17/99/89: Use of simulation and machine learning to identify key levers for maximising the disability benefit of intravenous thrombolysis in acute stroke pathways

1. Summary of Research

Background: Stroke is a leading cause of death and disability worldwide. Currently the only licensed drug treatment for acute stroke is thrombolysis with alteplase, the benefit of which is critically time-dependent. There is significant variation between hospitals both in rates of thrombolysis use and door-to-needle times for ischaemic stroke.

Aim: Our aim is to use simulation and machine learning technologies to identify key levers of improvement in thrombolysis use and speed, developing this analysis to be run as part of the routine quarterly national stroke audit. Qualitative research will be undertaken to maximise the acceptance and influence of these techniques.

Methods: Discrete event simulation allows for the prediction of the effect of changing key aspects of the acute stroke pathway (e.g. change in speed). Machine learning techniques allow for an understanding of differences in decision making between different hospitals, and offer the potential for 'exporting' decision making from one location to another. For example, training a machine-learning model based on decision making in a set of benchmark hospitals acknowledged to be centres of clinical excellence, allows an estimation of the effect of similar decision making in different hospitals which might have a different patient mix. All models will be built in Python, allowing easy transfer of techniques. Qualitative research will be based on 1:1 interviews, focus groups and workshops.

Pilot work: Methods have been trialled in seven hospitals in which thrombolysis use for stroke ranged from 7% to 14% of admitted patients. Three factors were pivotal in governing thrombolysis use: (1) the proportion of patients with a known stroke onset time, (2) pathway speed, and (3) predisposition to use thrombolysis for those patients canned with time to treat. A pathway simulation model could predict the potential benefit of improving individual stages of the clinical pathway speed, whereas a machine learning model could predict the benefit of 'exporting' clinical decision making from one hospital to another, whilst allowing for differences in patient population between hospitals. By applying both techniques together, we found a realistic ceiling of 15-25% in use of thrombolysis across different hospitals and more importantly, in the hospitals studied, a realistic opportunity to double the number of patients with no significant disability following treatment with thrombolysis.

Summary of planned work: Models will be refined and developed to run on the national stroke data set as part of routine quarterly audit. This will involve defining a reference group of benchmark hospitals for the decision-making (machine learning model). Pilot work has shown good promise for these techniques, but there is significant scope for optimising models, testing different types of machine learning models, and combining multiple machine learning models. We will pilot different ways of visualising output of models, and conduct qualitative research to understand how best to present model output to maximise their influence and impact in the clinical community.

2. Background and Rationale

Stroke is a leading cause of death and disability worldwide, with an estimated 5.9 million deaths and 33 million stroke survivors in 2010(1). In England, Wales and Northern Ireland 85,000 people are hospitalised with stroke each year(2), and stroke is ranked third as a cause of disability-adjusted life years in the UK over the last 25 years(3). Currently the only licensed drug treatment for acute stroke is thrombolysis with Alteplase, the benefit of which is critically time-dependent(4) with little or no benefit after 4.5 hours from stroke onset. Frustratingly, over the fifteen years since European licencing, the population benefit from thrombolysis has been limited by slow uptake of the treatment, and in-hospital delays to the administration of thrombolysis(5–7).

In England, Wales and Northern Ireland the national stroke audit 'SSNAP' (see below) records that 11.2% of acute stroke patients receive thrombolysis, but use in individual acutely admitting stroke centres varies from 0 to 24.5%(2). The lowest 10% of acutely admitting stroke teams administer thrombolysis to fewer than 5.9% of patients, whereas the top 10% administer thrombolysis to more than 16.7%. Time from arrival to thrombolysis ('door-to-needle') also varies significantly. The fastest 10% of hospitals have door-to-needle times of 40 minutes or less, whereas the slowest 10% have door-to-needle times of 85 minutes or more(2). There is therefore considerable variation between hospitals in the use, and speed, of thrombolysis for acute stroke patients, and the overall use of thrombolysis and the high inter-hospital variation has not changed in the last four years.

There have been many studies of barriers to the uptake of thrombolysis(8–10). Eissa et al.(8) divided barriers into pre-admission and post-admission phases. Pre-admission barriers included poor patient response (not recognising symptoms of a stroke and not calling for help soon enough) and paramedic-related barriers (such as adding delays in getting the patient to an appropriate hospital in the fastest possible time). Hospital-based barriers include organisational problems (delay in recognising stroke patients, delays in pathway, poor infrastructure) and physician uncertainty or lack of experience leading to low use of thrombolysis. There has been significant discussion on how services may best be organised to optimise the effectiveness of thrombolysis(11).

