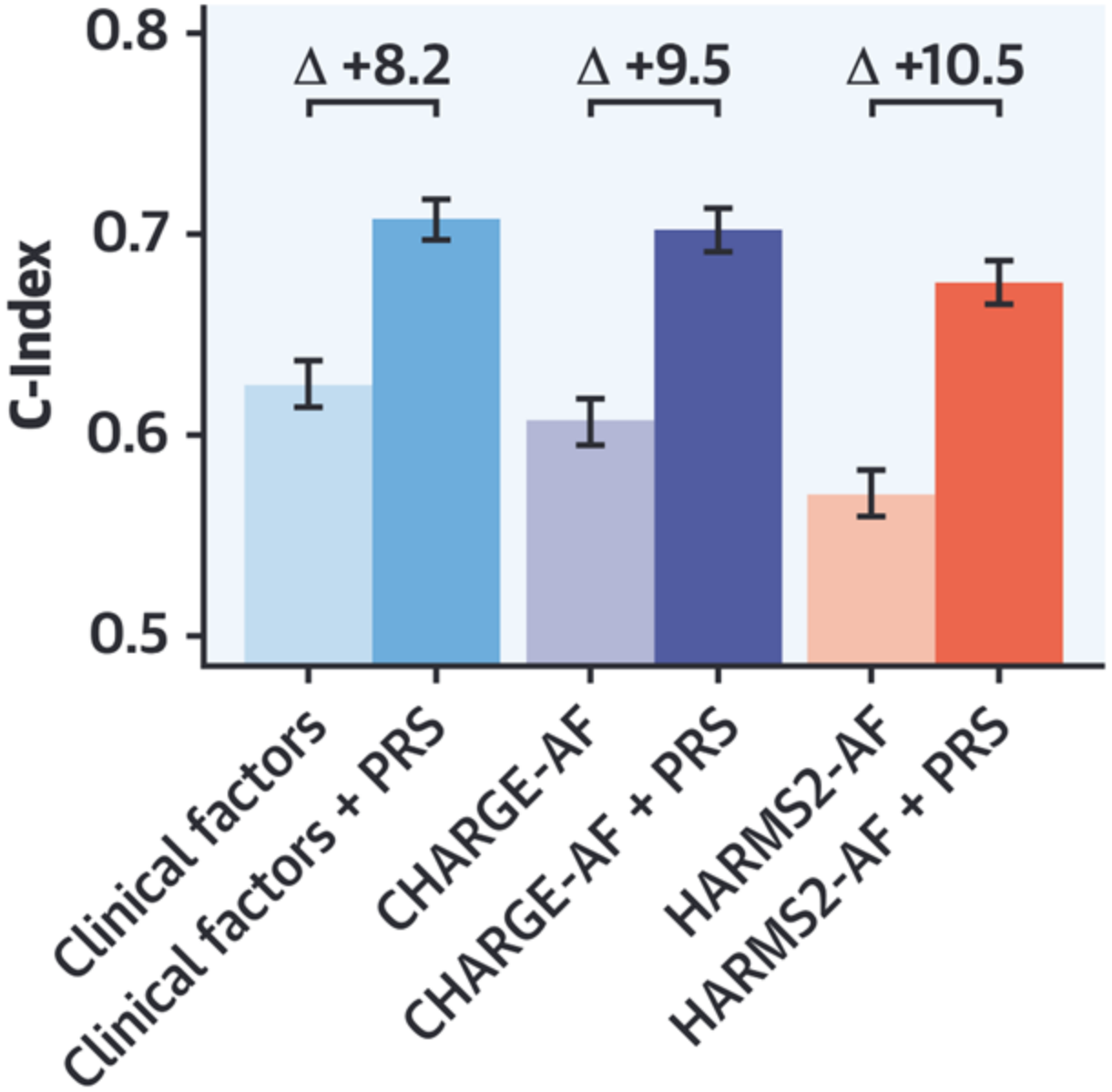
| | Remainder of population | Top 8% of $GPS_{CAD}$ distribution | P-value |
|---|---|---|---|
| Number of individuals | 265,859 | 23,119 | |
| Coronary artery disease | 7,061 (2.7%) | 1,615 (7.0%) | < 0.001 |
| Age, years | 56.9 (8.0) | 56.7 (8.1) | < 0.001 |
| Male sex | 120,673 (45%) | 10,410 (45%) | 0.29 |
| Hypertension | 73,982 (28%) | 7,477 (32%) | < 0.001 |
| Type 2 diabetes | 5,240 (2.0%) | 613 (2.7%) | < 0.001 |
| Hypercholesterolemia | 35,042 (13%) | 4,559 (20%) | < 0.001 |
| Current smoking | 24,399 (9.2%) | 2,200 (9.5%) | 0.09 |
| Family history of heart disease | 94,117 (35%) | 10,101 (44%) | < 0.001 |
| Body mass index, $kg/m^2$ | 27.3 (4.7) | 27.6 (4.8) | < 0.001 |
| Lipid-lowering therapy | 43,923 (17%) | 5,589 (24%) | < 0.001 |

Khera AV, et al. Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations. Nature Genetics. 2018

# PRS for atrial fibrillation

Additive beyond clinical risk scores

ASPREE (older population: age ≥ 70)




Fransquet PD, *et al*. Genomic Risk Prediction of Incident Atrial Fibrillation in Older Individuals Without Prior Cardiovascular Disease. *JACC: Advances*. 2025

# PRS for atrial fibrillation

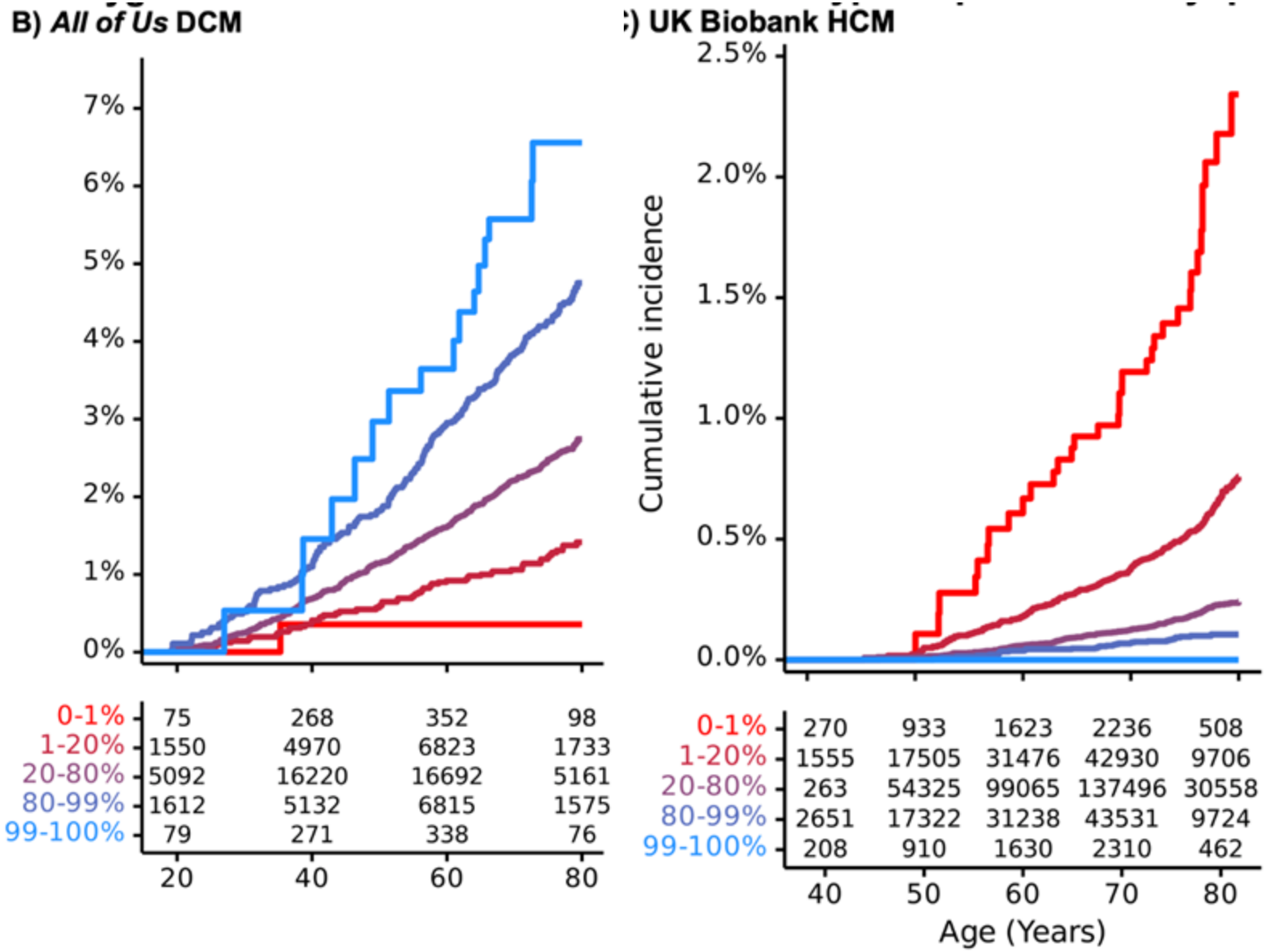
Greater ‘lift’ for women (but lower overall risk)

| | |
|---|---|
| CHARGE-AF in women | |
| CHARGE-AF alone | 0.606 (0.016) |
| CHARGE-AF with AF-PRS & PCs | 0.716 (0.014) |
| C-Index Δ | +11% |
| CHARGE-AF in men | |
| CHARGE-AF alone | 0.620 (0.018) |
| CHARGE-AF with AF-PRS & PCs | 0.686 (0.016) |
| C-index Δ | +6.6% |

Fransquet PD, *et al*. Genomic Risk Prediction of Incident Atrial Fibrillation in Older Individuals Without Prior Cardiovascular Disease. *JACC: Advances*. 2025

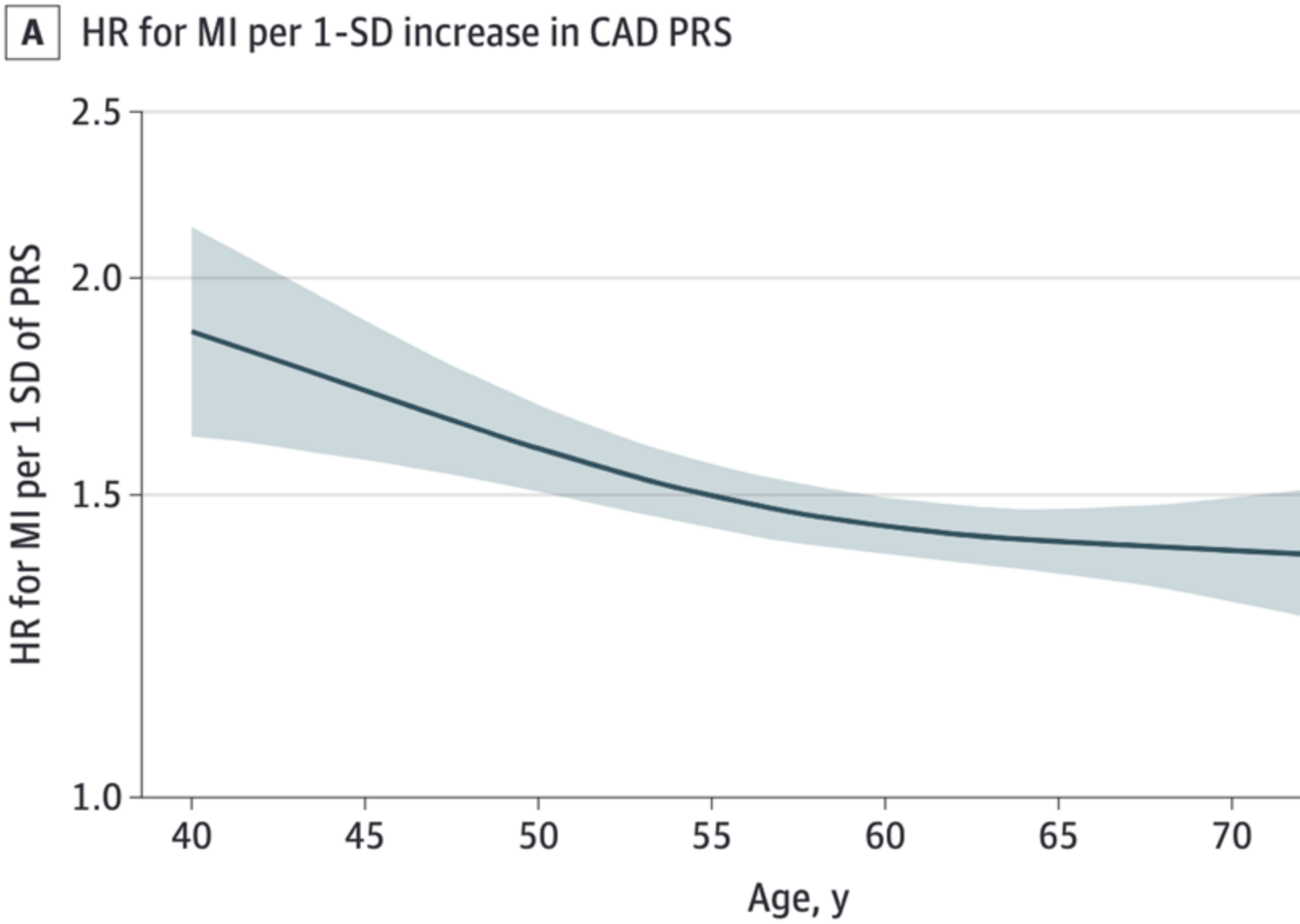
# PRS for cardiomyopathies

Pirruccello JP. Whole-genome sequencing analysis of left ventricular structure and sphericity in 80,000 people. medRxiv 2025:2025.08.22.25334019



# Polygenic scores for myocardial infarction

More informative for the young.

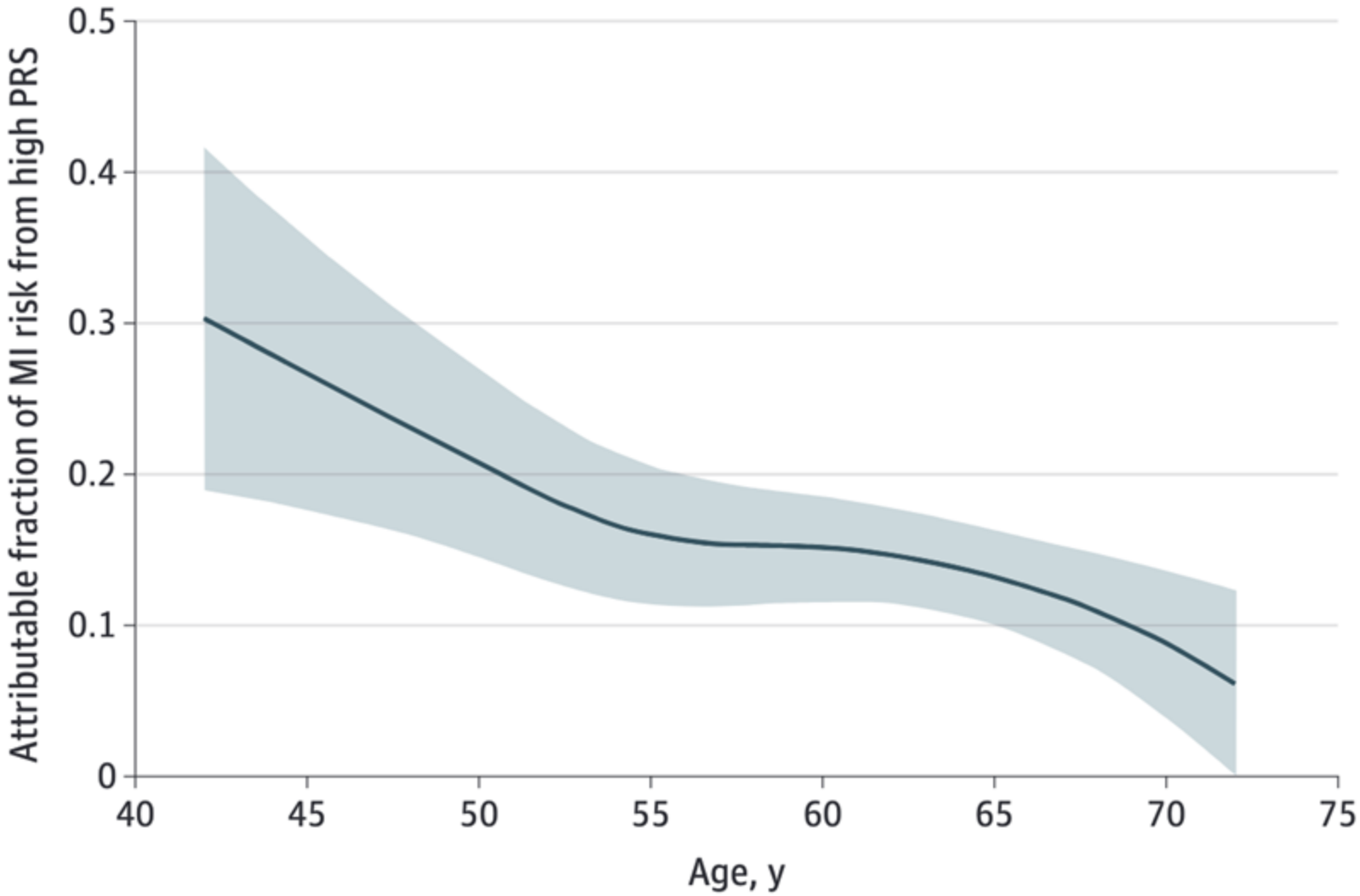



Marston NA, *et al*. **Predictive Utility of a Coronary Artery Disease Polygenic Risk Score in Primary Prevention**. *JAMA Cardiology*. 2023.

