

# Stroke Audit Machine Learning (SAMueL) Advisory Group November 2022

Investigating variation in thrombolysis use with clinical pathway  
simulation and explainable AI

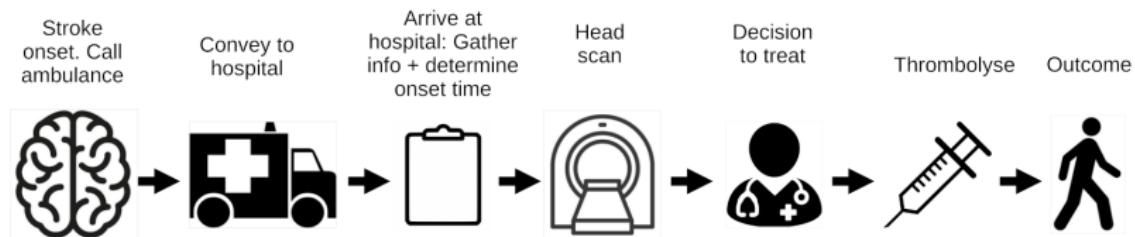
November 2022

# Outline

- 1 Modelled stroke pathway
- 2 SAMueL-1 summary
- 3 Machine learning - learning and comparing decisions to thrombolysis between hospitals
- 4 Stroke outcome modelling based on times to treatment with thrombolysis and thrombectomy
- 5 HQIP data request
- 6 Qualitative research
- 7 Further project materials

Modelled stroke pathway

# Breaking down the emergency stroke pathway into key steps



We can model key changes to pathway:

- What if the pathway were faster?
- What if hospital determined the stroke onset time in more patients?
- What if clinical decision-making was like that of *benchmark* hospitals?  
(Predict what treatment a patient would receive at other hospitals).

We model these changes with a hospital's own patient population, to allow for inter-hospital variation in patient population characteristics.

SAMueL-1 summary

# SAMueL-1 Summary: What is the problem?

There is a gap between target thrombolysis use (20%) and actual thrombolysis use (11–12%) in emergency stroke care

Clinical expert opinion on what *should be* happening



What is happening?



Unknown onset time or  
arrived too late to treat



Not suitable for treatment  
with thrombolysis



Treated with thrombolysis



Potentially treatable, but not  
treated with thrombolysis

## SAMueL-1 Summary: What did we test?

We used clinical pathway simulation and machine learning to analyse a series of *what if?* questions:

- What if arrival-to-treatment time was 30 minutes?
- what if all hospitals determined stroke onset time as frequently as an *upper quartile* hospital (a hospital ranked 25 out of 100, for determining stroke onset time).
- What if decisions to thrombolyse were made according to a majority vote of 30 benchmark hospitals?

For each hospital we use their own patients to ask these questions, to allow for differences in local patient populations.

# SAMueL-1 Summary: What did we find?

We found that making all these changes would increase thrombolysis use in England and Wales to 18–19%. Out of every 10 patients who were potentially treatable but did not receive treatment, we found the cause to be:



Hospital processes  
were **too slow**



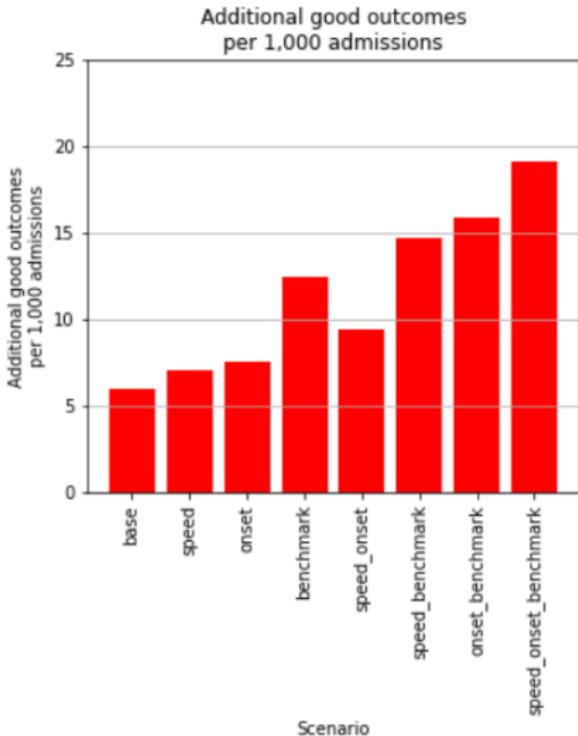
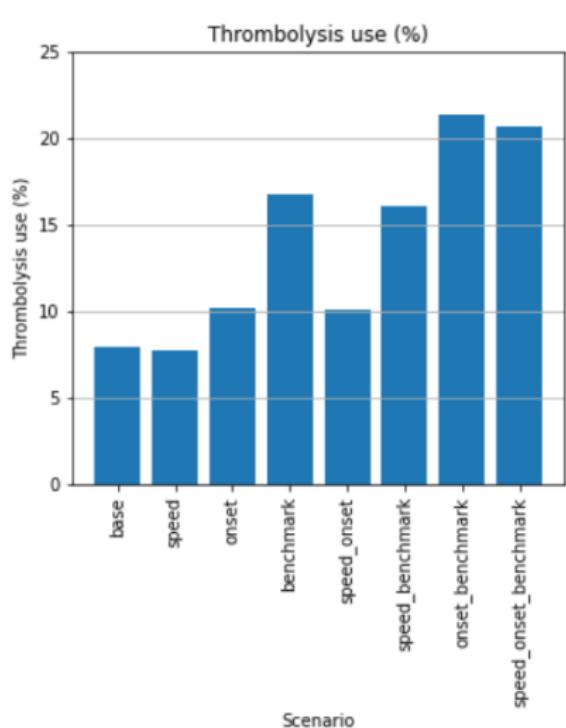
Stroke onset time was not  
**determined** when it  
potentially could have been



Doctors chose not to use thrombolysis  
when other higher-thrombolytic  
hospitals would have done

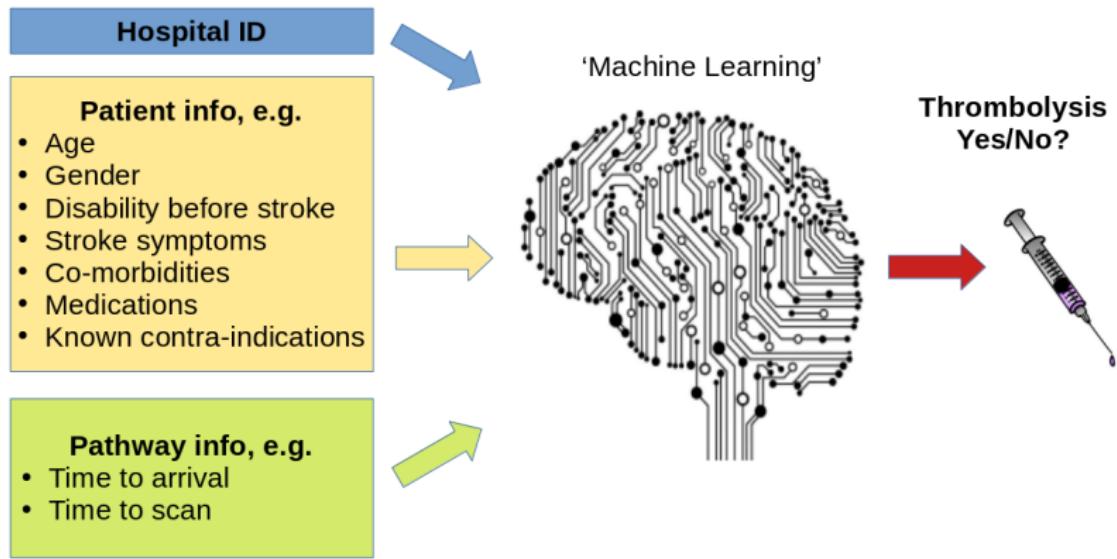


# Applying our models at hospital level



Machine learning - learning and comparing decisions  
to thrombolyse between hospitals

# Machine learning overview



Machine learning (and nearly all *artificial intelligence*) is based on the simple principle of recognising similarity to what has been seen before.

We accessed 240,000 emergency stroke admissions in England and Wales over three years.

# Model accuracy, and simplification

Our machine learning models use XGBoost classification, and are based on all patients who arrive within 4 hours of known stroke onset.