Analysis of patient pathway data coupled with computer simulation has previously allowed investigation and improvement of thrombolysis use in individual hospitals - increasing both the number of patients treated and reducing door-to-needle times(12,13). These models have usually focused solely on the speed of the acute stroke pathway from arrival at hospital to treatment with thrombolysis(12). Interest in the use of simulation for improving the performance of the acute stroke pathway has reached an interest such that a common framework has been proposed(14).

Pathway modelling based on simulating process steps allows for good simulation of the speed of the stroke pathway, but cannot easily model differences in clinical decision making. We were interested in testing whether a model could dissect out the variation in thrombolysis rate that is dependent upon differences in patient populations (e.g. age or stroke severity) in different hospitals, and from the differences that are dependent on the culture of decision making at different hospitals (e.g. more cautious vs more aggressive clinical decision making). A variety of machine learning techniques now exist(15), which are able to make good predictions on pre-existing multidimensional data over a binary or categorical outcome variable (such as whether a patient receives thrombolysis or not). These have the potential to add modelling of clinical decision making to a model of the acute stroke pathway, with the aim of predicting what decision (to thrombolyse or not) would be made for the same patient in different hospitals. Models may also be trained on a reference standard set of hospitals (regarded as centres of clinical excellence) and use of thrombolysis for any patient predicted using that 'benchmark clinical

decision making model'. Machine learning has three key advantages for our approach: 1) it may use a variety of techniques (ranging from more traditional statistical regression models through to state-of-the-art Deep Learning Neural Networks) which may be combined into one outcome using a technique known as 'ensemble learning', 2) Machine Learning is highly scalable, with framework developed for dealing with very large numbers of patients each of which might have very many 'features' recorded. Models may, for example, in time be scaled to also make use of any imaging data available, 3) Machine Learning models may continually learn from new data without having to re-fit all previous and recent data together.

SSNAP is the prospective national stroke audit of in-patient stroke care in England, Wales and Northern Ireland, funded by the Healthcare Quality Improvement Partnership (HQIP) and hosted by the Royal College of Physicians of London and King's College London. Since inception in 2013, SSNAP now has over 300,000 case records from 127 acutely admitting hospitals. The Sentinel Stroke National Audit Programme (SSNAP) provides an opportunity to train models using data from all acute stroke hospitals in Engalnd, Wales and Northern Ireland, extracting learnings at both the generic and local level, with the ability to feed back, through the established quarterly audit process, to all hospitals.

Evidence explaining why this research is needed now: As detailed above there is considerable and persistent variation in the use and speed of thrombolysis, a time-critical treatment for stroke, which limits the disability benefit to individuals and the population from this cost-effective treatment. Advances in national audit data collection, and advances in scalable computational methods for pathway simulation and machine learning make this a timely project to introduce these advanced analytical tools into SSNAP's quarterly national stroke audit reports.

3. Aims and objectives

Simulation and machine learning

Our aim is to extend previous work on stroke thrombolysis pathway simulation in three significant ways: 1) to create a generic stroke thrombolysis pathway simulation model that could be readily applied to any hospital, 2) extend the analysis to include factors other than door-to-needle times, with special focus on differences in clinical decision making as analysed and modelled with machine learning techniques, and 3) use a modelling framework that is open source and fast enough to run routine analysis on all UK hospitals. We will also structure the model to make it suitable for extension to include mechanical thrombectomy, an emerging treatment for the most severe form of ischaemic stroke.

The combined simulation and machine learning would then be used in the quarterly national stroke audit, estimating the potential use of thrombolysis and the associated clinical benefit, by improving pathway speeds and processes, and by applying clinical decision making similar to the benchmark centres of clinical excellence. It is hoped that by applying machine learning model in an audit setting, though valuable alone, may also potentially lead to 'expert' advisory systems that may support clinical decision making (especially by less experienced clinicians).

Simulation and machine learning has been piloted using data from seven regional hospitals in the South West of England (see section on pilot work).

Qualitative research

A critical question of applying this type of advanced computational techniques is 'will the feedback change clinical practice for the better?' Qualitative research will be conducted with an overall objective to

determine individual and consensus physician perspectives and concerns towards the use of simulation and machine learning in reviewing and improving clinical practice in stroke thrombolysis.

Qualitative methods will be used to:

- Explore current attitudes and rationale for the use of thrombolysis for ischaemic stroke, in order to establish reasons for the variance in the use and speed of thrombolysis.
- Elicit physician perspectives on simulation and machine learning feedback, to understand how our results are best presented in a way that is useful and likely to have an impact on their practice.
- Identify potential routes for the implementation of machine learning feedback, to inform and improve future stroke management.
- Explore and anticipate possible unintended consequences of stroke pathway changes.