# Polygenic scores for myocardial infarction

Little information loss when removed from model predicting MI in older adults.

Lots of information loss when removed from model for younger adults.

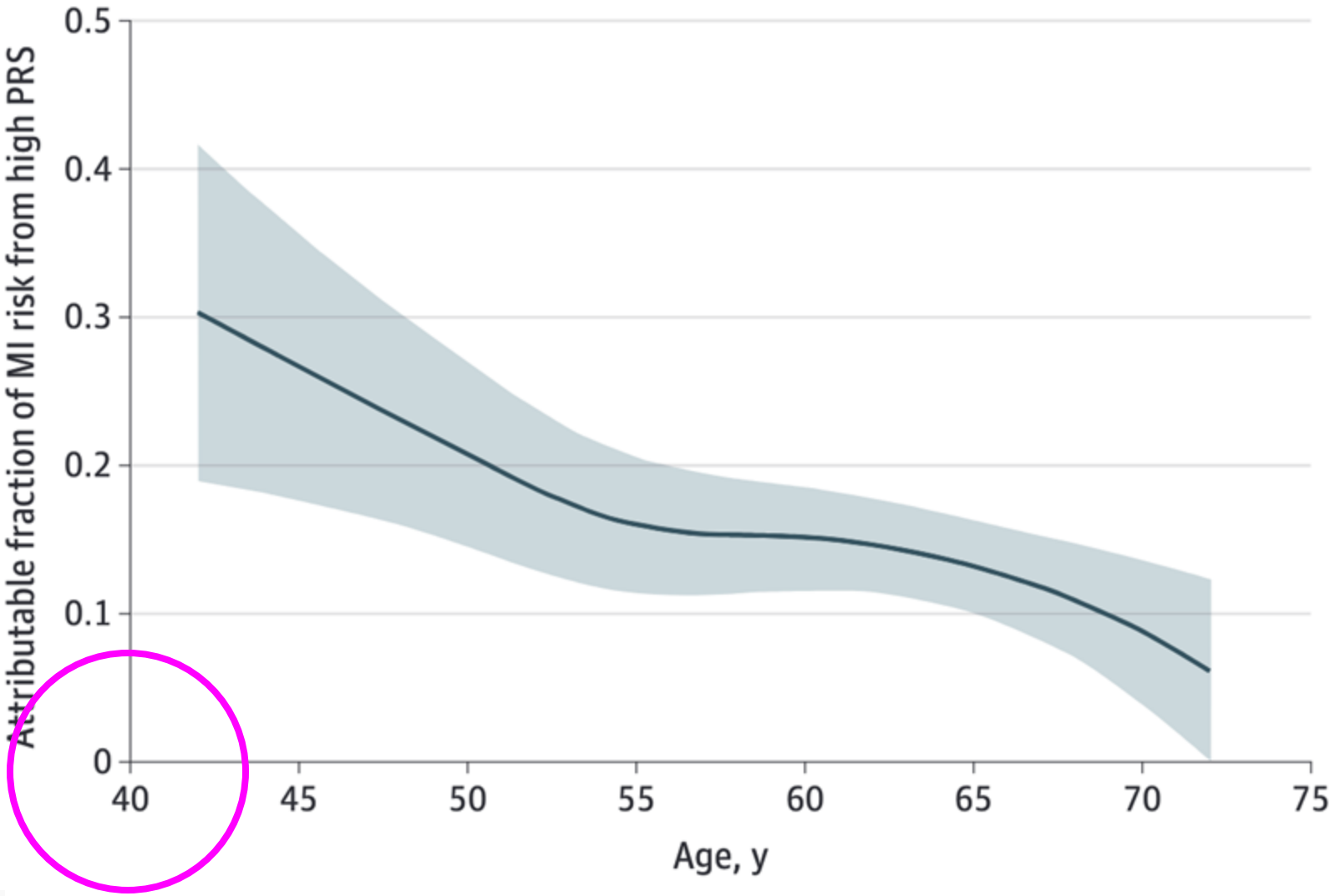



Marston NA, *et al*. **Predictive Utility of a Coronary Artery Disease Polygenic Risk Score in Primary Prevention**. *JAMA Cardiology*. 2023.

# Polygenic scores for myocardial infarction

Little information loss when removed from model predicting MI in older adults.

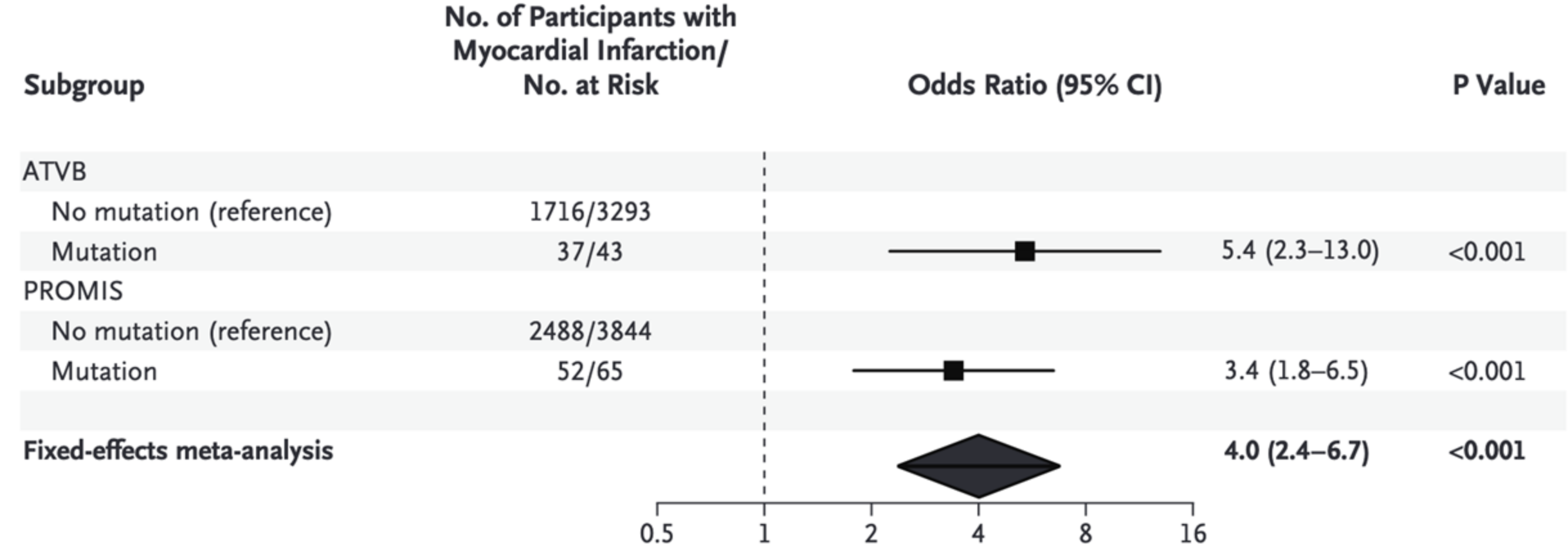
Lots of information loss when removed from model for younger adults.




Marston NA, *et al*. **Predictive Utility of a Coronary Artery Disease Polygenic Risk Score in Primary Prevention**. *JAMA Cardiology*. 2023.

# Somatic variation and MI under age 50

**B** **CHIP and Early-Onset Myocardial Infarction**



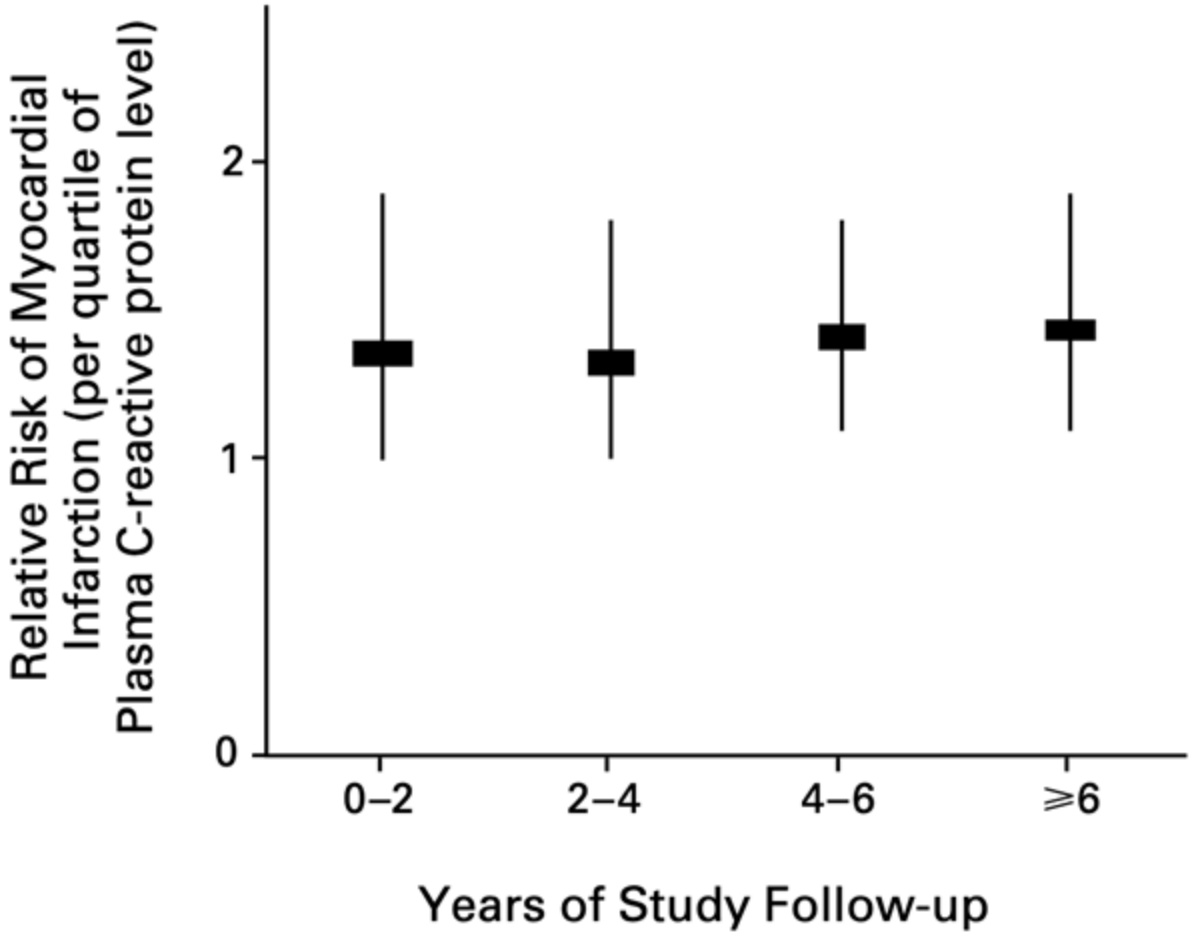

Jaiswal S, et al. Clonal Hematopoiesis and Risk of Atherosclerotic Cardiovascular Disease.

# I’d like to know:

- Will we be able to develop tools to infer appropriate allele weights for ancestry-specific alleles, mitigating ancestry bias?
- If we do achieve heritability-saturating polygenic scores, what does that buy us in terms of prediction for CV disease?
- Why does the (age•PRS) interaction not hold for non-ASCVD?

# Proteomics

# CRP: the original proteomic marker

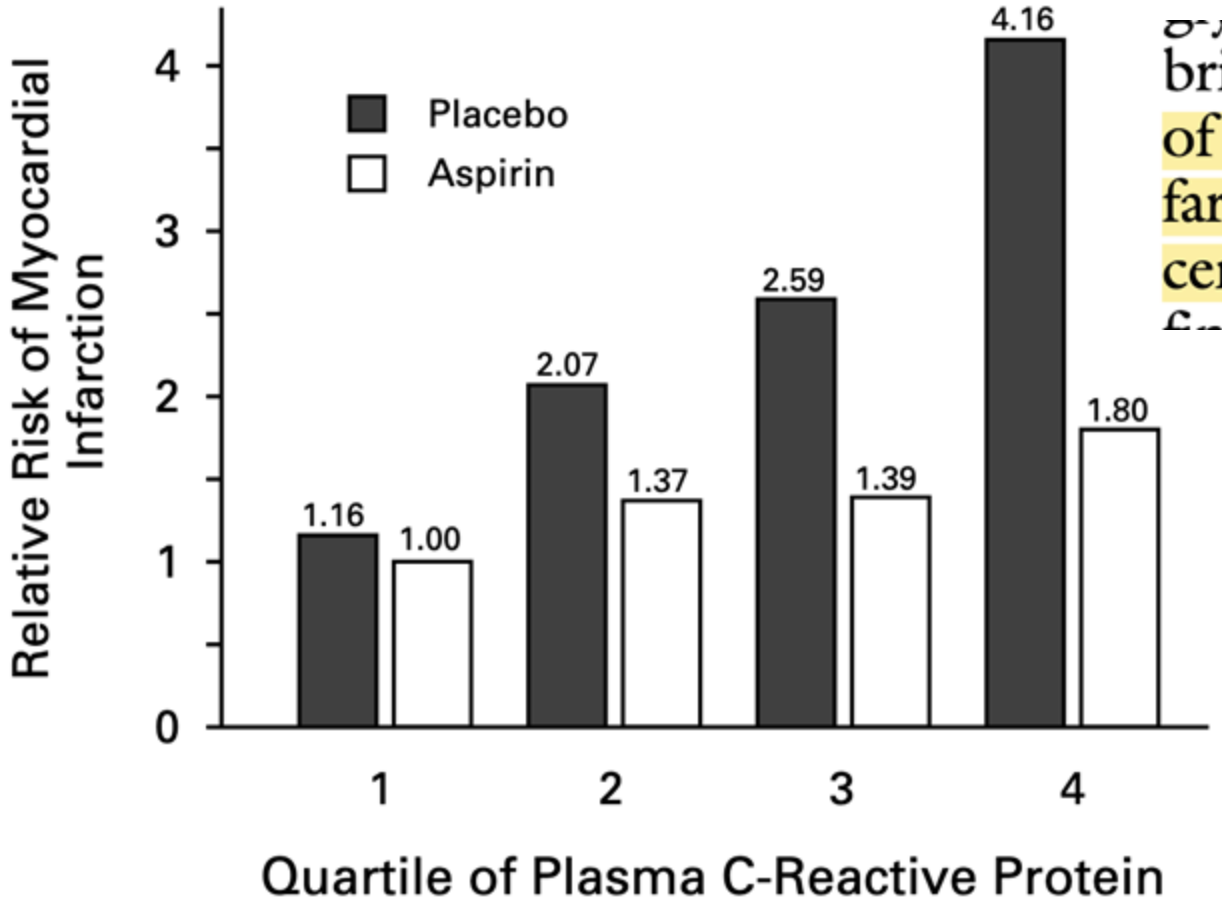



## DISCUSSION

These prospective data indicate that the base-line plasma concentration of C-reactive protein in apparently healthy men can predict the risk of first myocardial infarction and ischemic stroke. In addition, the risk of arterial thrombosis associated with the level of C-reactive protein was stable over long periods and was not modified by other factors, including smoking status, body-mass index, blood pressure, or the plasma concentration of total or HDL cholesterol, tri-

Ridker PM, *et al*. Inflammation, Aspirin, and the Risk of Cardiovascular Disease in Apparently Healthy Men. *New England Journal of Medicine*. 1997.

