

Stroke Audit Machine Learning (SAMueL)

Learnings from explainable machine learning

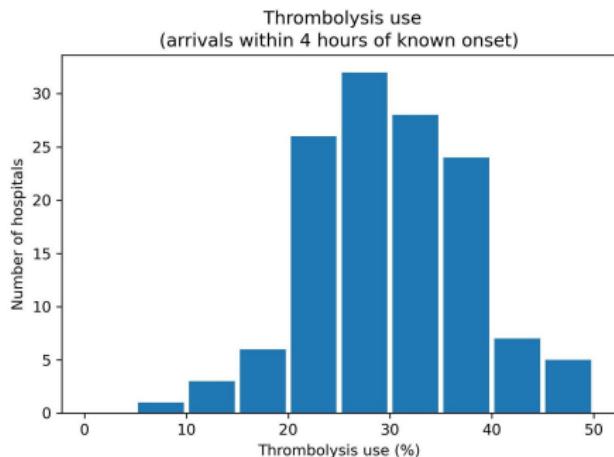
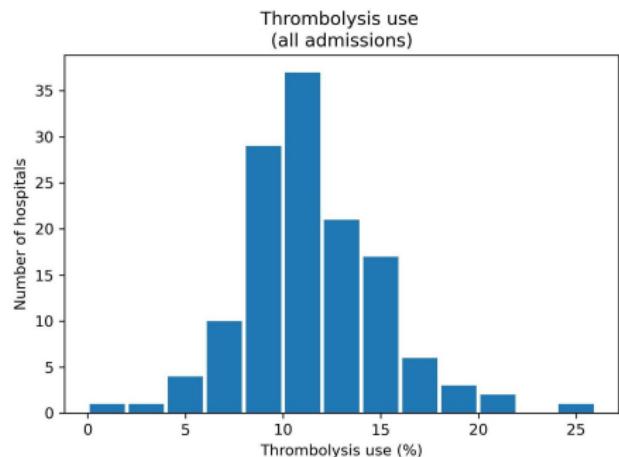
Kerry Pearn¹, Michael Allen^{1,3}, Anna Laws¹, Richard Everson³,
Martin James^{1,2}

¹University of Exeter Medical School ²Royal Devon University Healthcare NHS Foundation Trust ³University of Exeter Institute of Data Science and Artificial Intelligence

January 2023

Thrombolysis rates vary a lot

Here we show the range of thrombolysis use across the 132 acute stroke centres in England and Wales.



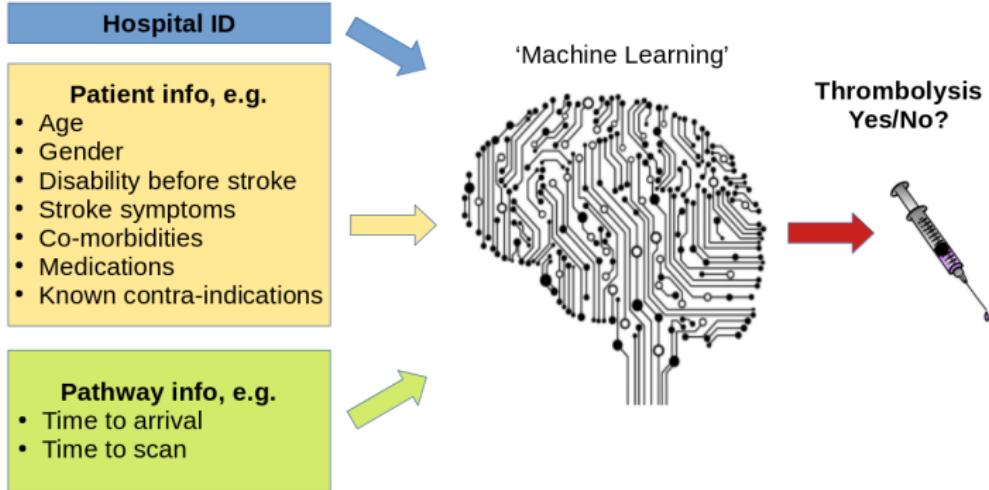
How much of this variation is due to differences in decisions hospitals make on who they would give thrombolysis to?