The full model has 61 patient features:

- Overall accuracy = 85.2%
- Best combined sensitivity and specificity = 84.3%
- ROC AUC = 0.921

A simplified model with 8 features

- Overall accuracy = 84.8%
- Best combined sensitivity and specificity = 83.8%
- ROC AUC = 0.916

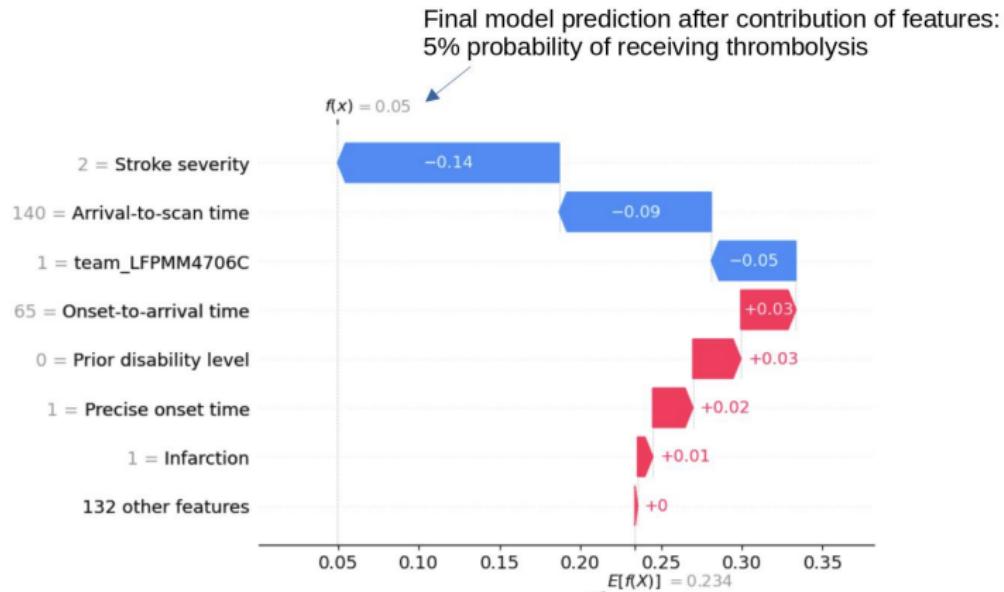
The 8 features of the simplified model are:

- ① Arrival-to-scan time
- ② Stroke type (infarction/haemorrhage)
- ③ Stroke severity (NIHSS)
- ④ Precise or estimated stroke onset time
- ⑤ Prior disability level (mRS)
- ⑥ Stroke team
- ⑦ Use of AF anticoagulants
- ⑧ Onset-to-arrival time

There are only very weak correlations between the selected features with no R-squared being greater than 0.05.

# Explaining model predictions with SHAP values

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

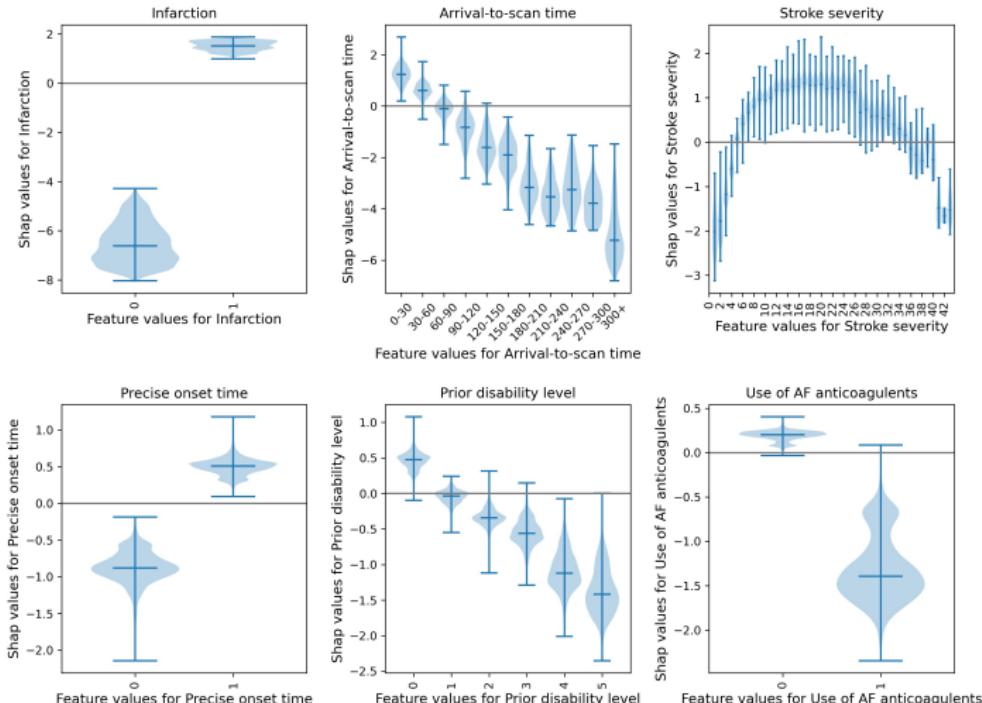


Base model prediction before contribution of features:  
23% probability of receiving thrombolysis

Note: The effect of team ID can depend on other patient characteristics.