# CRP: the original proteomic marker

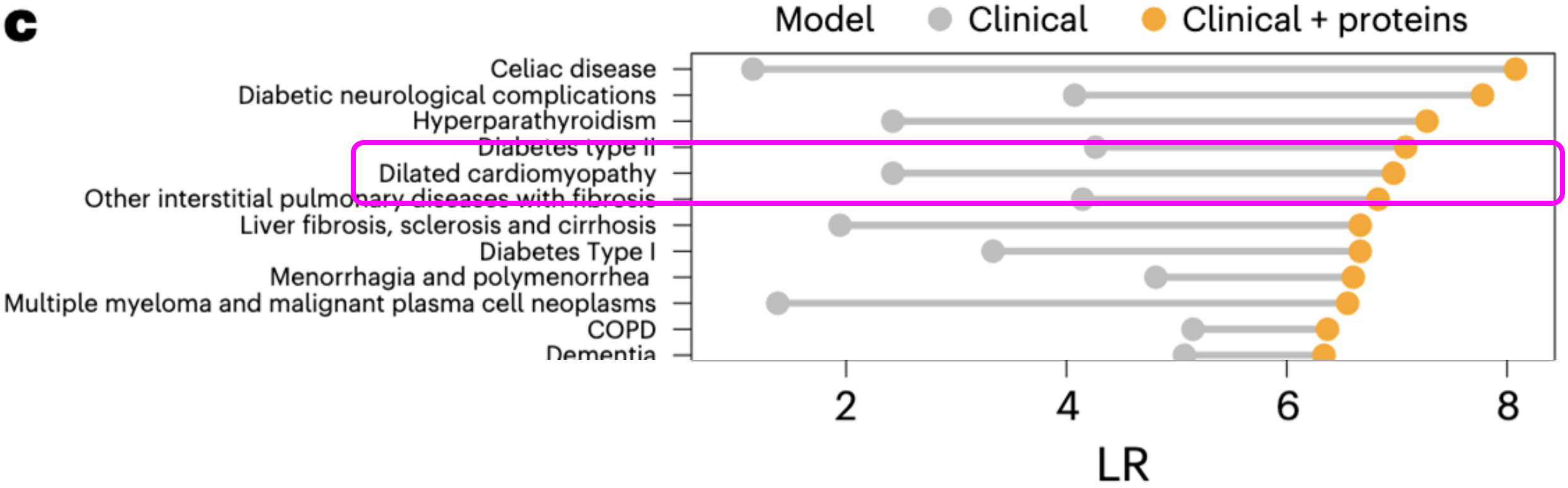



brinogen, or homocysteine. In contrast, the benefits of aspirin in reducing the risk of a first myocardial infarction diminished significantly with decreasing concentrations of C-reactive protein — an intriguing

(Interaction P < 0.05)

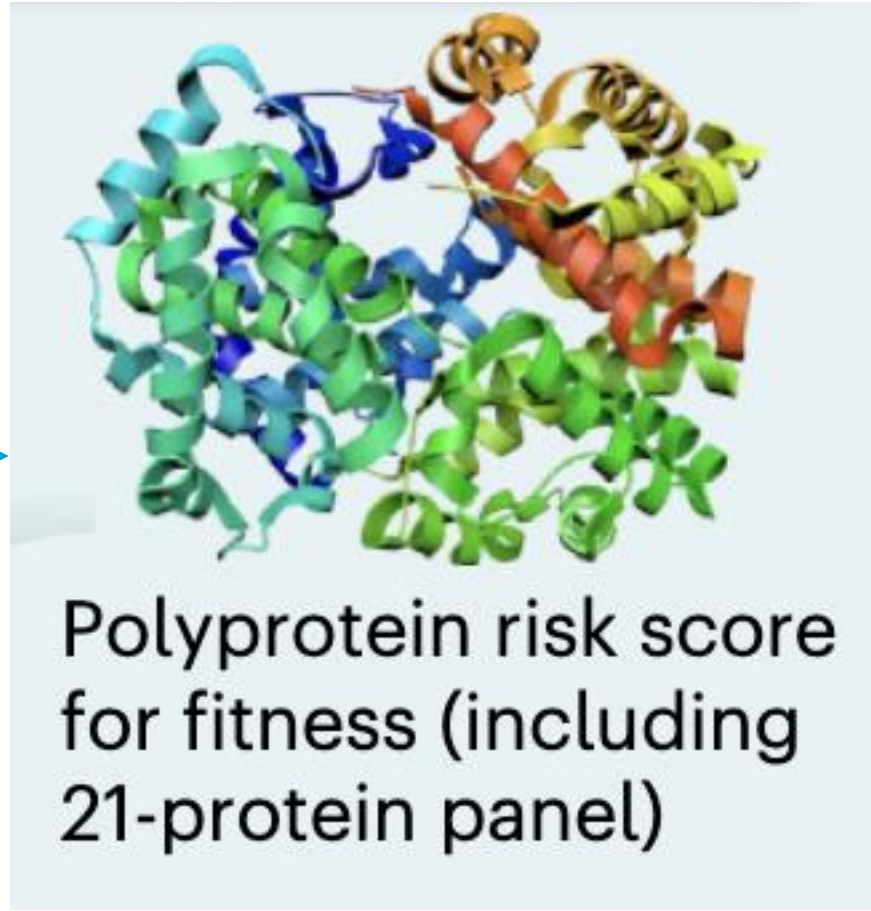
Ridker PM, *et al*. Inflammation, Aspirin, and the Risk of Cardiovascular Disease in Apparently Healthy Men. *New England Journal of Medicine*. 1997.

# Blood proteomics can improve risk stratification for cardiomyopathy

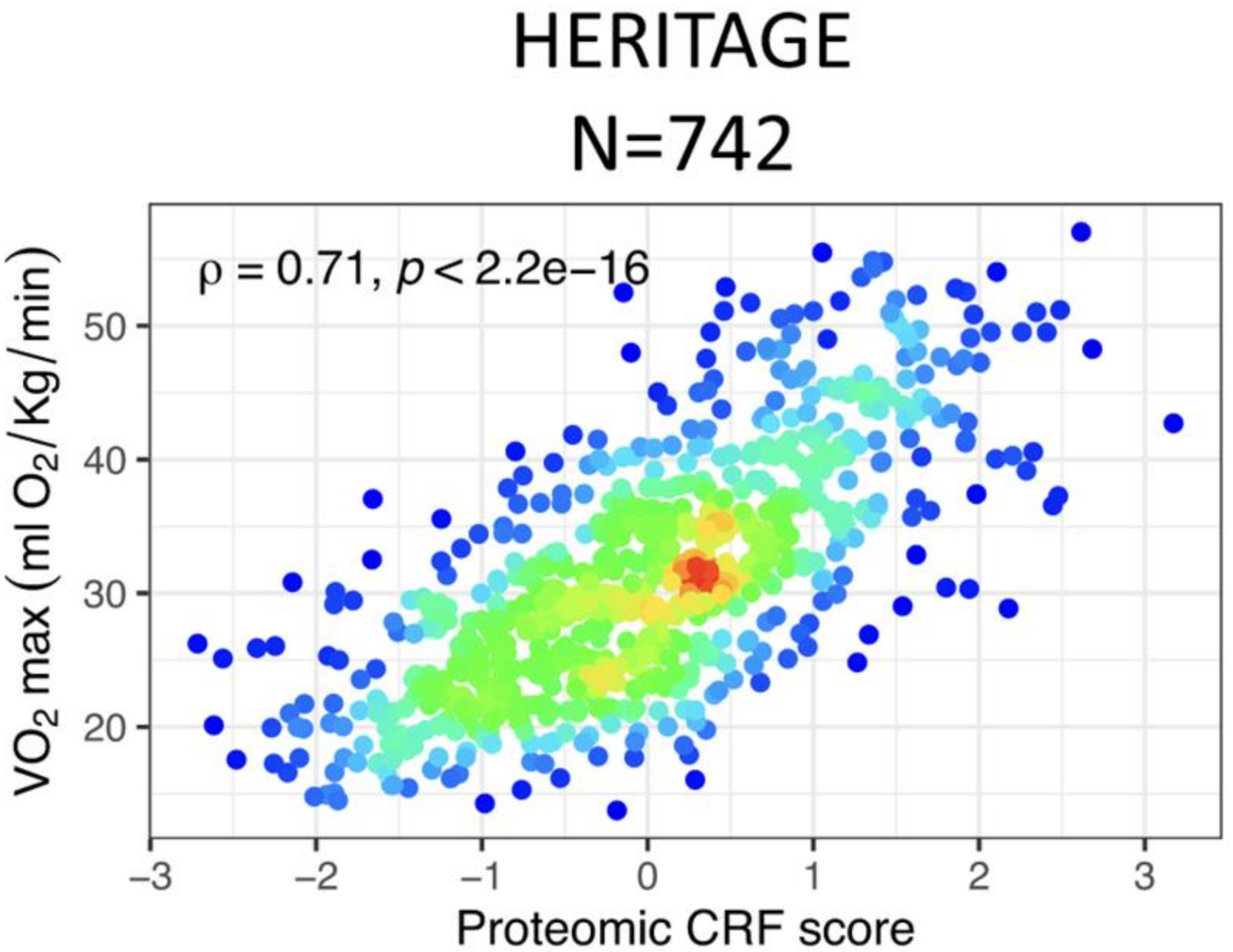



Carrasco-Zanini J, et al. Proteomic signatures improve risk prediction for common and rare diseases. *Nat Med*. 2024;30:2489–2498.

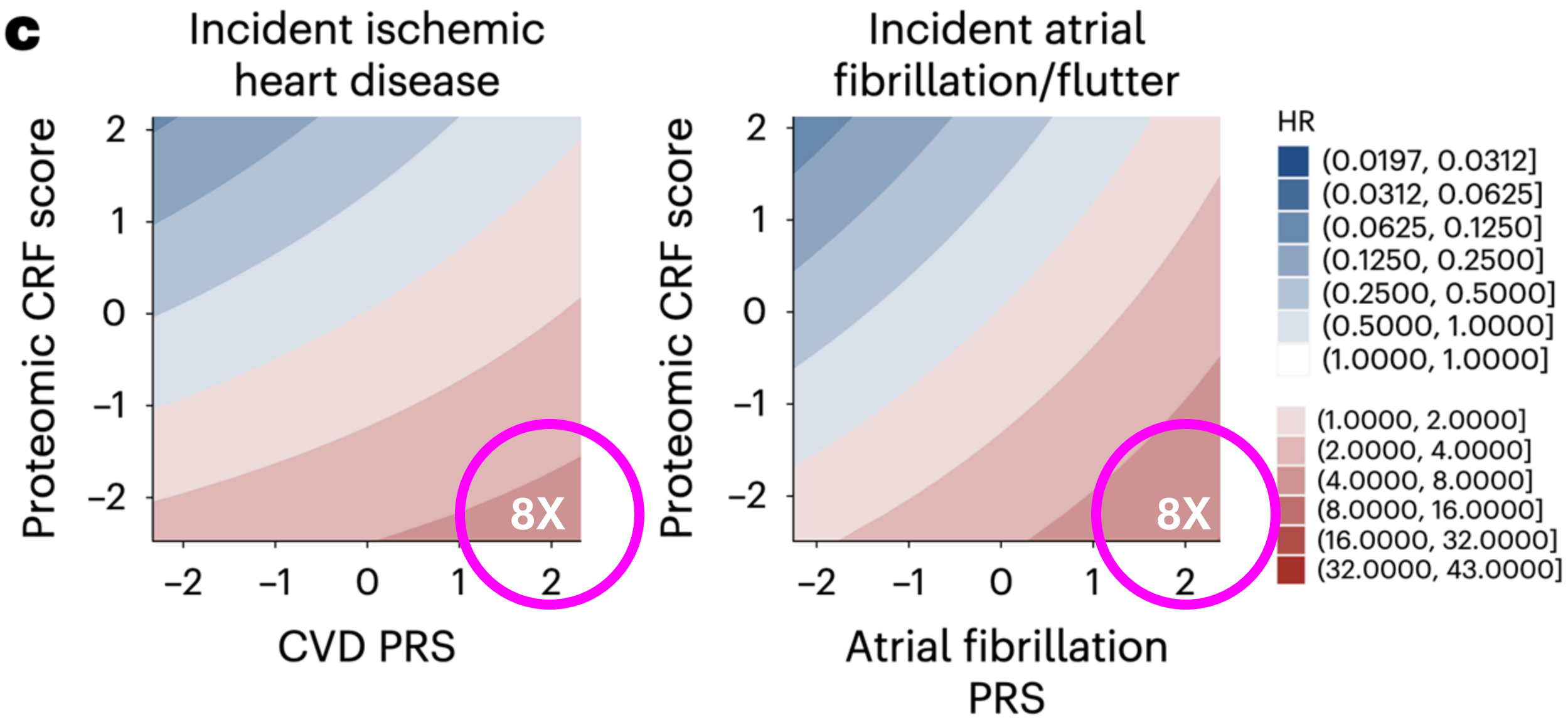
# Olink proteomic score for cardiorespiratory fitness




Perry AS, et al. Proteomic analysis of cardiorespiratory fitness for prediction of mortality and multisystem disease risks. *Nat Med*. 2024.

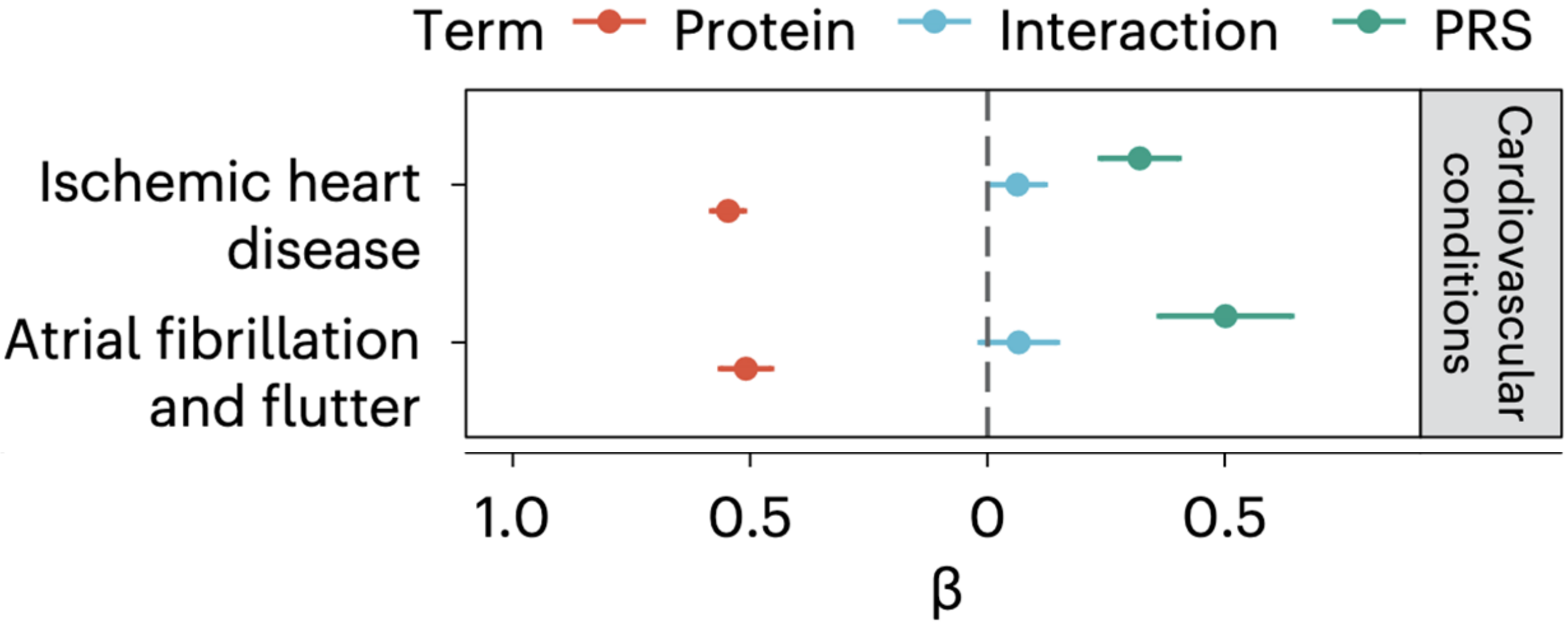



Perry AS, et al. Proteomic analysis of cardiorespiratory fitness for prediction of mortality and multisystem disease risks. *Nat Med*. 2024.




