How Well, and Why
Looking inside clinical AI

John Ehrlinger, PhD
Lead Data Scientist, CORR |
November 4, 2026



Tall Rounds®

At Work

Who here has used an AI tool for something at work?



Why? And Trust?

Keep your hand up if you could explain why it gave you that answer.

And if you trusted it enough to act on it.



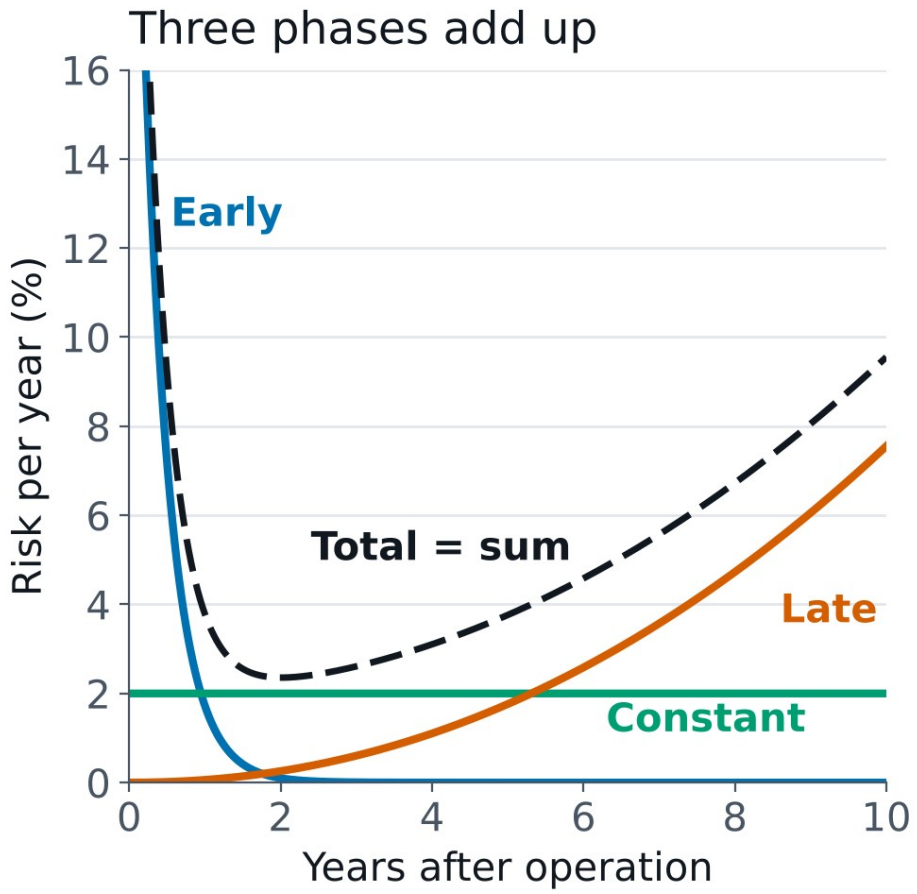
My Place

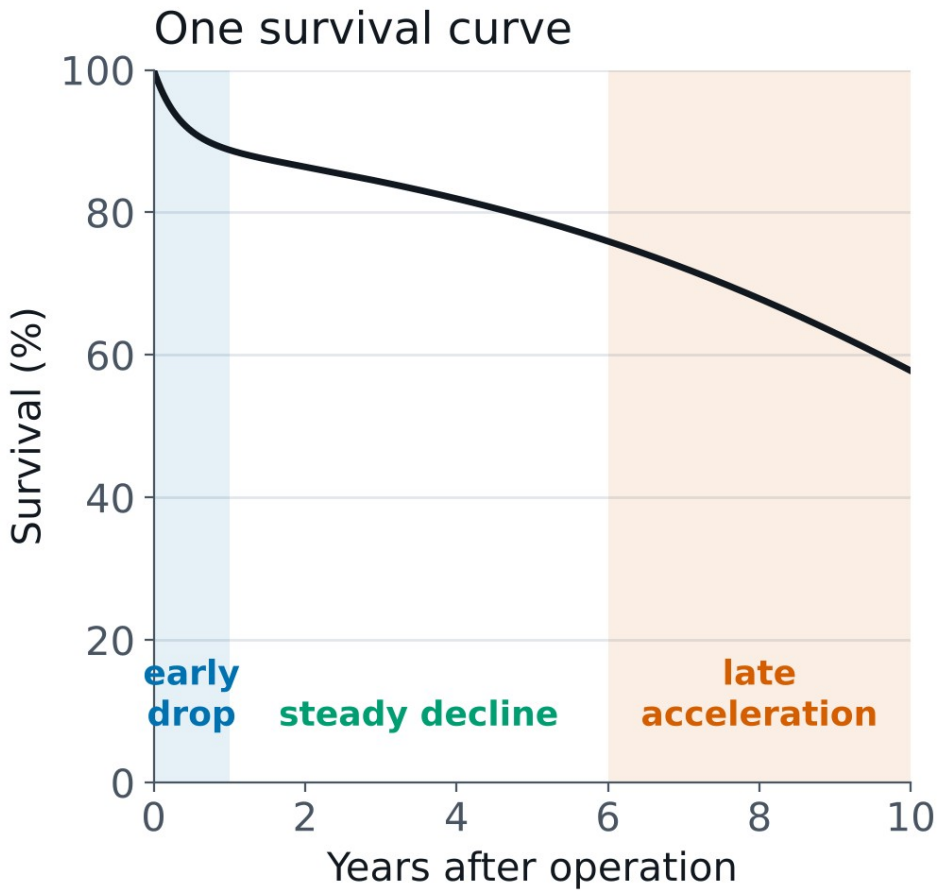
I am a statistician and engineer, not a clinician.

We build models, then look inside them.