Our Approach

- ▶ Train a machine learning algorithm to learn to predict to which patients each hospital would, or would not, give thrombolysis.
- ▶ Identify the patient characteristics which are most important in predicting use of thrombolysis, and produce a simpler machine learning model based on these characteristics.
- ▶ Identify *ideal* candidates for thrombolysis.
- ▶ Add an *explainable machine learning* layer* to understand patterns of use of thrombolysis.
- ▶ Investigate what we think would happen if every hospital received exactly the same 10k cohort of patients.
- ▶ Identify key subgroups of patients where hospitals differ in their use of thrombolysis.

*SHAP (Shapley Additive Values, which provide information on what contribution each patient characteristic makes to the final prediction of whether they will, or will not, receive thrombolysis).

Machine learning overview

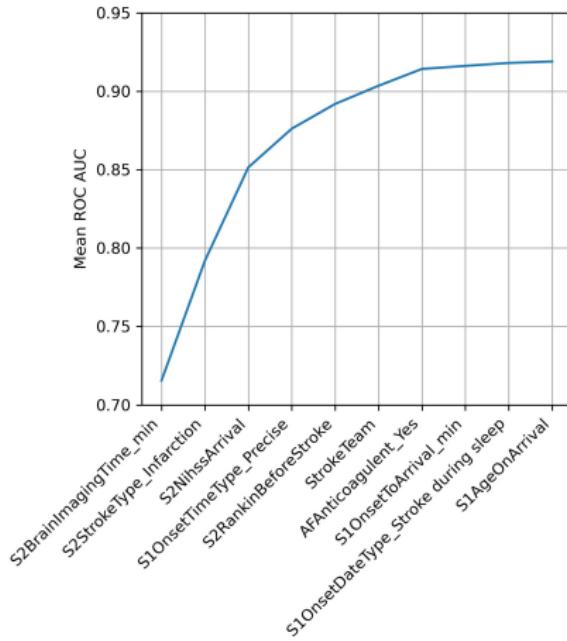
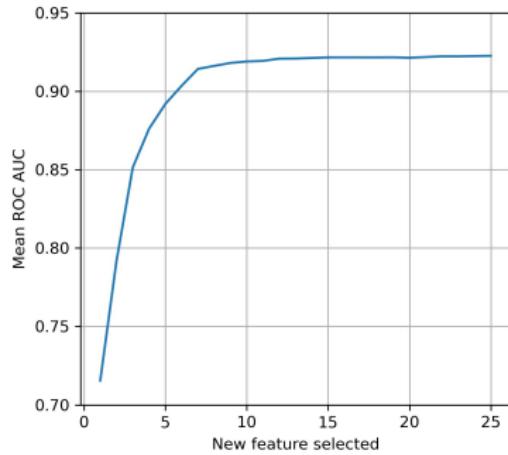


Machine learning is based on the simple principle of recognising similarity to what has been seen before.

We accessed 246,676 emergency stroke admissions in England and Wales over three years. Our machine learning models use XGBoost classification, and are based on those 88,928 patients who arrive within 4 hours of known stroke onset.