Perry AS, et al. Proteomic analysis of cardiorespiratory fitness for prediction of mortality and multisystem disease risks. *Nat Med*. 2024.

# PRS and proteomics seem to be additive, not multiplicative

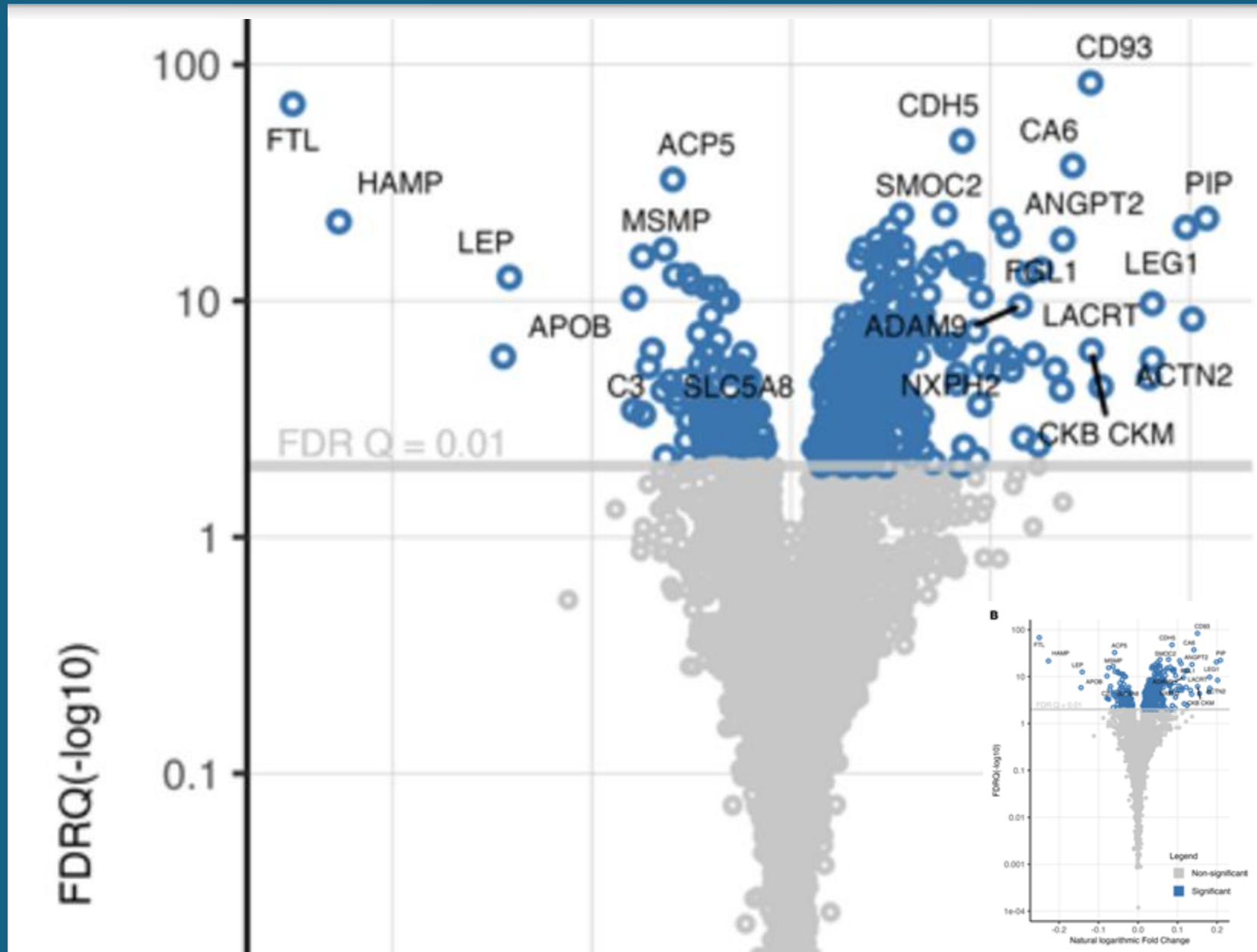



Perry AS, et al. Proteomic analysis of cardiorespiratory fitness for prediction of mortality and multisystem disease risks. *Nat Med*. 2024.

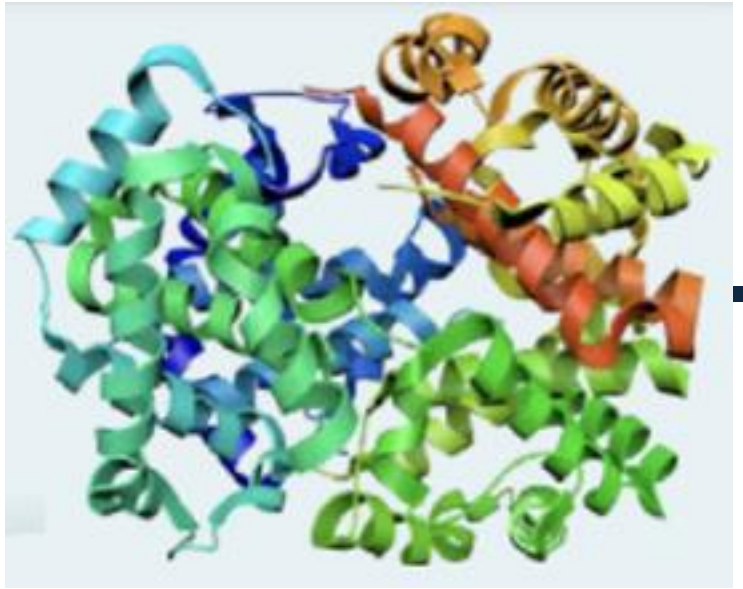
# Proteomic changes after 20 weeks of endurance exercise training

245 out of 4,914 proteins significantly changed after 20 weeks of exercise training (N=443)

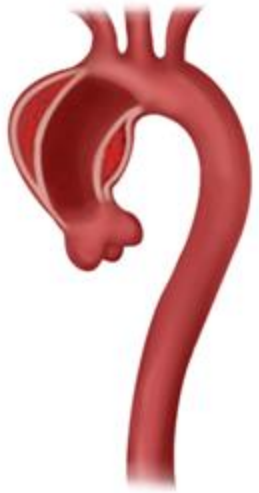
Robbins JM, et al. Plasma proteomic changes in response to exercise training are associated with cardiorespiratory fitness adaptations. JCI Insight 2023.



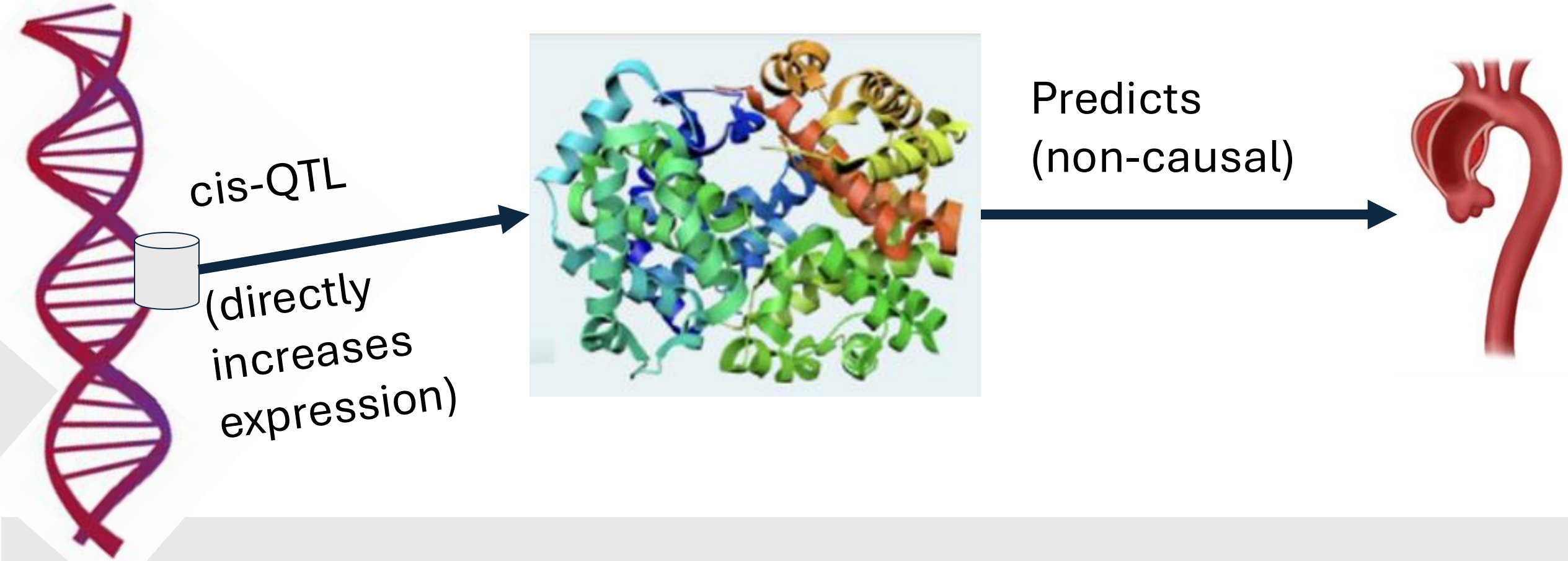
# We need to understand the genetic architecture of proteomics to maximally benefit



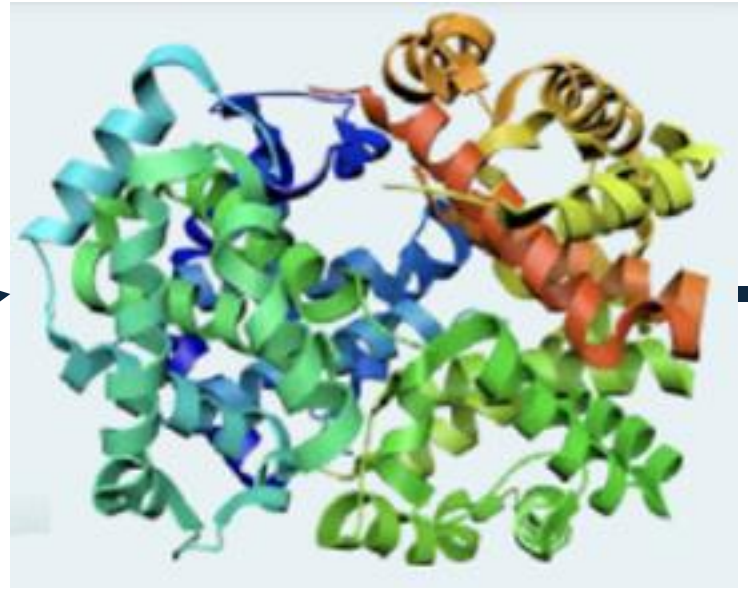
Predicts
(non-causal)



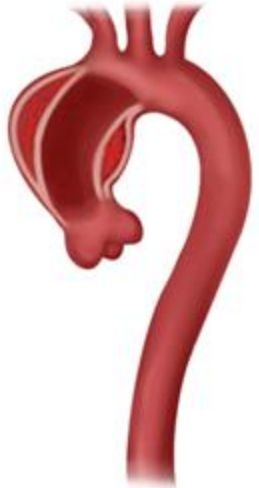
# We need to understand the genetic architecture of proteomics to maximally benefit



# We need to understand the genetic architecture of proteomics to maximally benefit

Predicts
(non-causal)



Here, cis-QTLs are simply noise - model them to improve your prediction

# I’d like to know:

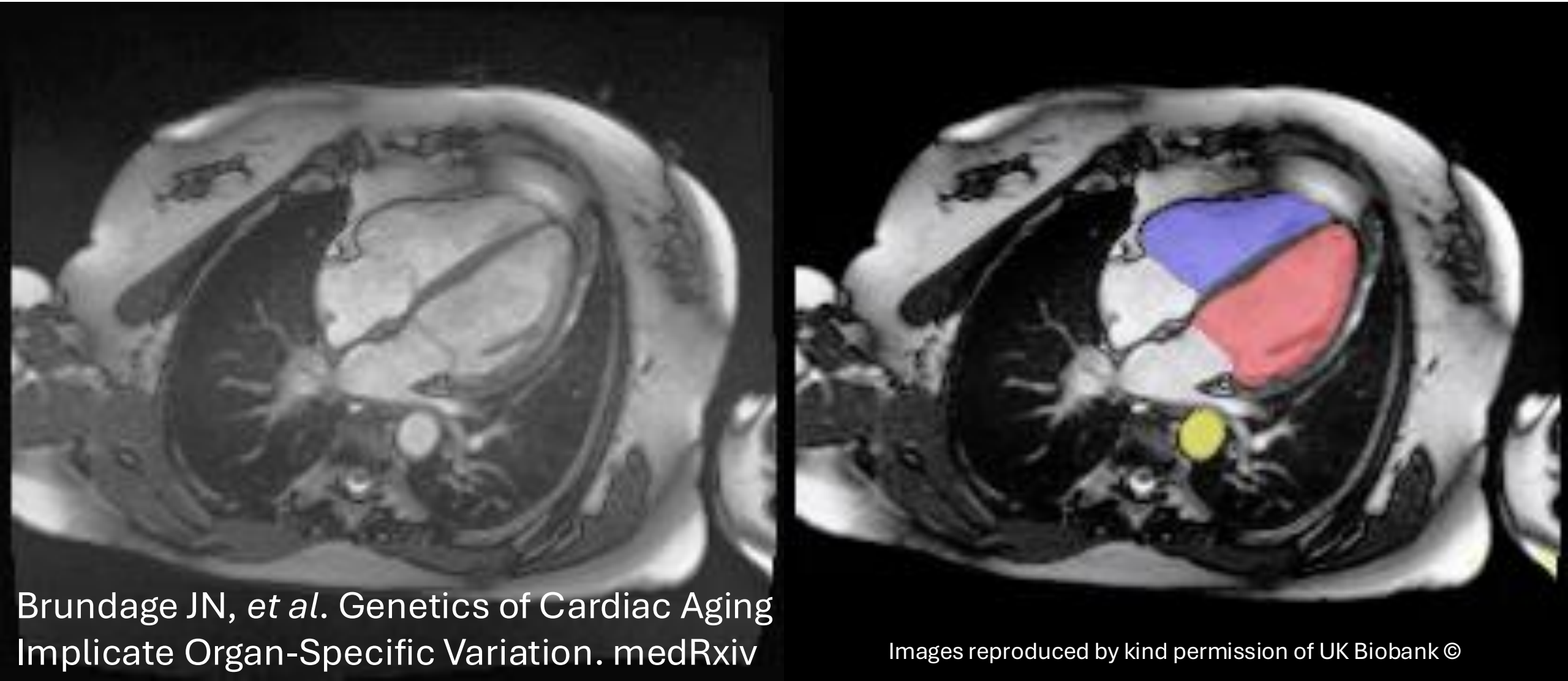
- Are there protein sets that can help us distinguish near-term risk? Ultra-long-term risk?
- What is the value of repeat measures?
- Could some be used as surrogate endpoints?