Three Phases, One Curve





Looking Inside

Regression: read the coefficients. Machine learning: none to read.

VIMP: how much accuracy depends on each factor

VarPro: which factors the model relies on

Partial dependence: how risk changes as one factor changes

ALE: the same, using only patients near each value

SHAP: what drove one patient's prediction



Pump Thrombosis

2% to 8%

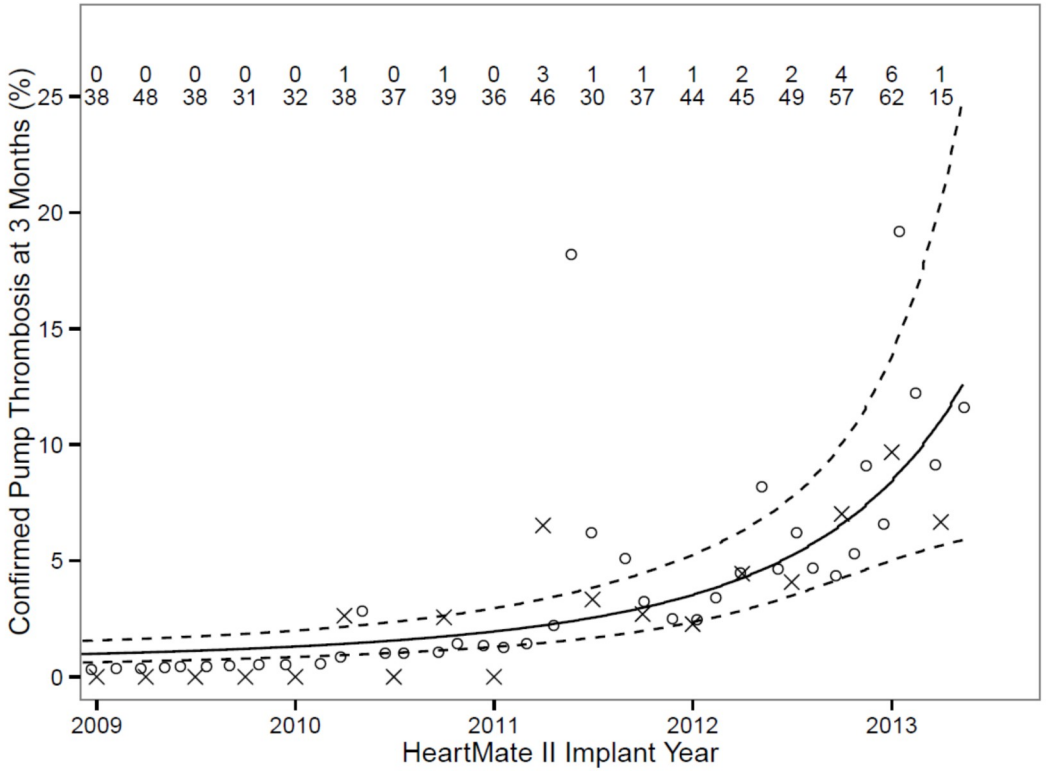
Confirmed thrombosis at 3 months, March 2011 to January 2013