Model accuracy, and simplification

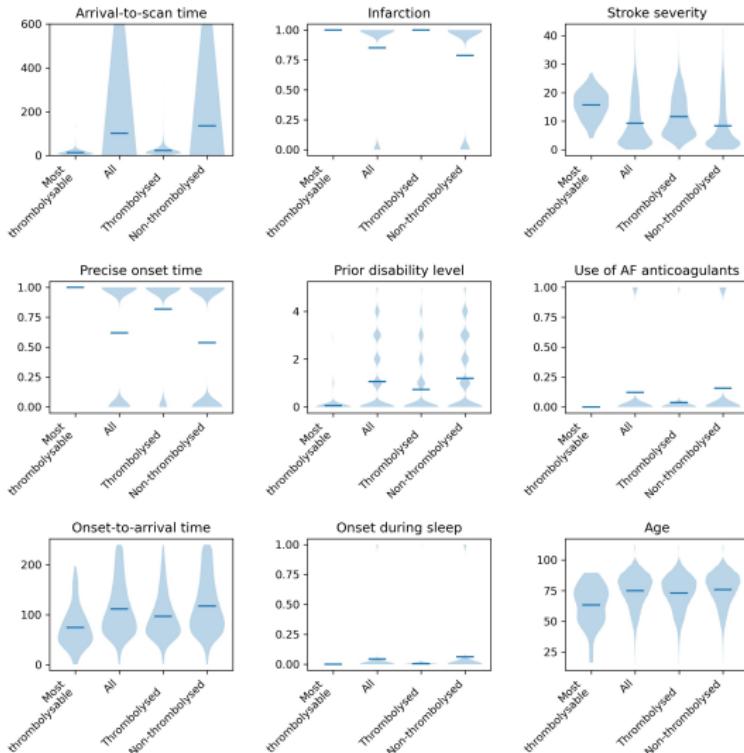
A model with all available 84 features had an ROC AUC of 0.922. A model with 10 features had an ROC AUC of 0.919.



Our simpler machine learning model will just use these 10 features.

What do the most thrombolysable patients look like?

For each hospital we identify the patient with the highest probability of receiving thrombolysis. Here we compare the range of feature values for these 'Most thrombolysable' patients with 1) All patients, 2) Thrombolysed patients, and 3) Non-thrombolysed patients.



What do the most thrombolysable patients look like?

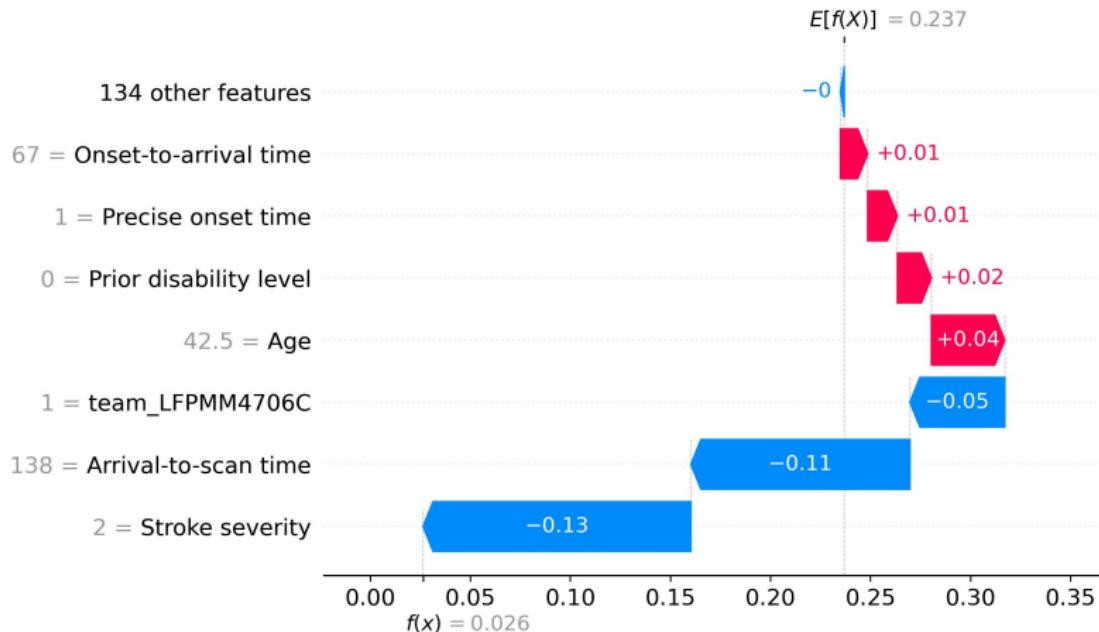
Compared with other patients, the most thrombolysable patients:

- ▶ Had shorter arrival-to-scan times
- ▶ Had an infarction stroke type
- ▶ Had stroke severity with NIHSS 5-25
- ▶ Had a precise onset time
- ▶ Had lower pre-stroke disability
- ▶ Were not taking anticoagulant medication
- ▶ Had shorter onset-to-arrival times
- ▶ Did not have onset during sleep
- ▶ Were younger

But, can we learn to be more *quantitative* than this?