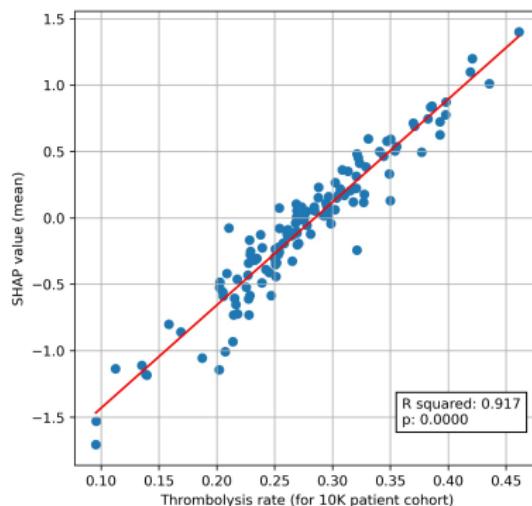
# What drives use of thrombolysis across all hospitals?

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



# Comparing hospital SHAP value with the predicted thrombolysis rate if all hospitals say the same 10k patients

- We can assess each hospital's 'propensity to use thrombolysis' by passing the same 10k cohort of patients through all hospital prediction models.
- We can compare this 10k thrombolysis rate to the average hospital SHAP.
- These two different methods of isolating the effect of the hospital on thrombolysis use give very similar results.



## Investigating how hospitals differ in thrombolysis decision-making (Patient 1: Base patient)

Assuming there are no reasons not stated here to exclude a patient from use of thrombolysis, would you give this patient thrombolysis?

- 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

Our model predicts 131 out of 132 (99%) hospitals would give this patient thrombolysis.

## Investigating how hospitals differ in thrombolysis decision-making (Patient 2: Milder stroke)

Assuming there are no reasons not stated here to exclude a patient from use of thrombolysis, would you give this patient thrombolysis?

- Onset to arrival = 80 mins
- Arrival to scan = 20 mins
- Infarction = Yes
- NIHSS = 4
- Prior disability level = 0
- Precise onset time = Yes
- Use of AF anticoagulents = No

Our model predicts 97 out of 132 (73%) hospitals would give this patient thrombolysis.

## Investigating how hospitals differ in thrombolysis decision-making (Patient 3: Pre-stroke disability)

Assuming there are no reasons not stated here to exclude a patient from use of thrombolysis, would you give this patient thrombolysis?

- Onset to arrival = 80 mins
- Arrival to scan = 20 mins
- Infarction = Yes
- NIHSS = 15
- *Prior disability level = 3\**
- Precise onset time = Yes
- Use of AF anticoagulents = No

Our model predicts 114 out of 132 (86%) hospitals would give this patient thrombolysis.

\*Moderate disability; requires some help, but able to walk without assistance.

## Investigating how hospitals differ in thrombolysis decision-making (Patient 4: Estimated stroke onset time)

Assuming there are no reasons not stated here to exclude a patient from use of thrombolysis, would you give this patient thrombolysis?

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

Our model predicts 84 out of 132 (64%) hospitals would give this patient thrombolysis.

# Machine Learning key findings

General observations about thrombolysis use: The chance of receiving thrombolysis is increased by:

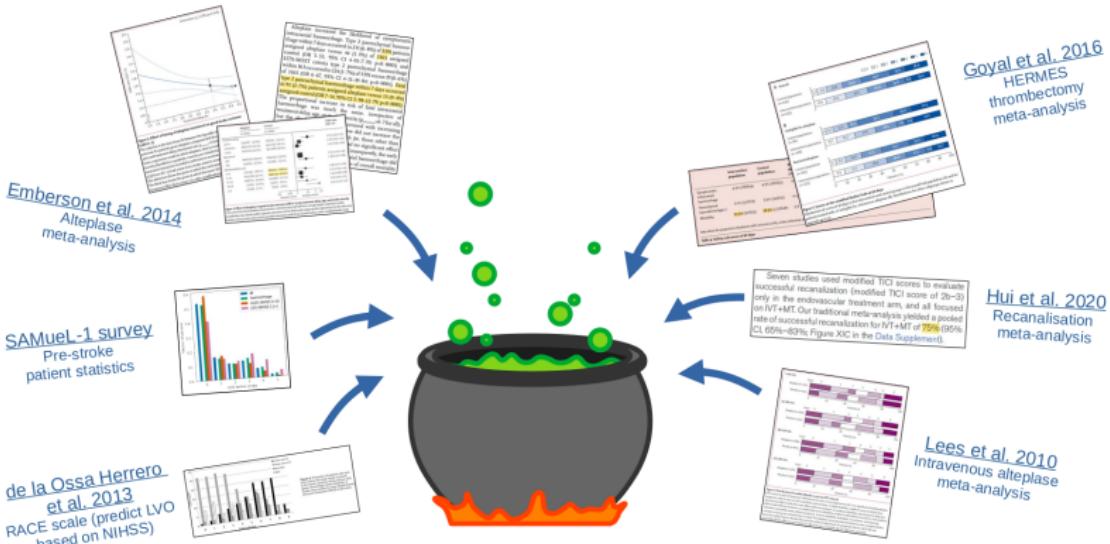
- *Shorter arrival-to-scan times*
- *Mid-level stroke severity*
- *Precise onset time*
- *Lower pre-stroke disability*

Lower thrombolysing units are particularly less likely to give thrombolysis to patients with:

- *Low or high stroke severity*
- *Higher pre-stroke disability*
- *Estimated onset time*

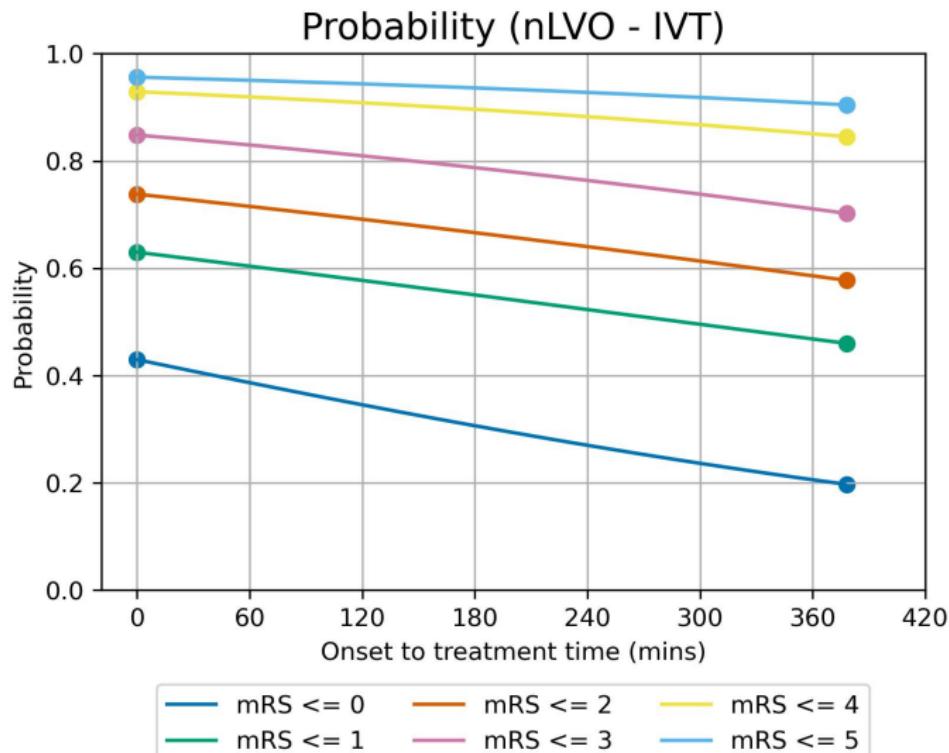
Stroke outcome modelling based on times to treatment with thrombolysis and thrombectomy