4. Research Plan / Methods

4.1 Design and theoretical/conceptual framework

The research consists of three key components, all relating to the use of thrombolysis in the acute stroke pathway.

1) Stroke pathway simulation.

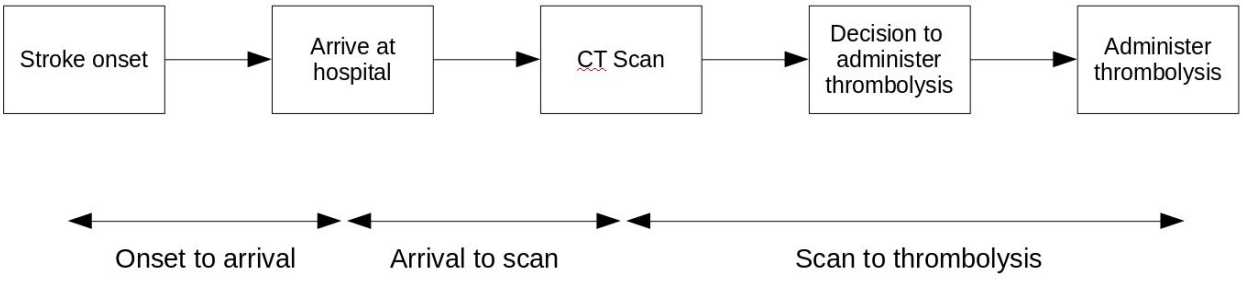
2) Patient level machine learning model on decision whether to administer thrombolysis (when there is time to administer thrombolysis).

3) Qualitative research exploring how physicians perceive the risks and benefits of using machine learning during audit to improve thrombolysis use for ischaemic stroke.

Pathway simulation model

(Project team members: Michael Allen, Kerry Pearn, Benjamin Bray, Martin James)

The pathway simulation model is coded in Python/NumPy and is based on sampling from distributions based on real-world data. Patients pass from one process stage to the next. The time spent in each stage is based on sampling from distributions derived from SSNAP data. Details of distributions used are given in Appendix 2.



For a patient to receive thrombolysis in the pathway model they must meet the following criteria: 1) stroke onset time known, 2) arrival at hospital within 4 hours of stroke onset, 3) have an ischaemic stroke and be judged to be eligible for thrombolysis, and 4) be within the licenced thrombolysis time window (4.5 hours and 3 hours onset-to-treatment time for patients aged under and over 80 respectively), when summing the process step times in the model. If a patient receives thrombolysis in the model then the probability of an additional good outcome (modified Rankin Scale [mRS] 0-1, no significant disability and able to carry out all usual activities) due to use of thrombolysis is calculated from the onset-to-treatment time and is based on the meta-analysis by Emberson *et al.(4)*.

If the pathway model is run without the machine monthlearning thrombolysis component (see next section), the likelihood of being given thrombolysis if scanned within 30 minutes is taken either from the hospital's own data on the proportion of patients who have time to receive thrombolysis (30 minutes licence window remaining after scan) and are given thrombolysis, or by using a published reference proportion (e.g. in the IST-3 trial 50% of stroke patients, if scanned with time left to treat, went on to receive thrombolysis(16)).

Model outputs: The primary outputs of the model are 1) an estimate of the proportion of patients receiving thrombolysis, and 2) expected clinical benefit achieved through use of thrombolysis (additional disability-free patients).

Model Validation: In order to validate the model, 3 years data (~225,000 patient records) will be split into two sets: Model parameters will be set using 75% of the data, and accuracy of the model compared with 25% test data not used for model training.

Scenario analysis: The model will be run for each hospital with key changes in the pathway, e.g. using upper quartile SSNAP data for determining stroke onset time, arrival-to-scan time, scan-to-treatment time, and the proportion of patients (with time left to treat) with decision to thrombolyse.

Pilot work: Please see section on pilot work for an example of use of the simulation model.

Clinical decision model (machine learning)

(Project team members: Richard Everson, Michael Allen, Kerry Pearn, Zhivko ZHelevl, Benjamin Bray, Martin James)

The clinical decision model aims to replicate the decision whether to give thrombolysis for any given patient at any given hospital. If patient features (characteristics) are kept unchanged, but the admission hospital is changed, the model should predict different expected decisions for the same patients in different hospitals. Differences in decisions between hospitals may be compared by passing the same randomly selected sample of patients to all hospitals (as differences between observed thrombolysis use between different hospitals may be complicated by differences in patient characteristics between different hospitals).