895 devices, three centers

Starling RC, et al. N Engl J Med 2014;370:33-40

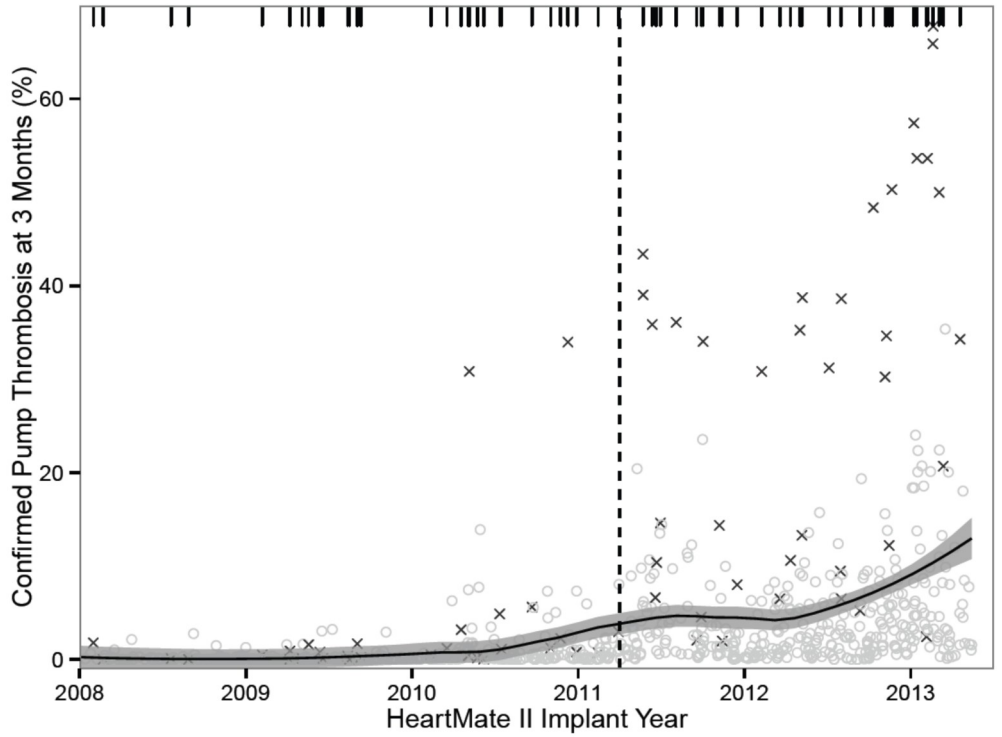


Two Models, One Answer



Hazard model (line), forest (circles)

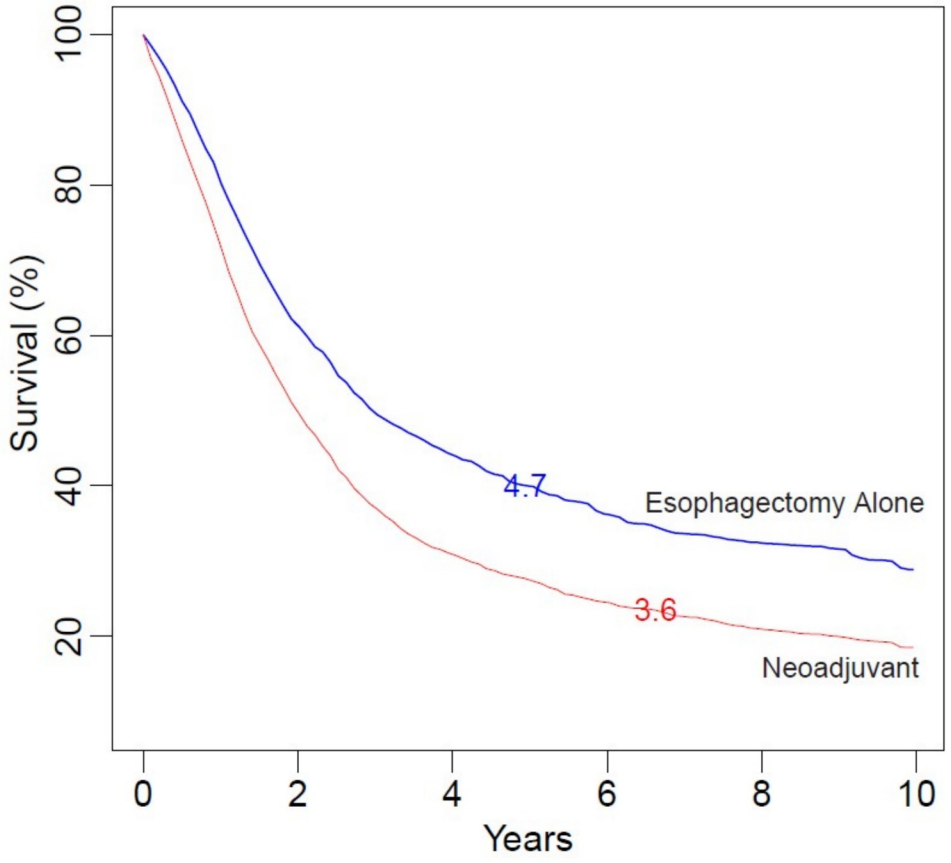
Starling RC, et al. NEJM 2014, Suppl. Fig. S6