# Deep learning: image models

# Masking the 4-chamber MRI

Raw MRI

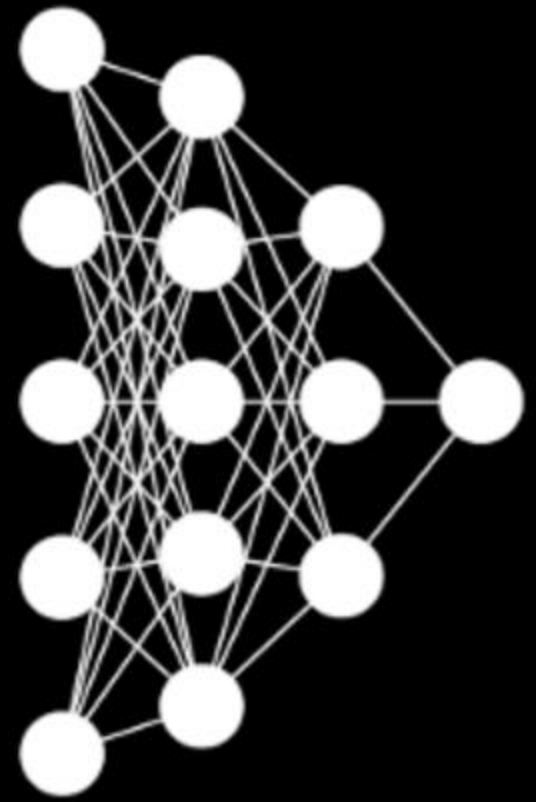
Segmented MRI



Brundage JN, *et al*. Genetics of Cardiac Aging Implicate Organ-Specific Variation. medRxiv

Images reproduced by kind permission of UK Biobank ©

# Estimating age from cardiac MRI CINEs: **error = signal**

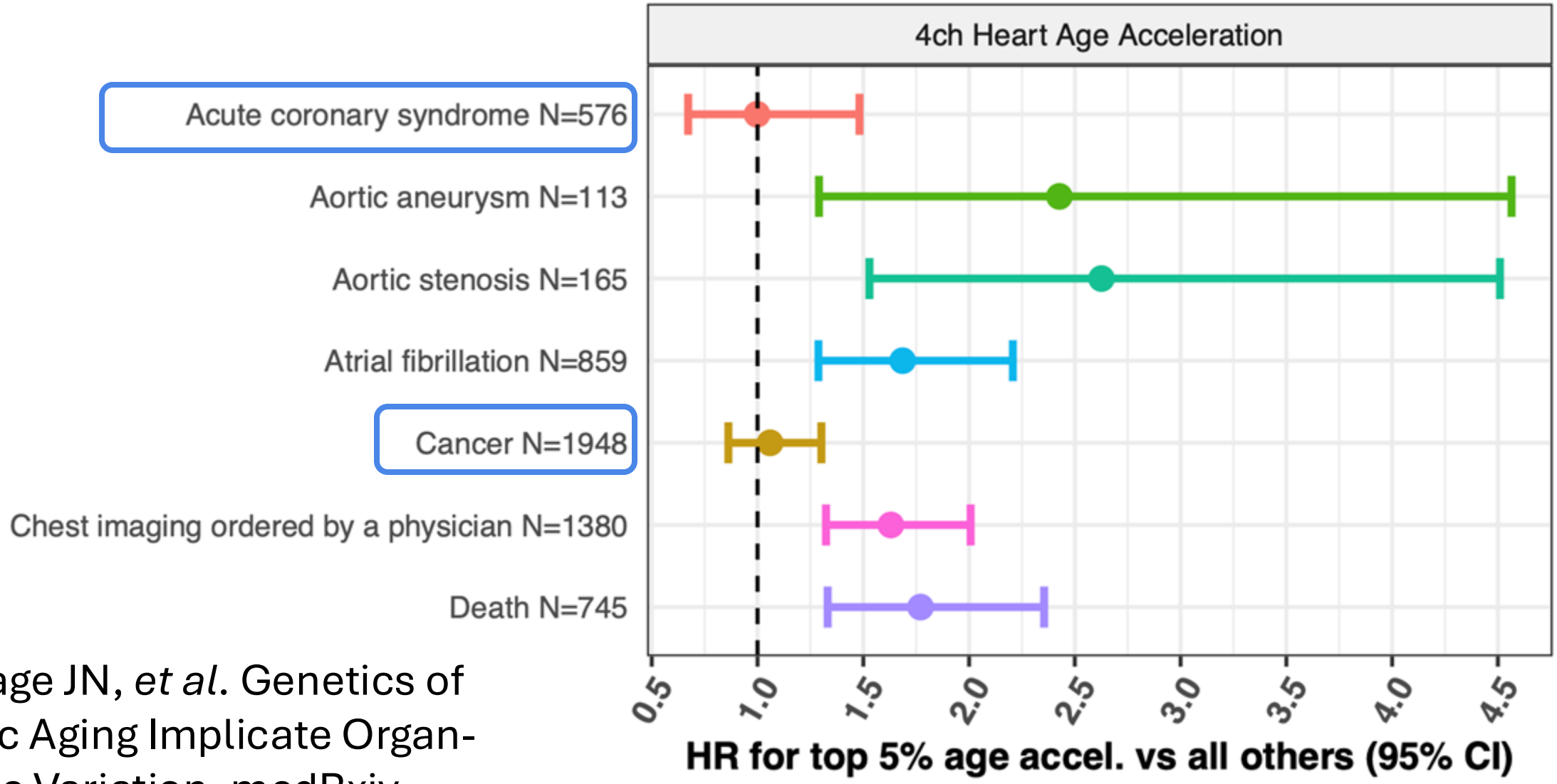


+7 years

Brundage JN, *et al*. Genetics of Cardiac Aging Implicate Organ-Specific Variation. medRxiv

Images reproduced by kind permission of UK Biobank ©

# Cardiac age acceleration is not associated with all age-related diseases




Brundage JN, *et al*. Genetics of Cardiac Aging Implicate Organ-Specific Variation. medRxiv

# Interpretability can be a lie

❌ Delete the left atrium, **run** the age model:

model loses ~all predictive power

➡️ Left atrium must be important!

Images reproduced by kind permission of UK Biobank ©

# Interpretability can be a lie

✅ Delete the left atrium, **train a new** age model:

model regains ~all predictive power

🤷

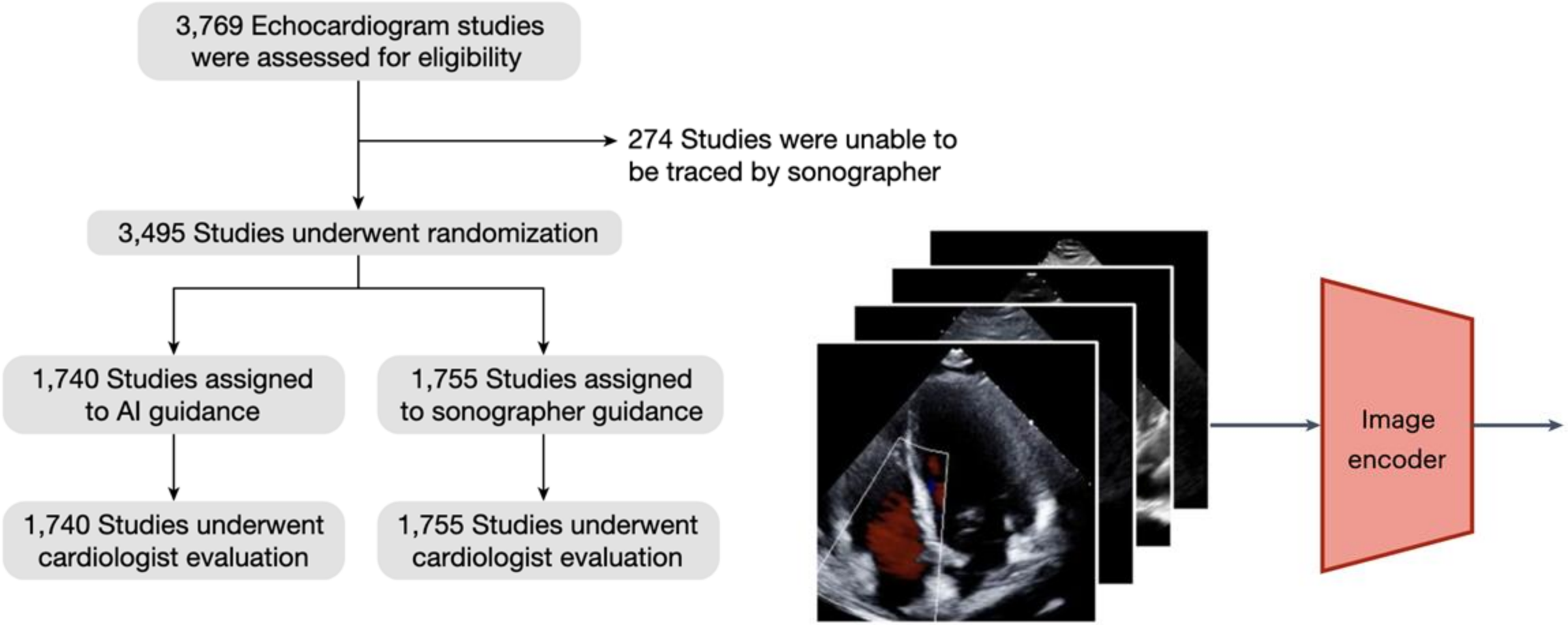
Images reproduced by kind permission of UK Biobank ©

# Opportunistic screening?

# Opportunistic screening?

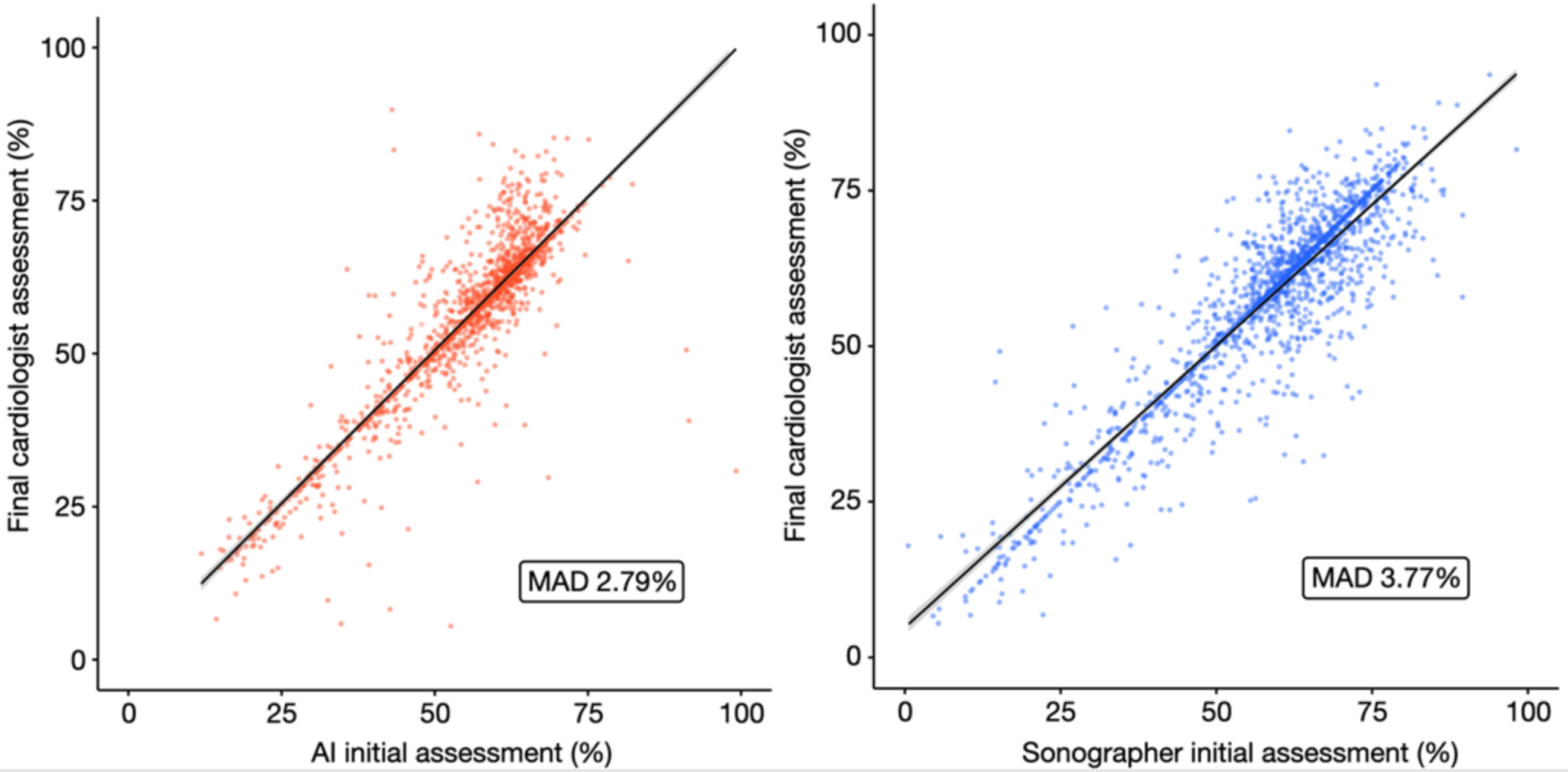
Implies that you had a reason for imaging in the first place…

# Broadening access to noninvasive imaging: automated LVEF measurement




He B, *et al*. Blinded, randomized trial of sonographer versus AI cardiac function

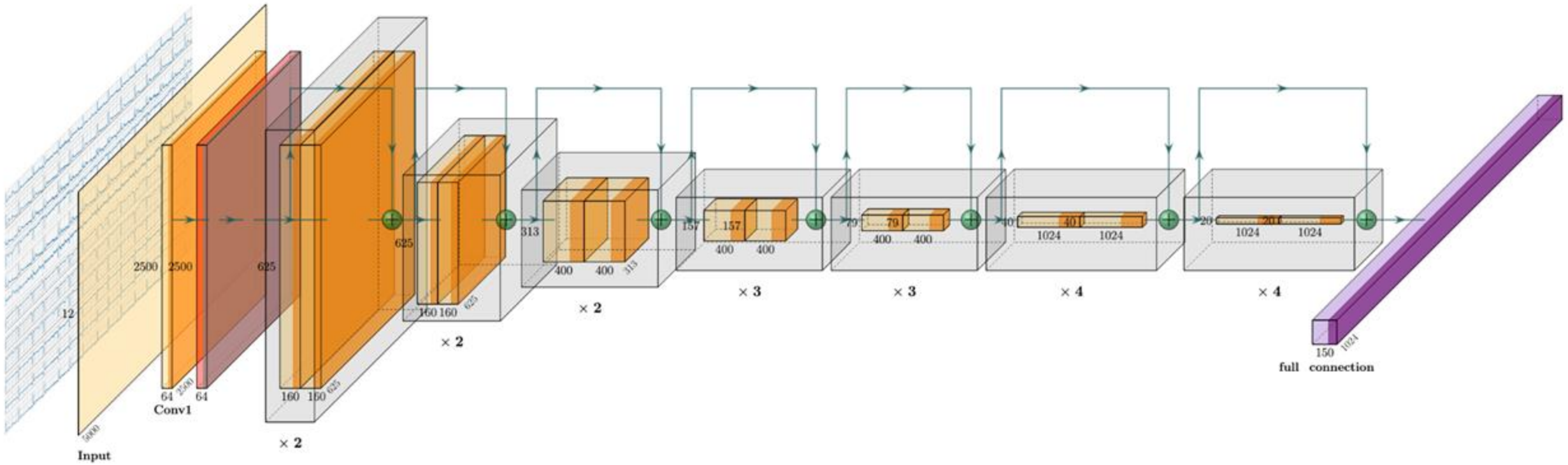
# Broadening access to noninvasive imaging: automated LVEF measurement




He B, *et al*. Blinded, randomized trial of sonographer versus AI cardiac function