Explaining model predictions with SHAP values

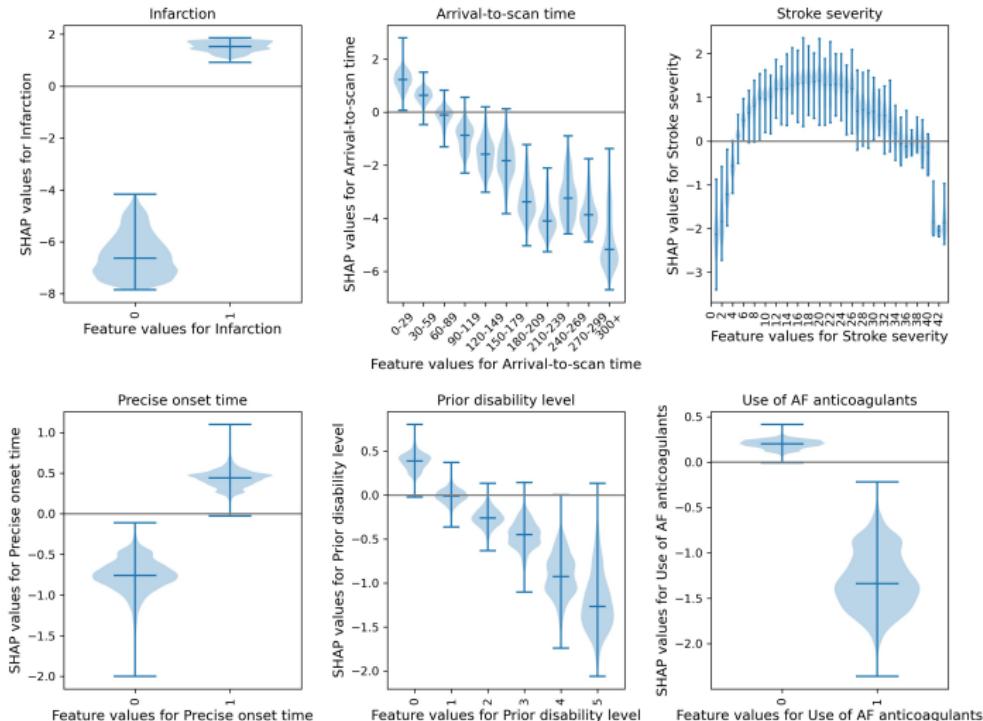
SHAP values show the influence of features (even for '*black box*' models).



$E[f(X)] = 0.237$

What drives use of thrombolysis across all hospitals?

Note: SHAP values here are *log odds*. Each step-change in value of ± 1 changes the chances of receiving thrombolysis about 3-fold. (Plots are in order of feature importance.)



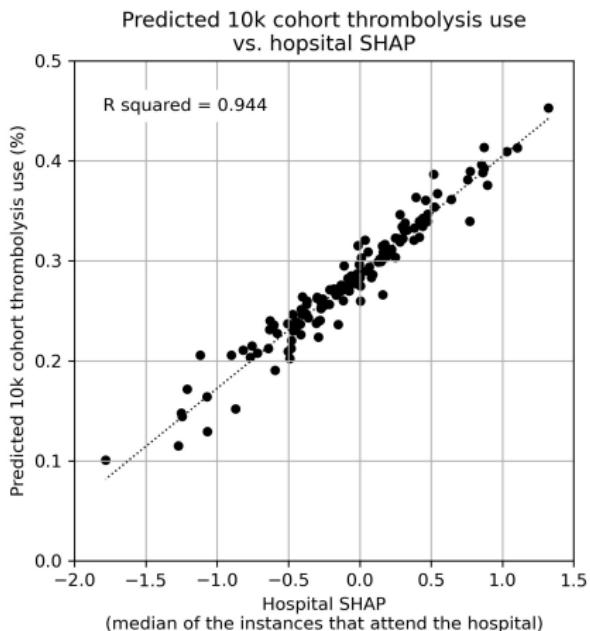
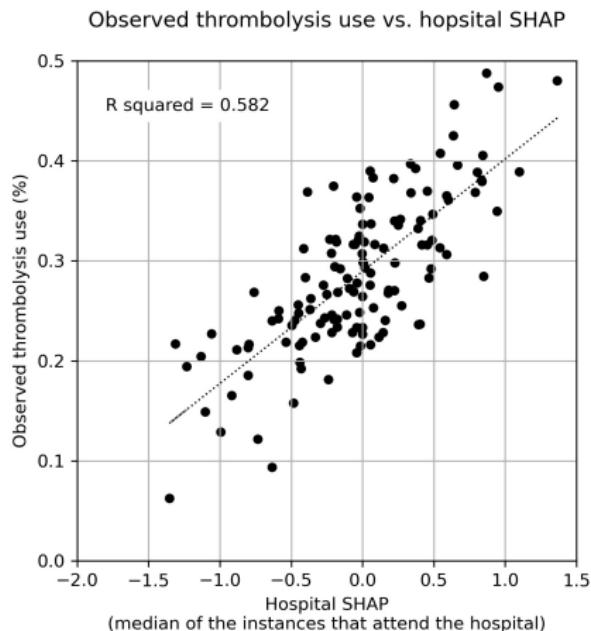
What drives use of thrombolysis across all hospitals?