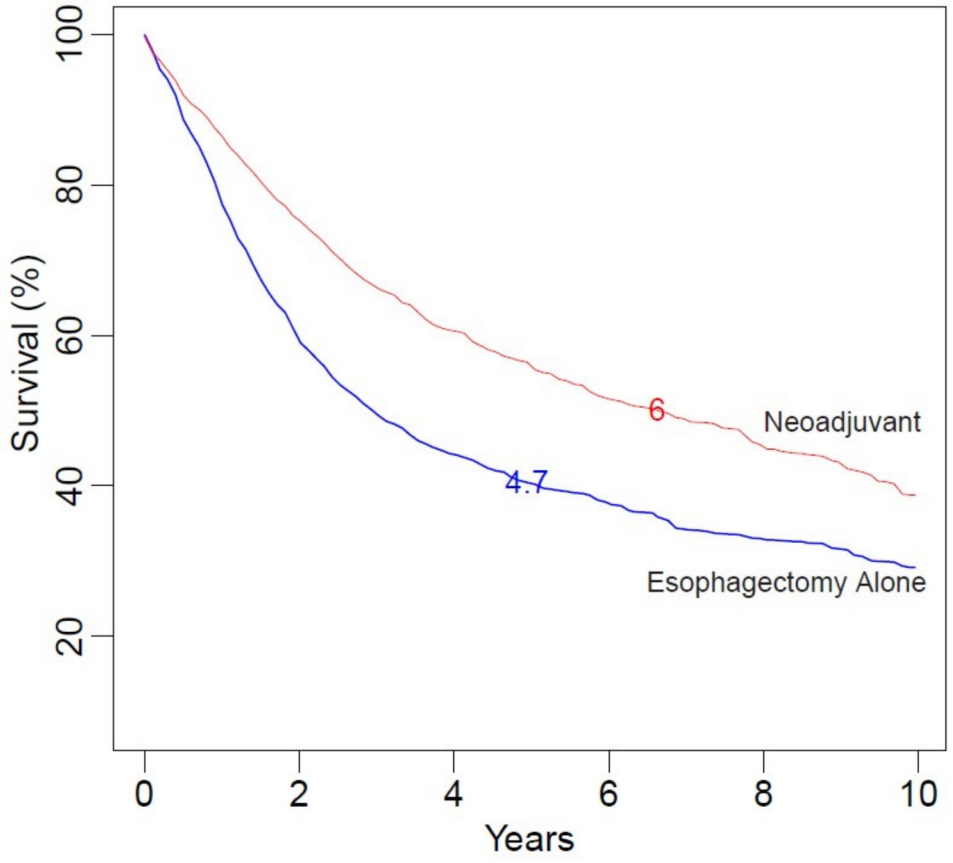


Forest: change point, 3/1/2011

Starling RC, et al. NEJM 2014, Suppl. Fig. S7A

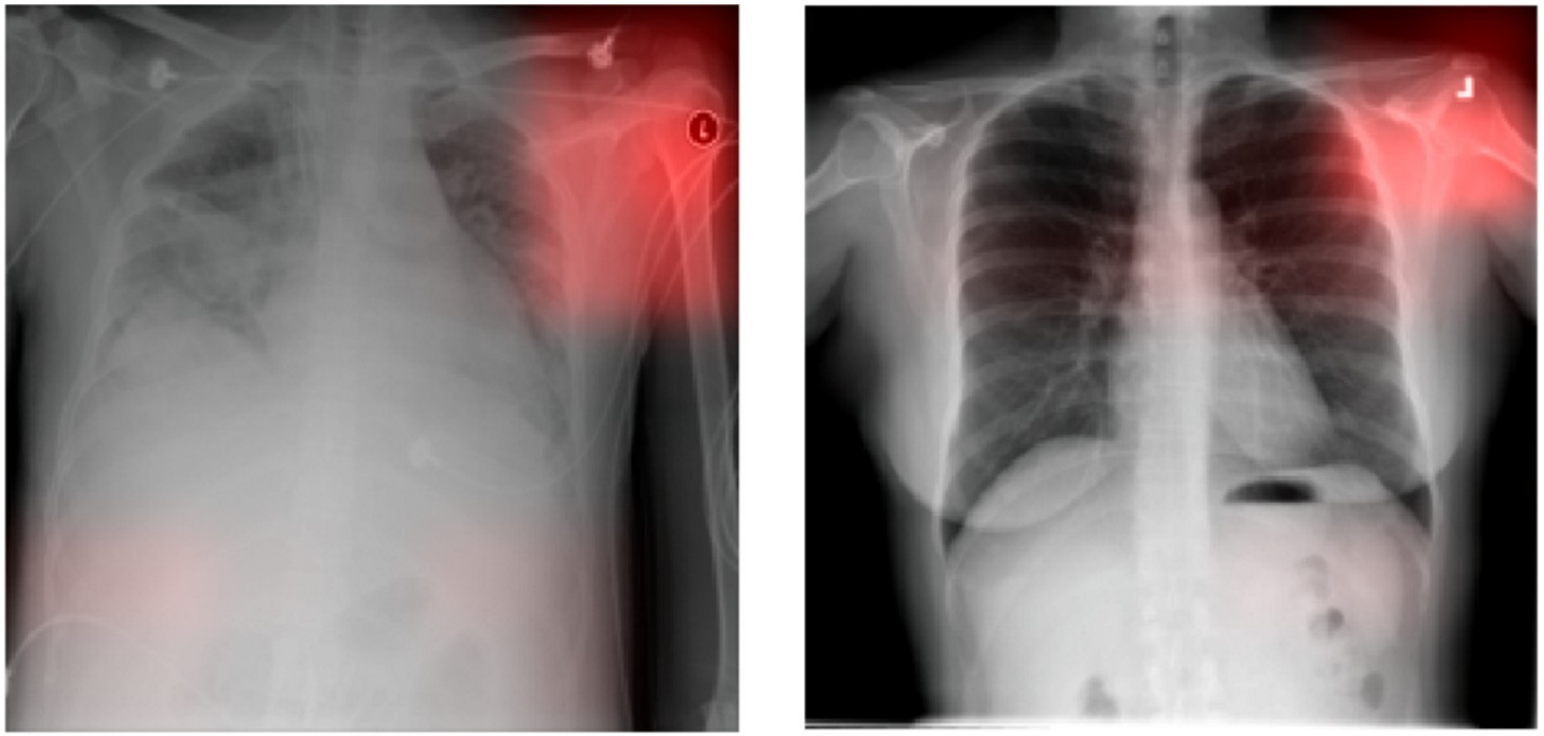
Two Patients, Two Answers





Rice TW, et al. J Thorac Oncol 2019;14:2164-75, Table 3

Images: What the Model Looked At



99.9%

of the time, a model could tell which hospital took the X-ray

Pneumonia: 34% of films at one hospital, about 1% at the other

Zech JR, et al. PLoS Med 2018;15(11):e1002683. CC BY 4.0

Language Models: Spoken With Authority

Large Language Models speak with authority.

Confident is not the same as correct.



Starting to See Inside

We are only now seeing what drives them, and they do not always learn the way people do.

A model's written reasoning is not always what it used.

Chen Y, et al. (Anthropic), Reasoning models don't always say what they think, 2025.

What You Can Do Now

Cite your sources: then open them

Use my documents: not your memory

What would change your answer?

Ask again: in different words

Test it: on cases where you know the answer

Was it tested on patients like mine?



Why, and Whether to Trust It

Do you know why it gave that answer?
Do you trust it?

All models are wrong, but some are useful.
George E. P. Box