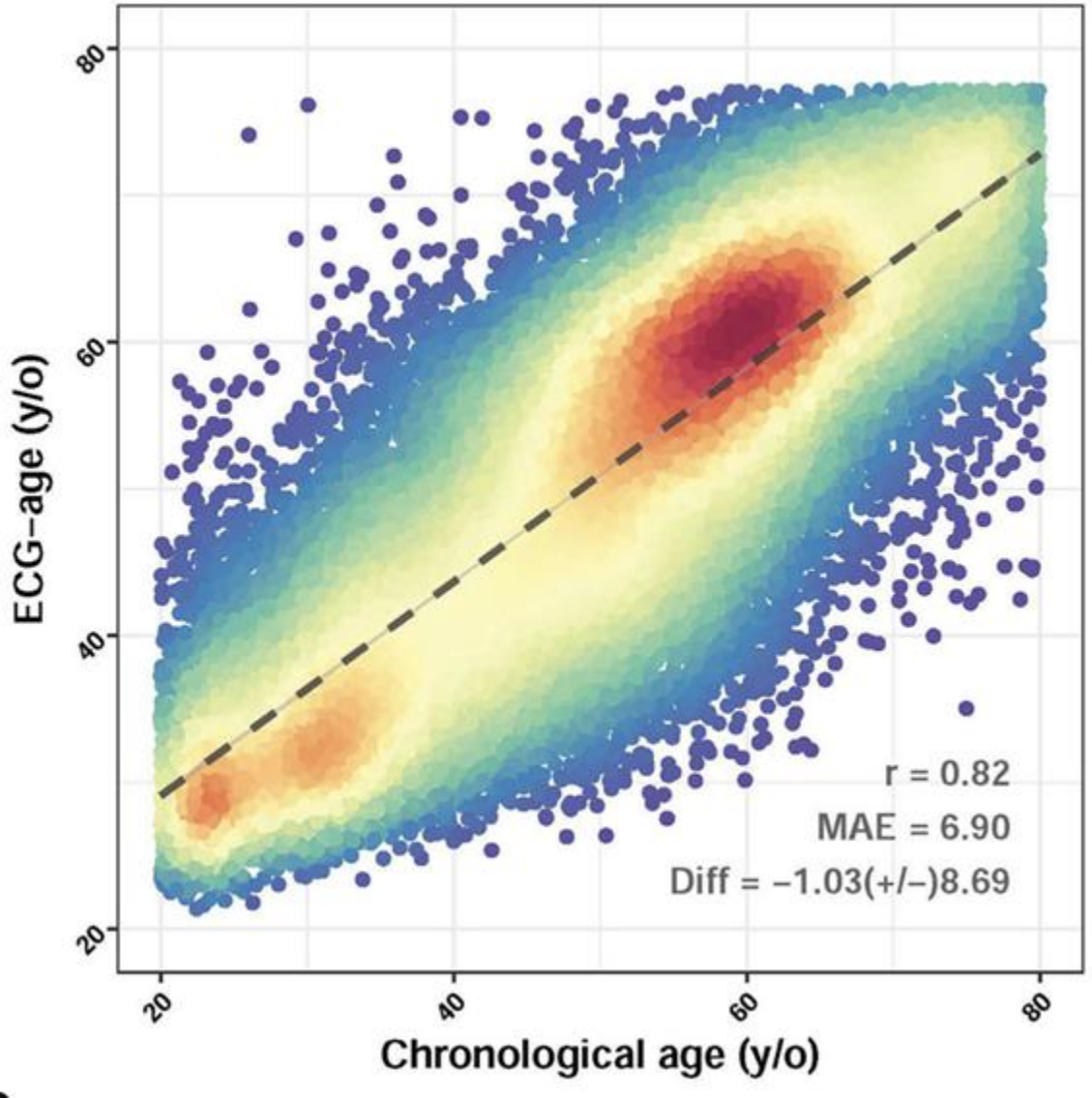
# Deep learning ECG models

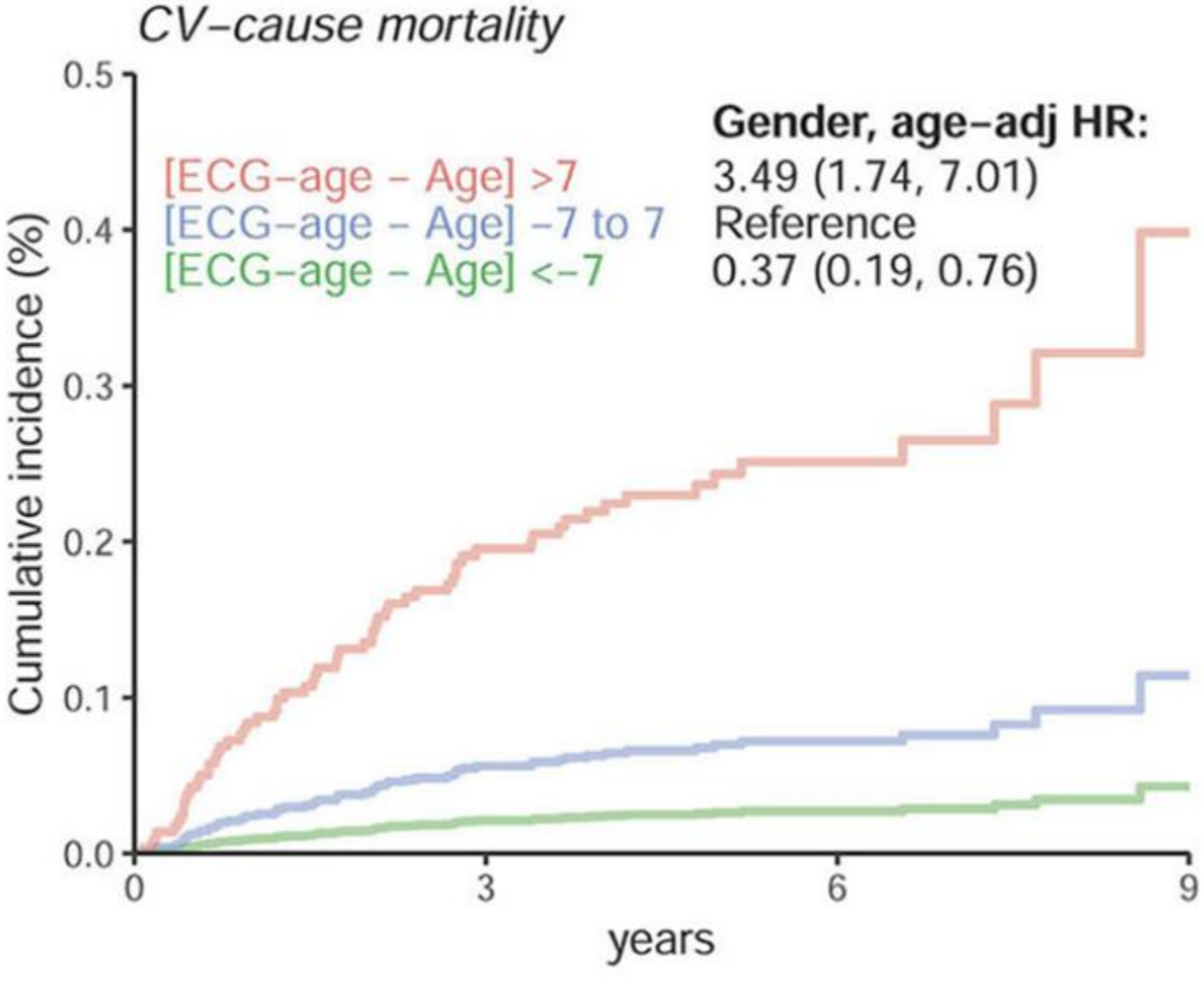
# Building an ECG representation




Li J, et al. An Electrocardiogram Foundation Model Built on over 10 Million Recordings with External Evaluation across Multiple Domains. arXiv. 2025.

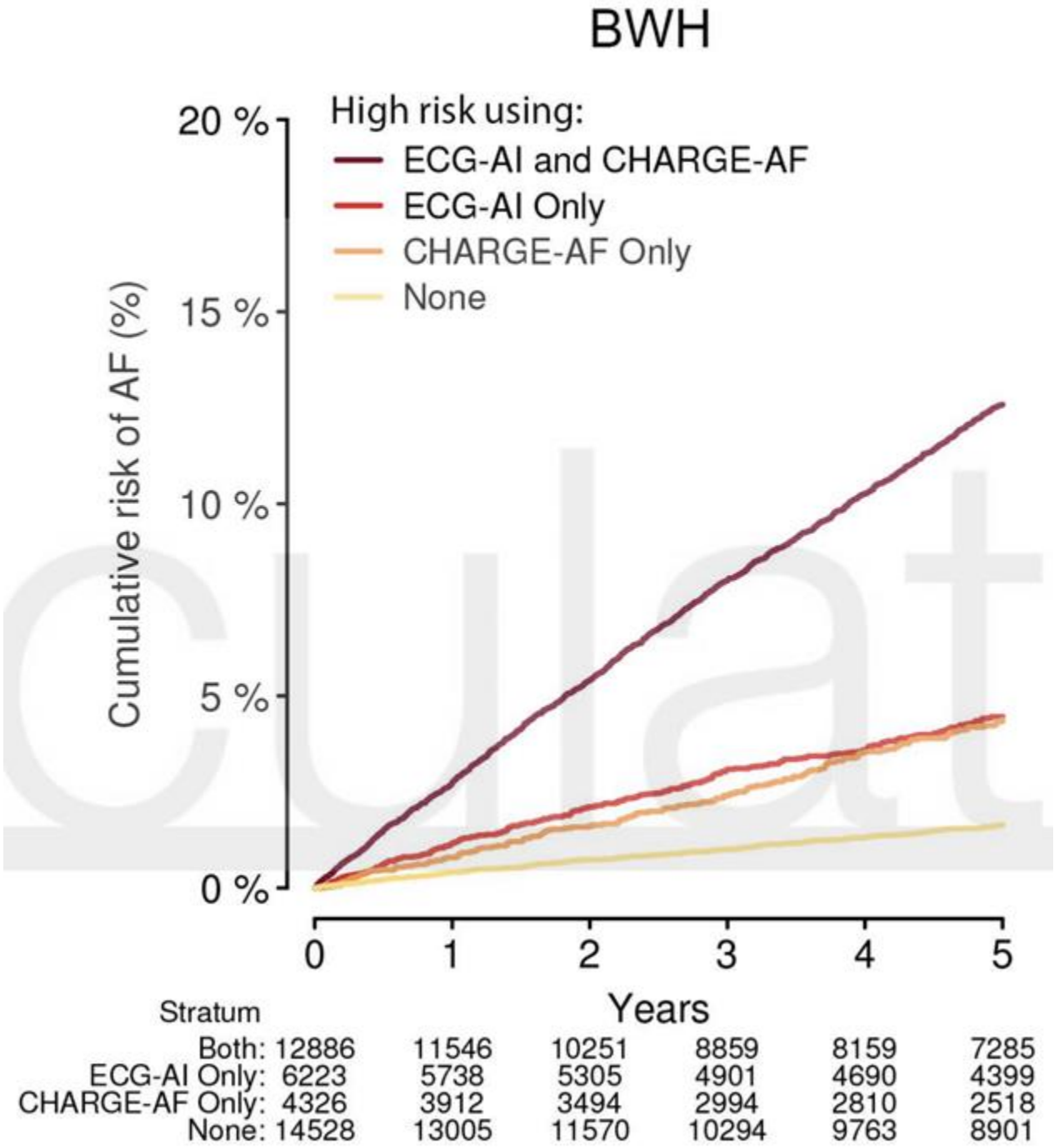
# Abstract representations: ECG-Age



# ECG models for atrial fibrillation

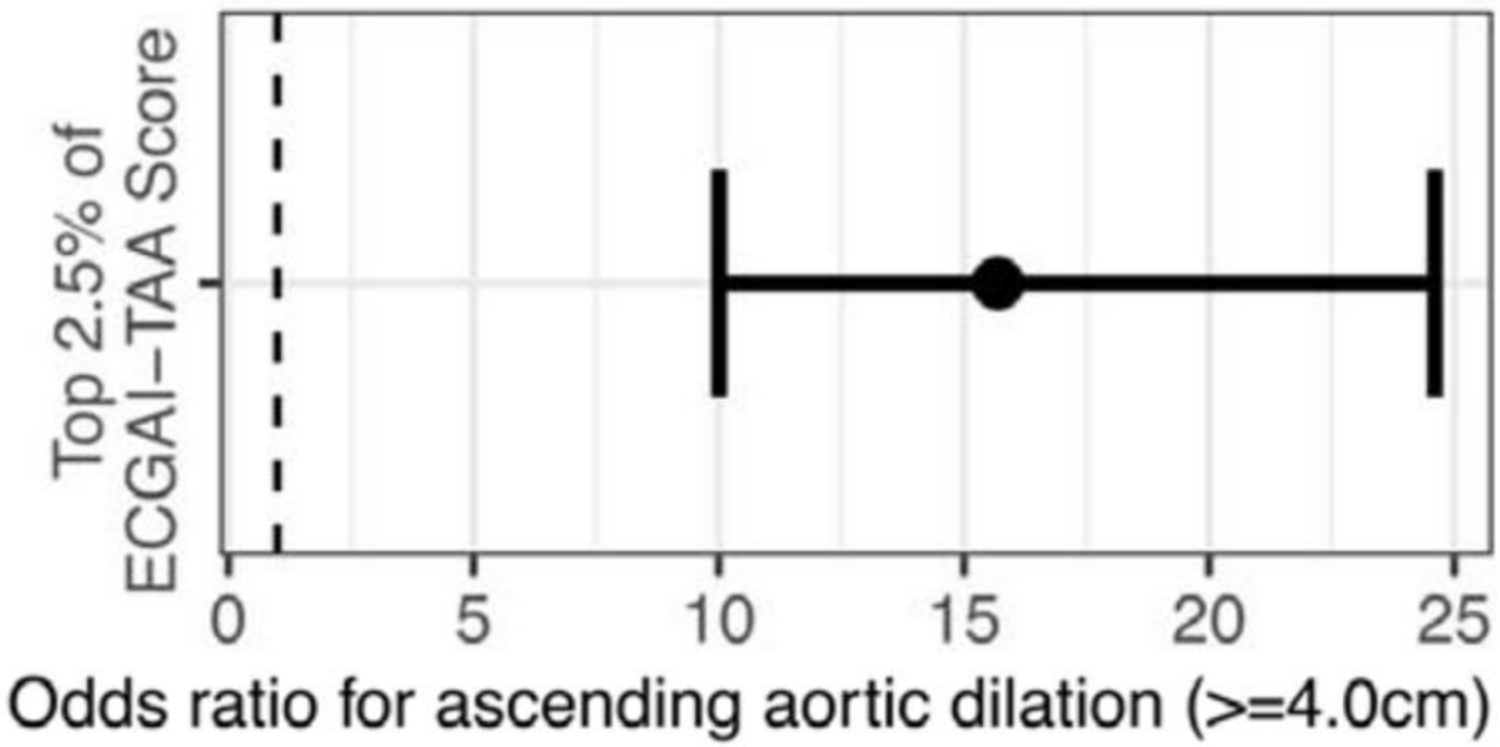
Non-redundant with clinical predictors, additive risk



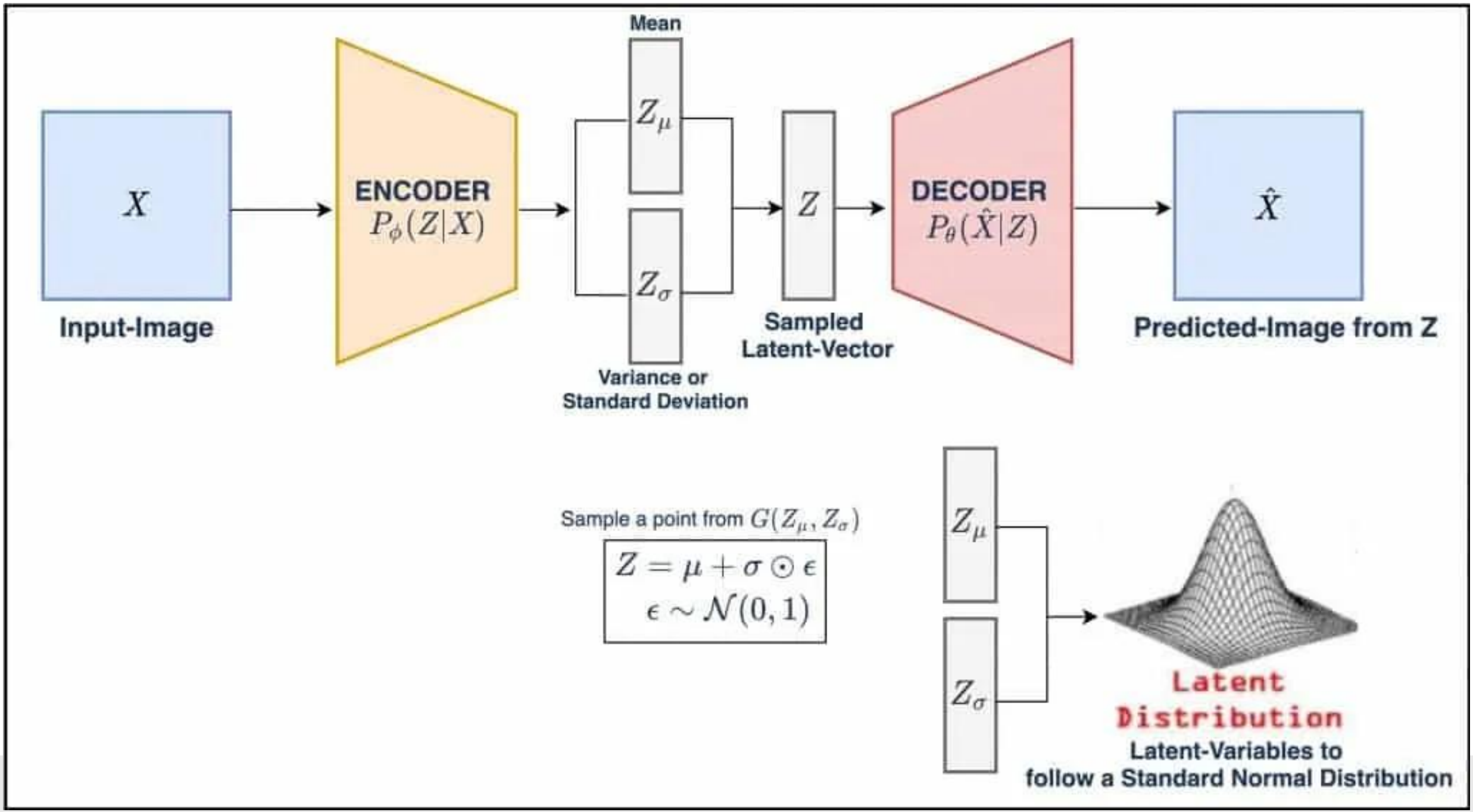

Khurshid S, *et al*. Electrocardiogram-based Deep Learning and Clinical Risk Factors to Predict Atrial Fibrillation. Circulation 2021