The XGBoost/SHAP model revealed that the odds of receiving thrombolysis:

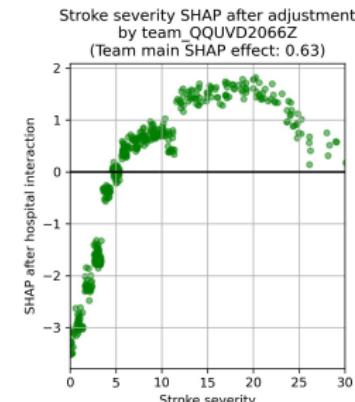
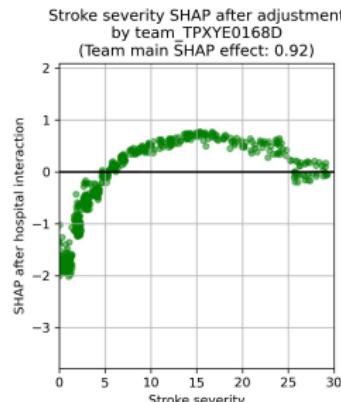
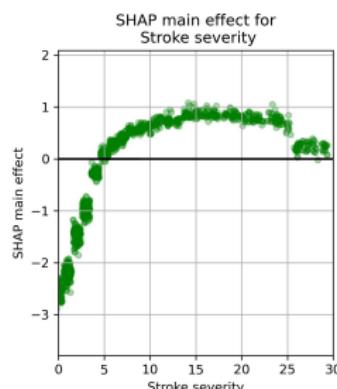
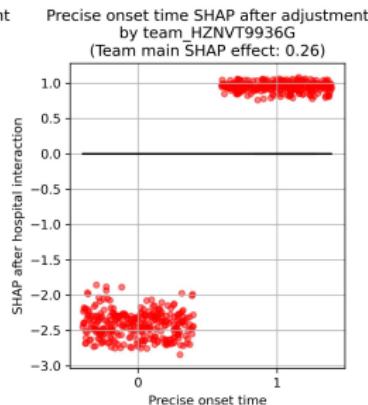
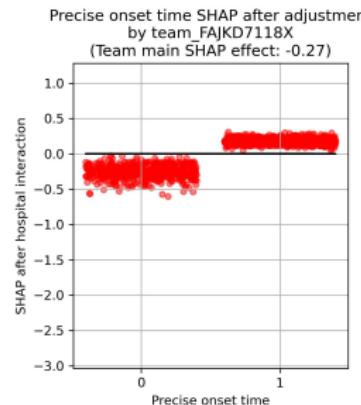
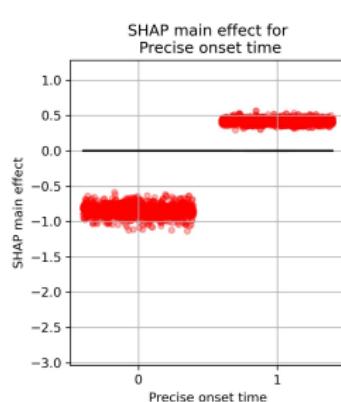
- ▶ Reduced 20 fold over the first 100 minutes of arrival-to-scan time.
- ▶ Varied 30 fold depending on stroke severity.
- ▶ Reduced 3 fold with imprecise onset time.
- ▶ Fell 5 fold with increasing pre-stroke disability.
- ▶ Varied 15 fold between hospitals.