# Model combines multiple data sources

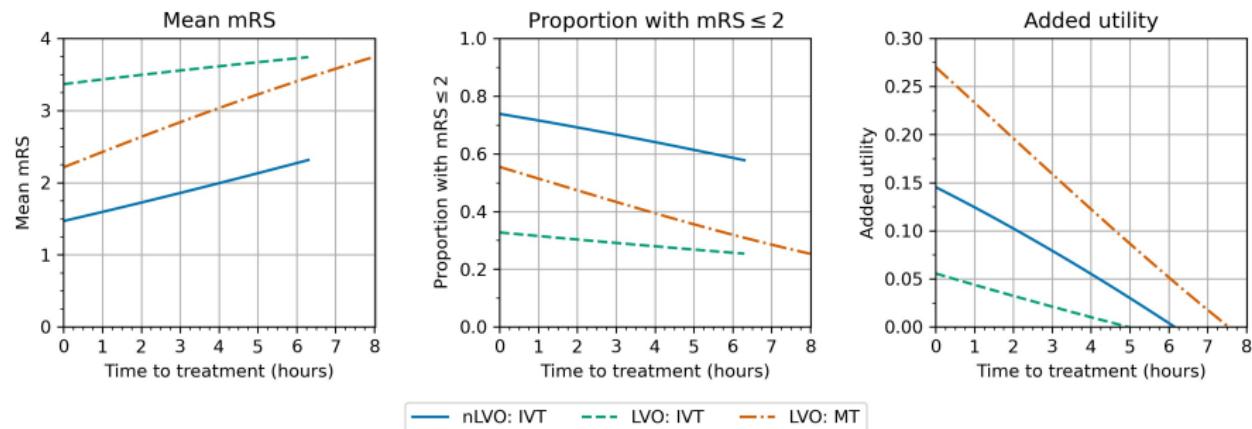


Sanity checks:	✓ <i>Test de la Ossa-Herrero</i> (Test de la Ossa-Herrero)	✓ <i>MR CLEAN</i> (MR CLEAN)	✓ <i>IST-3</i> (IST-3)
<b>Holodinsky et al. 2018</b> Pre-hospital transport model			
<b>Franseen et al. 2016</b> MR CLEAN clinical trial on recanalisation			
<b>IST-3 collaborative group et al. 2012</b> Thrombolysis meta-analysis			

# Prediction of mRS-level outcomes based on time to treatment



# Summary of benefit and added utility by stroke type and treatment type



HQIP data request

## HQIP data request

- Request for new data has been submitted to HQIP
- New data will de-anonymize hospitals, and include ambulance times
- Knowledge of hospital will allow linkage to SSNAP organisational audit data
- We had informal input from HQIP, and agreement that our data request can be considered as anonymised data, but there remains a small risk that HQIP will not consider data sufficiently anonymised.

Qualitative research

## Qualitative research key themes

*"Finding out what does it take to get AI stuff USEFULLY incorporated [adopted] into a national audit"*

Qualitative research has been reviewed, and we are now proposing three components:

- Observation case studies
- Interviews (individual and group)
- Model and AI development stakeholder meetings

## Observation case studies

- 2 weeks each at three hospitals
- Observation of emergency stroke care
- Interviews about SAMueL analysis for that hospital
- If possible, a (virtual) stakeholder workshop at each hospital, with modellers

## Interviews (individual and group)

- Broaden input to wider than three main case study hospitals

## Model and AI development stakeholder meetings

Engage stakeholders with SAMueL analysis, seek feedback, and refine models and presentation.

This may occur outside of formal qualitative research.

Stakeholders include:

- Individual hospitals (stroke teams)
- Integrated Stroke Delivery Networks
- NHS-E Communities of Practice
- *Key Informers*, e.g. you!, Sally Evans (NHS-E National Stroke Programme Clinical Policy Unit)

## Timelines

- Qualitative work was delayed by time taken to recruit, and by re-assessment of qualitative plans.
- HRA submission now due January 9<sup>th</sup>.
- Active research to start April.
- Work depends on approval by HRA REC *and* NIHR (approval of amendment to protocol). We believe risk of rejection by NIHR is low as we are providing some more in-depth research at no requested extension to project timelines and cost.

Further project materials

## Further project materials

- Key project documents:  
<https://github.com/samuel-book/samuel-2-reference>
- SHAP work:  
[https://samuel-book.github.io/samuel\\_shap\\_paper\\_1/](https://samuel-book.github.io/samuel_shap_paper_1/)
- Stroke outcome modelling:  
[https://samuel-book.github.io/stroke\\_outcome/](https://samuel-book.github.io/stroke_outcome/)



# Reserve slides

# When will low thrombolysing units not use thrombolysis when higher thrombolysing would?

Here, a high SHAP shows when a low-thrombolysing unit will reject use of thrombolysis when a higher thrombolysing hospital would use thrombolysis. (Plots are in order of feature importance.)