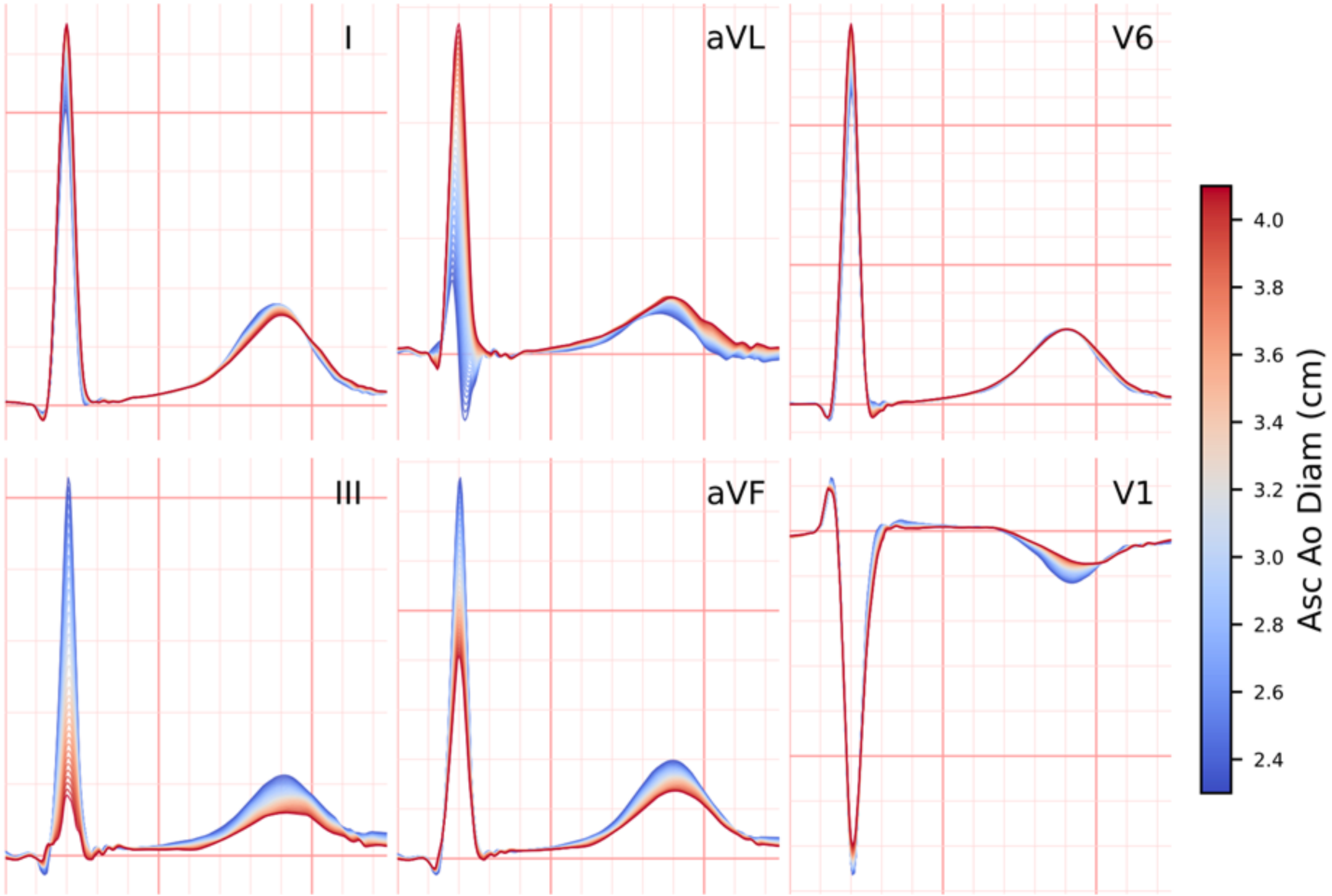
# ECG models for aortic aneurysm

Not just informative in the middle of the distribution, but help identify individuals with clinically relevant dilation




Demarais ZS, *et al*. Estimating ascending aortic diameter from the electrocardiogram. medRxiv


Mean
$X$
ENCODER
$P_\phi(Z|X)$
$Z_\mu$
$Z_\sigma$
$Z$
DECODER
$P_\theta(\hat{X}|Z)$
$\hat{X}$
Input-Image
Variance or
Standard Deviation
Sampled
Latent-Vector
Predicted-Image from Z
Sample a point from $G(Z_\mu, Z_\sigma)$
$Z = \mu + \sigma \odot \epsilon$
$\epsilon \sim \mathcal{N}(0,1)$
$Z_\mu$
$Z_\sigma$
Latent
Distribution
Latent-Variables to
follow a Standard Normal Distribution

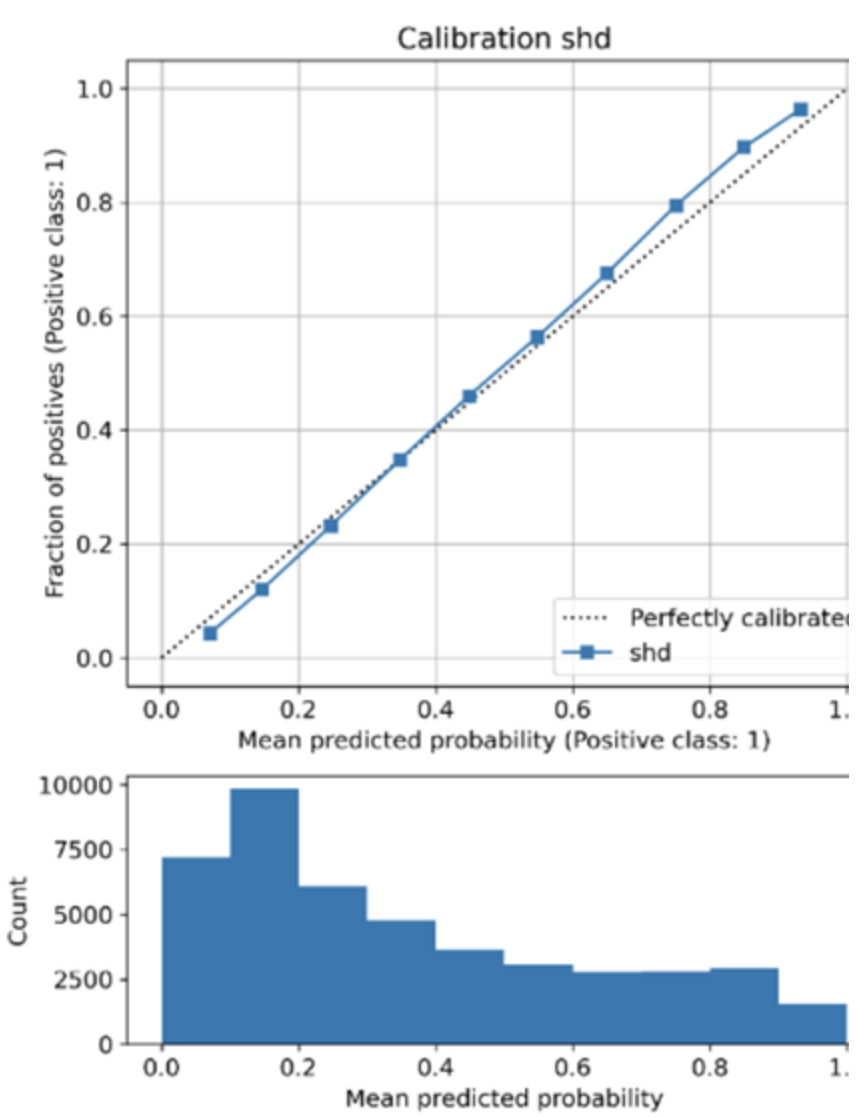


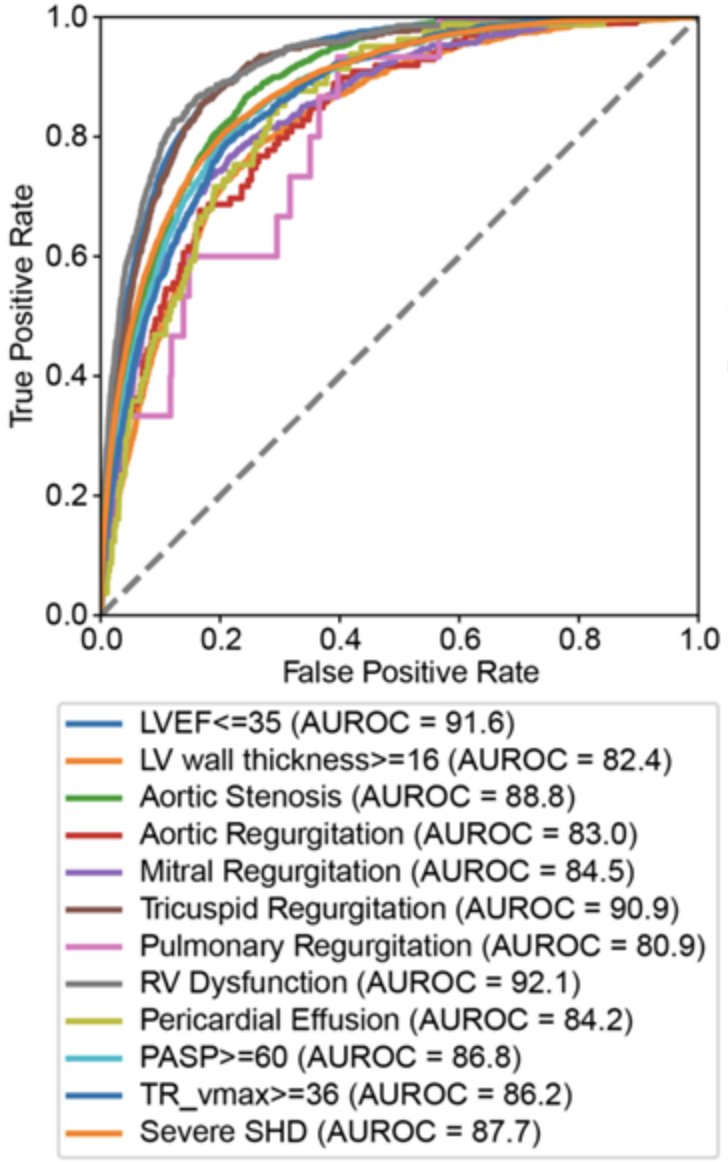

Demarais ZS, *et al*. Estimating ascending aortic diameter from the electrocardiogram. medRxiv 2025:2025.08.19.25333786.

Reproduced by kind permission of UK Biobank ©

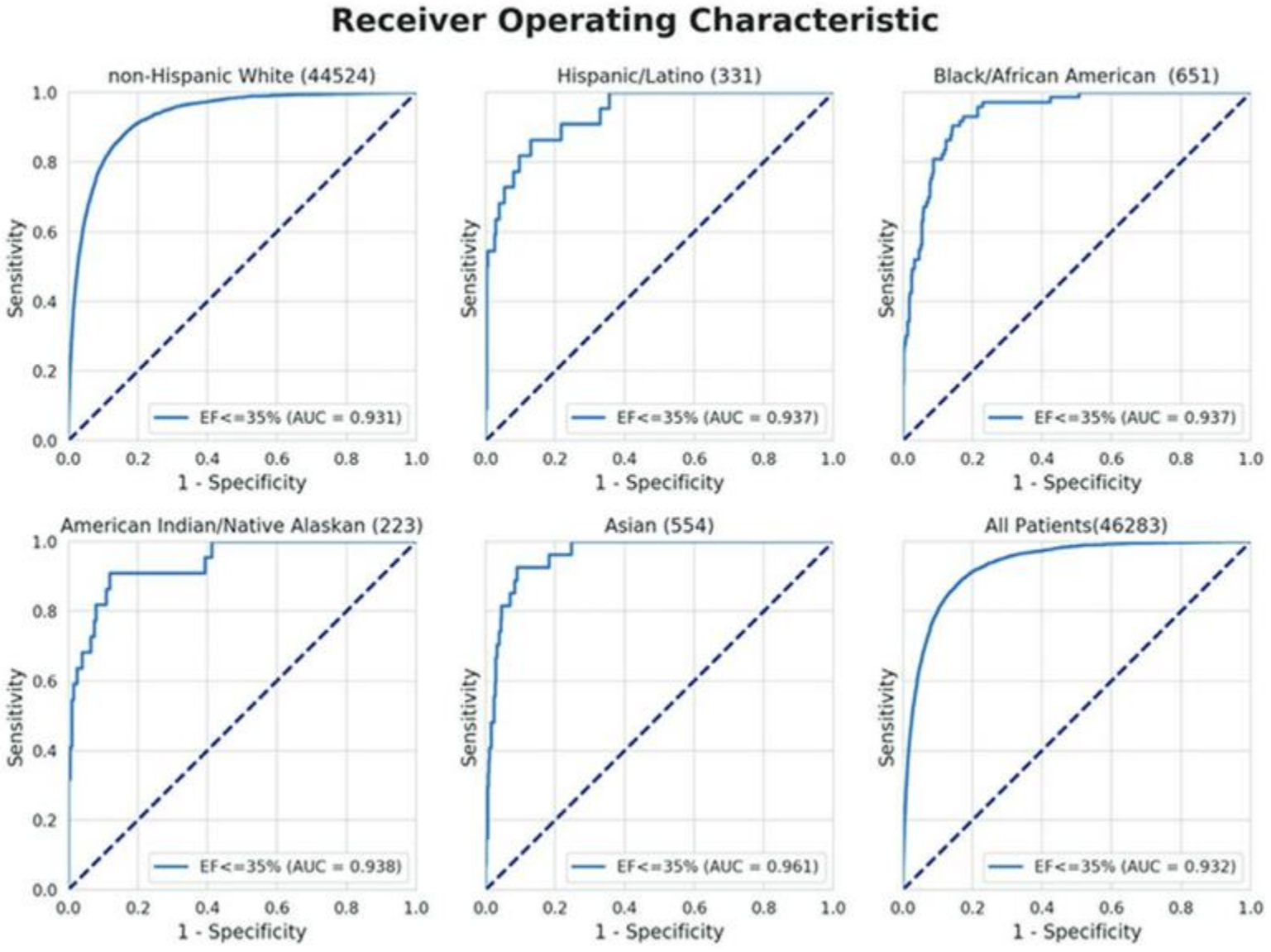
# EchoNext: detecting cardiac structural abnormality from ECG

Poterucha TJ, *et al*. Detecting structural heart disease from electrocardiograms using AI. Nature. 2025.