Hospital SHAP predicts a hospital's predisposition to use thrombolysis

Here we show the relationship of hospital SHAP with the hospital thrombolysis use for 1) the hospitals own patients, and 2) a common 10k cohort of patients.

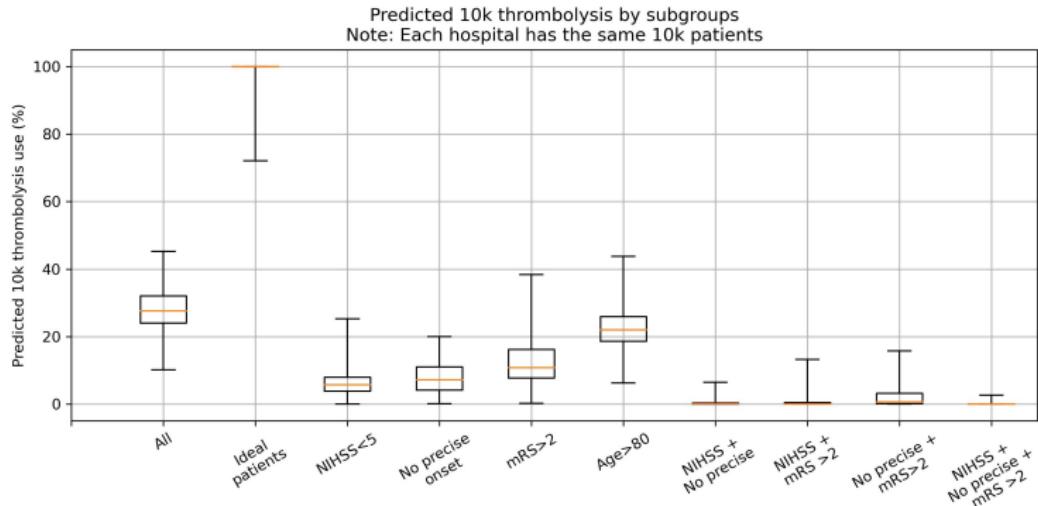


How general effects may be modified by individual hospitals



Thrombolysis in subgroups of patients (predicted use in a common 10k cohort at each team)

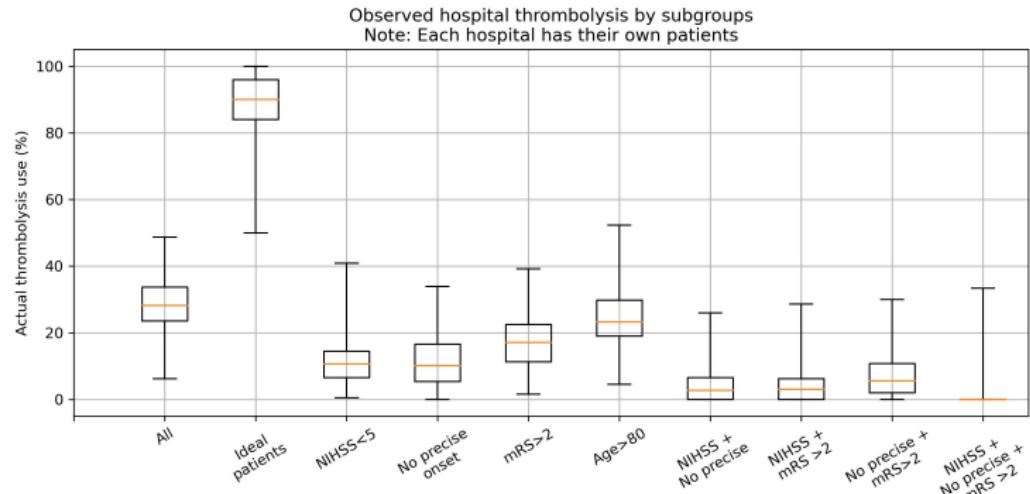
The range of the predicted thrombolysis use across the 132 hospitals for subsets of the 10k patient cohort. The subsets of patients to the right of *Ideal patients* all have the ideal patient characteristics with the specified changes



An *ideal patient* has: Stroke severity NIHSS in range 10-25, Arrival-to-scan time < 20 minutes, Stroke type = infarction, Precise onset time = True, Prior disability level (mRS) = 0, No use of AF anticoagulants, Onset-to-arrival time < 90 minutes, Age < 80 years, Onset during sleep = False

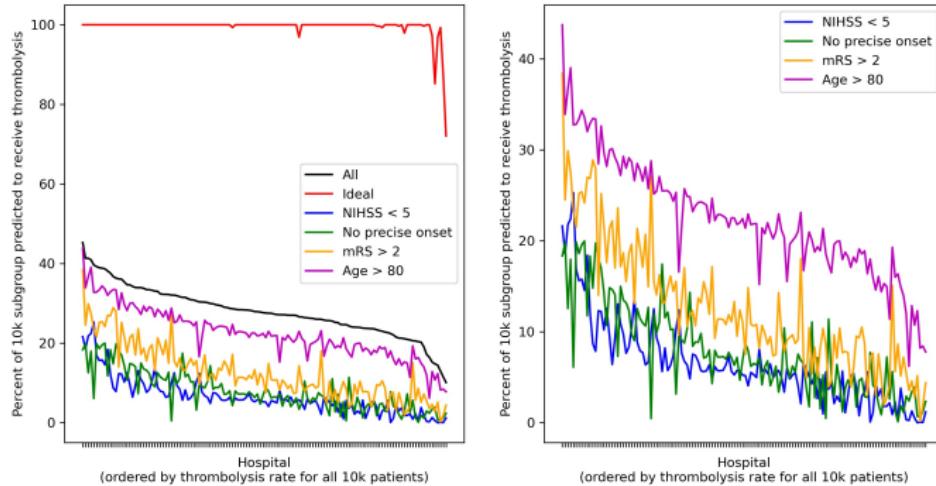
Thrombolysis in subgroups of patients (observed use in each hospitals own patients)

The range of the observed thrombolysis use across the 132 hospitals for subsets of the hospitals own patients. The subsets of patients to the right of *Ideal patients* all have the ideal patient characteristics with the specified changes.



An *ideal patient* has: Stroke severity NIHSS in range 10-25, Arrival-to-scan time < 20 minutes, Stroke type = infarction, Precise onset time = True, Prior disability level (mRS) = 0, No use of AF anticoagulants, Onset-to-arrival time < 90 minutes, Age <80 years, Onset during sleep = False

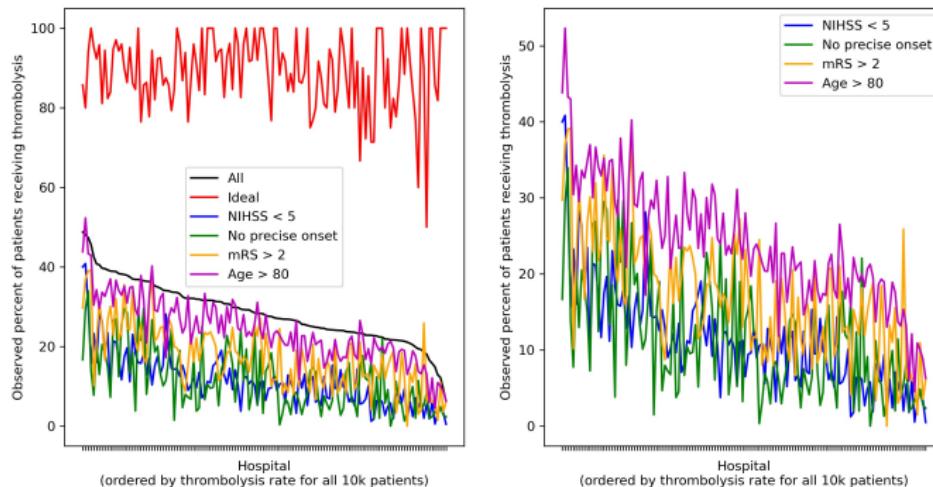
Thrombolysis in subgroups of patients (predicted use in a common 10k cohort at each hospital)



Hospitals with lower thrombolysis use were particularly less likely to give thrombolysis to patients with milder strokes, prior disability, or patients with imprecise onset time.