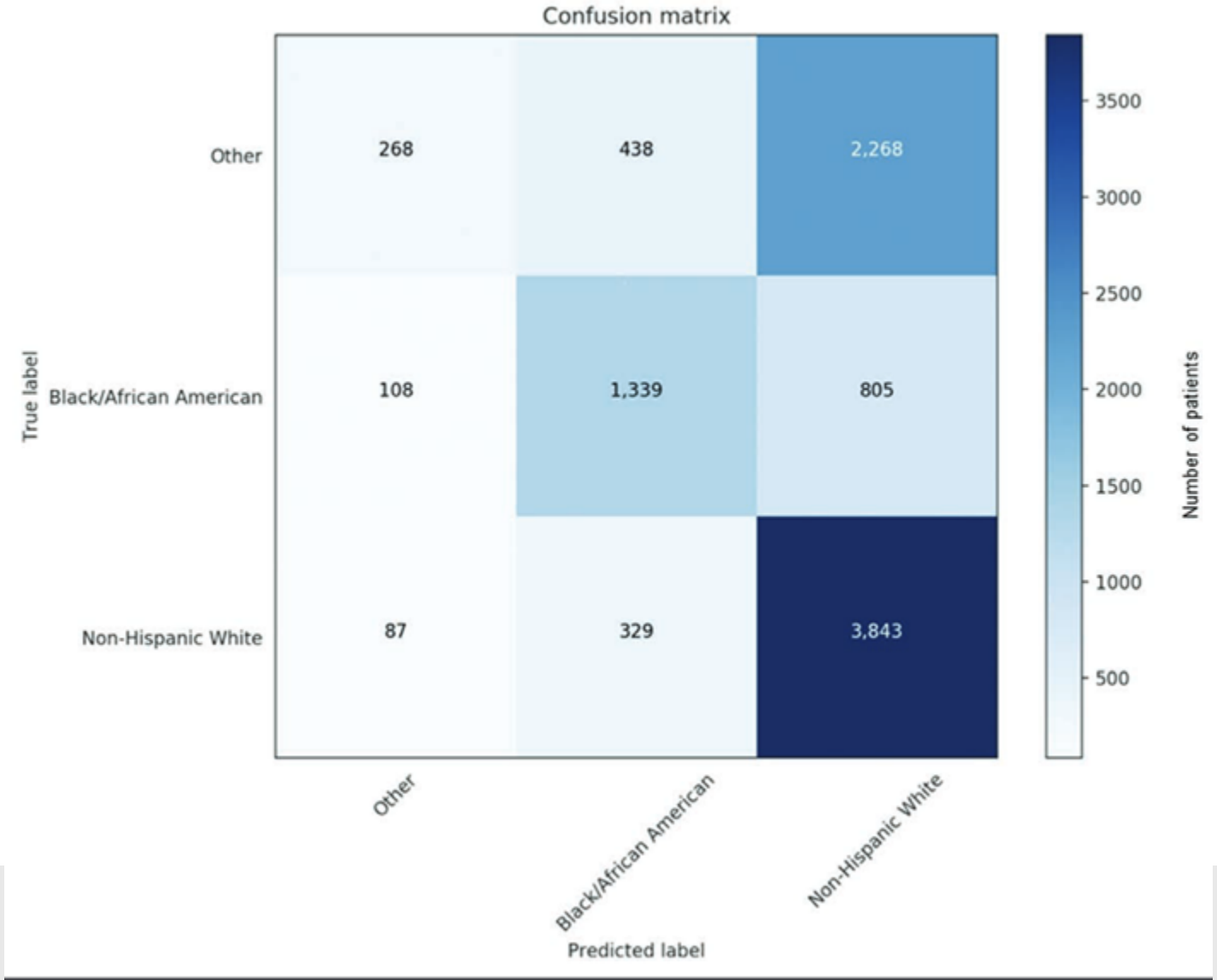



# Little bias by race and ethnicity for one ECG-LVEF model




Noseworthy PA, *et al*. Assessing and Mitigating Bias in Medical Artificial Intelligence: The Effects of Race and Ethnicity on a

# … but nonzero ability to distinguish race and ethnicity raises possibility of future bias



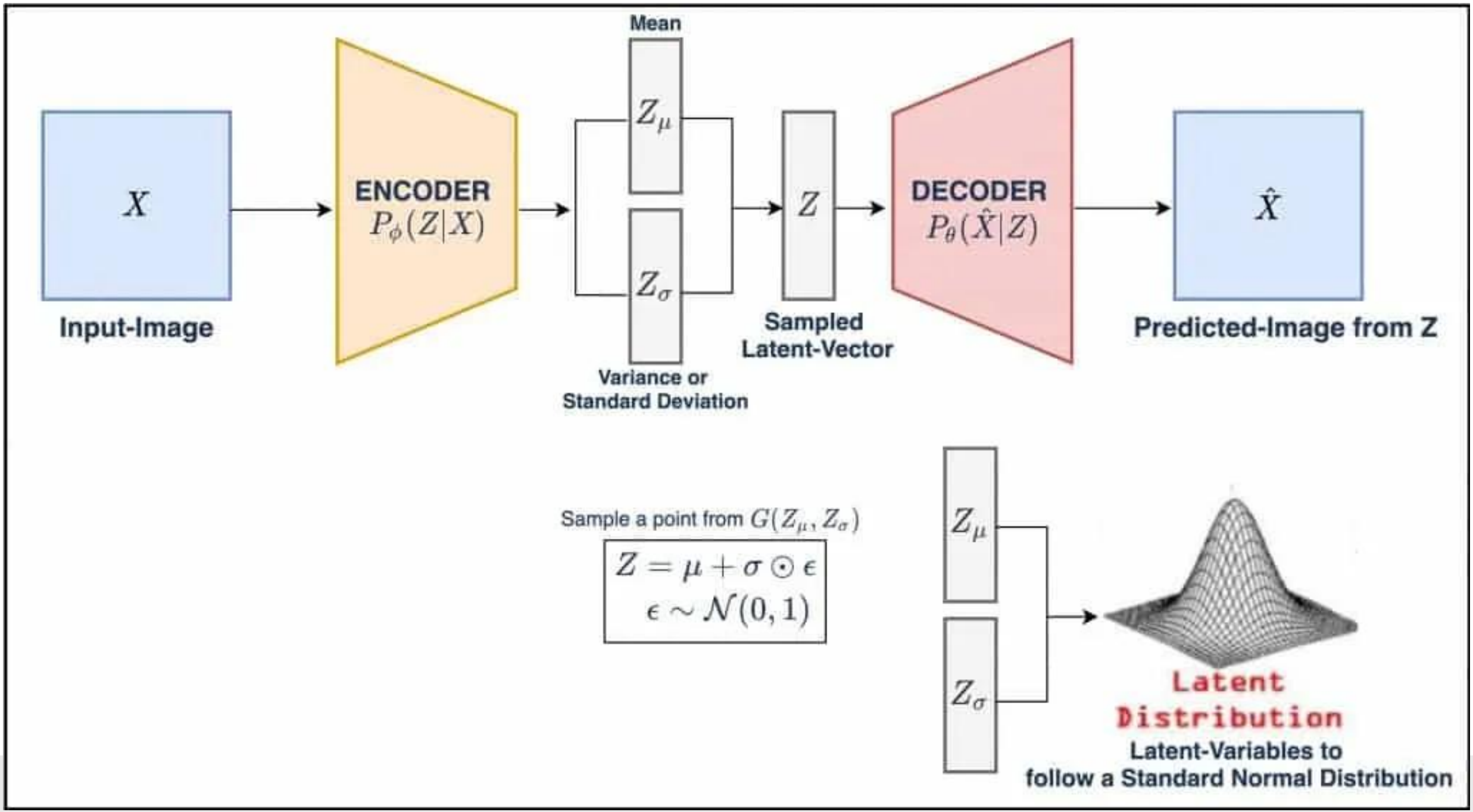

Noseworthy PA, *et al*. Assessing and Mitigating Bias in Medical Artificial Intelligence: The Effects of Race and Ethnicity on a

# I’d like to know:

- Have the ECG models saturated, or can we extract further improvements by continuing to scale up?

- Is there an ECG equivalent to the genetic concept of *heritability*, which would let us know this answer theoretically and not just empirically?

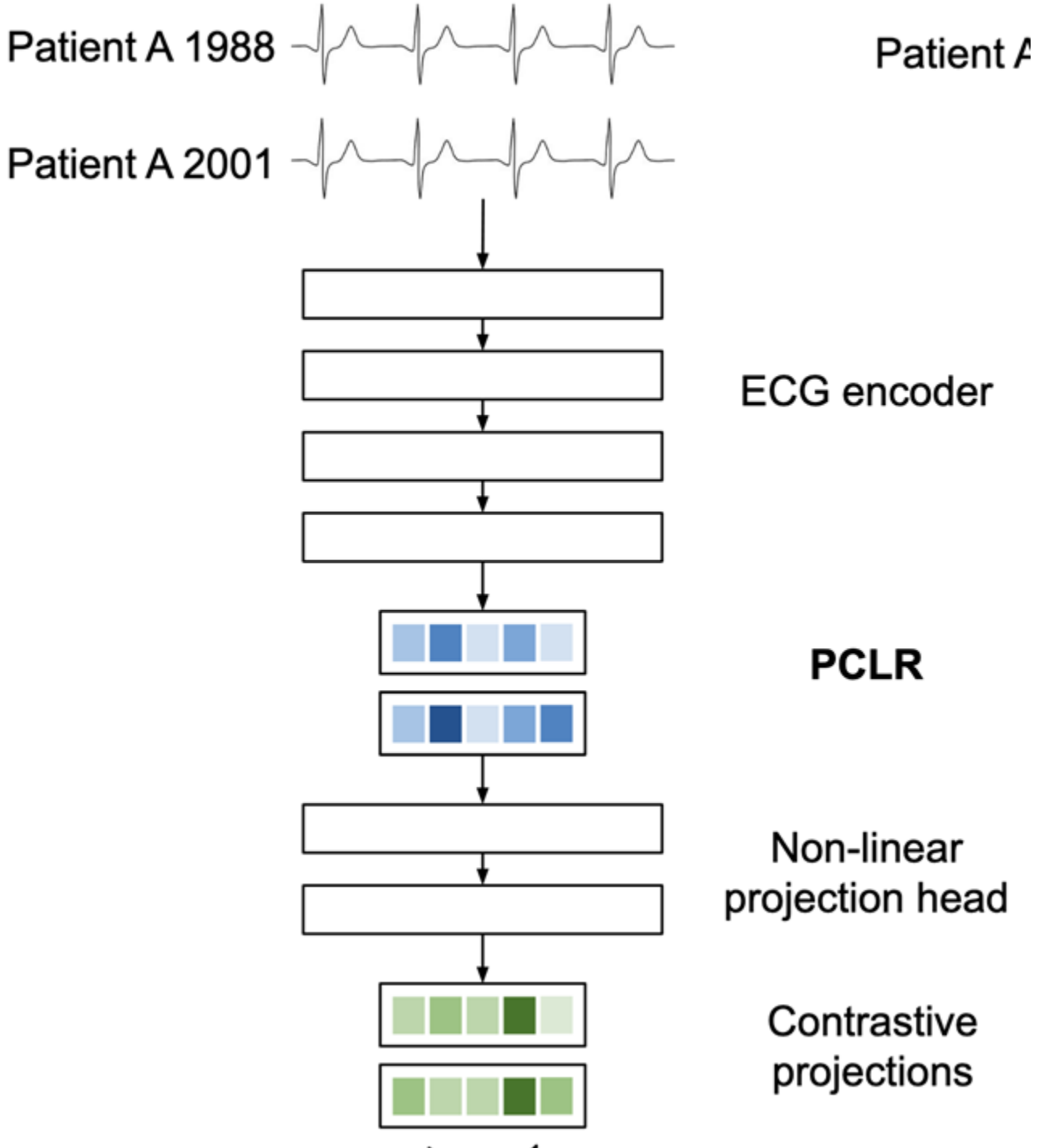
# Patient as latent space


Mean
$X$
ENCODER
$P_\phi(Z|X)$
$Z_\mu$
$Z_\sigma$
$Z$
DECODER
$P_\theta(\hat{X}|Z)$
$\hat{X}$
Input-Image
Variance or
Standard Deviation
Sampled
Latent-Vector
Predicted-Image from Z
Sample a point from $G(Z_\mu, Z_\sigma)$
$Z = \mu + \sigma \odot \epsilon$
$\epsilon \sim \mathcal{N}(0,1)$
$Z_\mu$
$Z_\sigma$
Latent
Distribution
Latent-Variables to
follow a Standard Normal Distribution

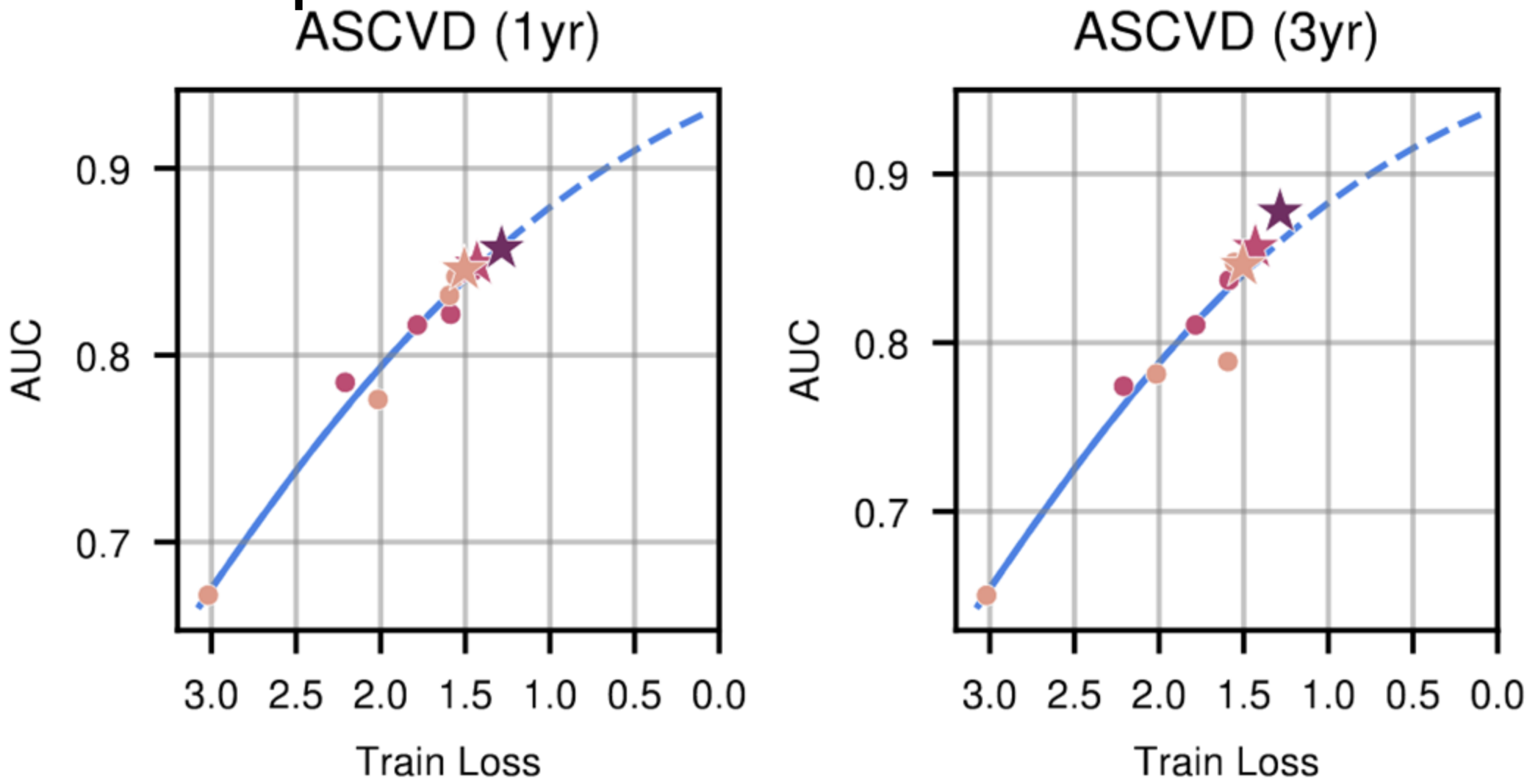
# Latent embeddings

Here, for ECG, but generally applicable.

Diamant N, Reinertsen E, Song S, Aguirre A, Stultz C, Batra P. Patient Contrastive Learning: a Performant, Expressive, and Practical Approach to ECG Modeling. arXiv:210404569 [cs, eess]. 2021.



# EHR-based generative models for short-term risk prediction



# Is it really about the patient?

Could these be powerful tools? Yes: physicians are smart.

The contents of the EHR are, in part, a function of what *physicians* decide to investigate; how they do so; and what they find.

May be appropriate to think of the EHR as a “sensor” that probes what physicians are already thinking about.

# Final thoughts

- Genetic risk
  - A once-in-a-lifetime blood sample yields (a small amount of) information about every disease
  - Greatest relative information gain is in low-risk groups
  - Could we eventually understand the pathways well enough to let us observe them all, obviating genetic risk?
- Proteomics
  - Which markers most useful short term? Long-term?
  - Value of repeat measures?
  - Future use as surrogate endpoints?
- Deep learning models of sensor data
  - ECG seems much easier to deploy broadly than imaging…
- Latent models from EHR data
  - If these are just sending the physician's diagnostic workup, may have little value for our purposes.

# Challenges

- Differential performance & benefit across sex, race, ethnicity
- Implementation challenges:
  - Proprietary algorithms: $$$
  - Open-source algorithms: whose job?
- Few proven preventive interventions in aortopathy, valvular heart disease, and cardiomyopathy
  - These gaps represent research opportunities
- Possibility of surrogate endpoints
  - How can we understand value for people with low near-term risk but elevated long-term risk?
  - How to demonstrate validity of surrogate endpoints?
- Knowledge is harm
  - Consequence of information for insurance, privacy

# I used your fancy new tool; now what?

We never build predictions from just one lab value. (The PREVENT score is not just LDL.) Why should we do them from just one ML tool?

The output of these new tools should be seen as **inputs** to comprehensive new models that can give us actionable **outputs** to help make rational decisions.