Note: Subgroups, other than *ideal patients*, tend to track each other.

Thrombolysis in subgroups of patients (observed use in each hospitals own patients)



Note: Subgroups, other than *ideal patients*, tend to track each other, but with more noise than predicted thrombolysis use.

Investigating how hospitals differ in thrombolysis decision-making: Artificial patients

Base patient:

- ▶ Onset to arrival = 80 mins
- ▶ Arrival to scan = 20 mins
- ▶ Infarction = Yes
- ▶ NIHSS = 15
- ▶ Prior disability level = 0
- ▶ Precise onset time = Yes
- ▶ Use of AF anticoagulents = No

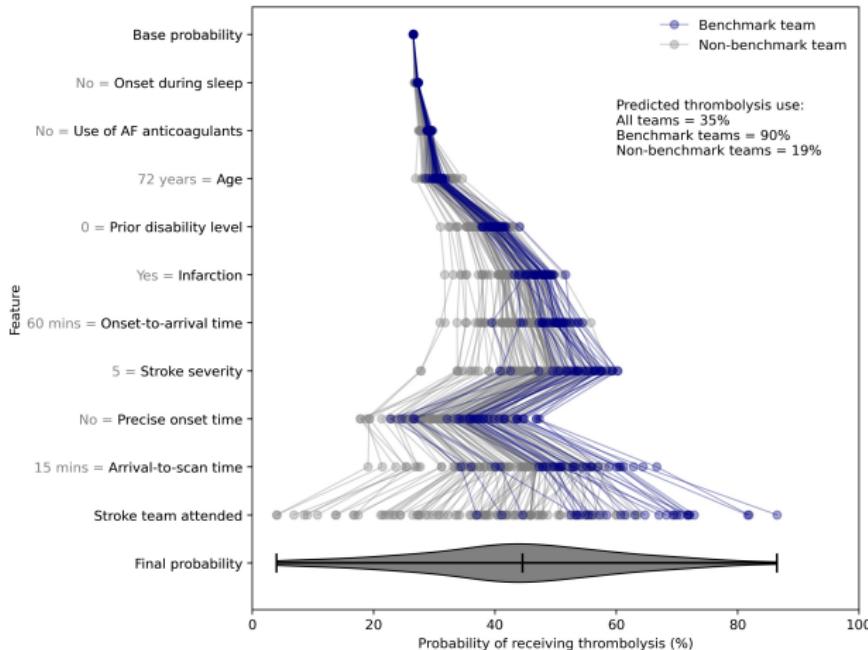
Proportion of hospitals predicted to give thrombolysis:

- ▶ Base patient: 99%

Base patient with:

- ▶ NIHSS = 4: 73%
- ▶ Pre-stroke mRS = 3: 86%
- ▶ Estimated stroke onset time: 64%

How would different teams respond to the same patient?



The patient starts with the same base probability, then each hospital reacts differently when presented with each of the patients characteristics and nudges the probability to a different level.

Summary

- ▶ The XGBoost/SHAP model revealed that the odds of receiving thrombolysis:
 - ▶ Reduced with increasing arrival-to-scan time.
 - ▶ Varied 30 fold depending on stroke severity.
 - ▶ Reduced with imprecisely known onset time.
 - ▶ Fell with increasing pre-stroke disability.
 - ▶ Varied 15 fold between hospitals.
- ▶ The hospital identification (hospital SHAP value) explained 58% of the variance in between-hospital thrombolysis use.
- ▶ Compared with hospitals with higher thrombolysis use, hospitals with lower use were particularly less likely to give thrombolysis to patients with milder strokes, prior disability, or patients with imprecise onset time.