The model may be trained using different benchmarks. For example the model may be trained using a subset of hospitals recognised for their clinical excellence in acute stroke care. Possible subsets include hospitals in London which were found, as a group, to have improved outcomes following stroke care reorganisation/centralisation(17), or subgroups identified within the national audit with other organisational characteristics of excellence e.g. high rates of direct admission to an acute stroke unit, high rates of early dysphagia screening, etc.).

This model is intended to make decisions based only on clinical presentation, assuming that there is sufficient time remaining in which to assess and give thrombolysis. Patients are included if they have

been scanned with 30 minutes remaining in the licence window to give thrombolysis. The model is coded in Python using available SciKitLearn, Tensorflow, and PyTorch machine learning libraries.

The model predicts whether an individual patient should receive thrombolysis or not from a set of 50 parameters defining the patient's characteristics, clinical well-being, and hospital attended (see section on pilot work for a list of features used in the pilot work) that would all be available to the stroke clinician at the time of their thrombolysis use decision-making. The models are supervised learning models based on a training set of data with known use of thrombolysis. The machine learning models to be used are:

- Random Forests
- Support Vector Machines (SVM): linear and rbf
- Neural network (including basic freed-forward neural networks, and more advanced PyTorch and/ or Tensoflow neural networks).
- Logistic regression
- K nearest neighbours
- Gaussian process models
- Novel decision-tree methods (based on application of current work ongoing at the University of Exeter).

In addition to single type machine learning models, ensemble models will be built. These take the output from multiple different machine learning models and have been shown to be able to produce better accuracy than any single method alone(18).

Model outputs: The model outputs for each patient record entered whether thrombolysis would be given (along with a measure of probability of the decision).

Model Validation: In order to optimise and validate the model, data 3 years data (~225,000 patient records) will be split into three sets (randomly selected from stratified data to ensure all sets have similar over thrombolysis use). One set (25%) of the data will be held back for final testing after model selection and optimisation. The remaining 75% will be used for training and testing for model selection and optimisation using stratified k-fold validation (where the group is iteratively split, e.g. into ten 90% training and 10% test sets, so that all samples are part of a test set once and only once). The primary measure of performance will be ROC area under curve, with other outcomes reported (sensitivity, specificity, overall accuracy, F1 measurements).

Model scenarios: The aim of model scenarios is to ask 'what if?' questions of the model, to test the expected effect of hypothetical changes to the system.

A reference set of patients may be passed to each of the models that are trained on decisions made for patients from a single hospital, and thrombolysis use can be predicted for each hospital on this common reference set of patients. This will help identify hospitals that appear to have unusual thrombolysis use decisions (either significantly higher or lower than average, or significantly different to a benchmark group) independent of differences in patient populations attending each unit.

Where clinical decision making appears to significantly differ from benchmark hospitals, a cohort of patients from a single hospital (non-elite) may be passed to a model trained on patients from a benchmark set of hospitals (elite). That way, the use of thrombolysis within the non-elite centre (the recorded thrombolysis use) can be compared with the potential decision-making as would have

happened for that patient if they attended an elite centre (modelled thrombolysis use). To aid audit, a small group of patients will be identified for each hospital where modelled (based on the model trained using the benchmark set of hospitals) and actual thrombolysis use differ.

Our pilot studies give a strong indication of the robustness of the machine learning methods. Nevertheless if these initial results are not borne out over particular time epochs or at particular hospitals we will investigate the reasons, for example, if there is a difference in practice between pilot hospitals. This will give insight into the practice of the stroke treatment pathway and we will will construct alternative models to model these data. As an example, if accuracy is lower than anticipated, the results of individual machine learning models may be compared. If models differ from each other then the most likely issue is that the models are each too weak, in a random fashion, to attain high accuracy. In this case increasing the number of models used, and combining results (an 'ensemble of weak learners') offers a popular approach to improving accuracy. If accuracy is low but different models agree on the decision then this points to a systematic difference between model and data (for example one clinician in any one centre always making a different type of decision than other clinicians in that centre), and further investigation should focus on what additional data should be collected to improve accuracy (e.g. collecting data at clinician level in this hypothetical example).

The machine learning model may be used independently, but may also be combined with the pathway simulation mode. The combined model allows for investigation of the potential benefits of improving pathway processes and applying clinical decision making aligned with centres of clinical expertise. The combination may therefore provide a useful and realistic target use of thrombolysis given a hospitals own patient group characteristics.

Pilot work: We have performed extensive pilot work for project (for both simulation and machine learning aspects), and have included a separate section below on that work.

Qualitative research

(Project team members: Julia Frost, Ken Stein, Kristin Liabo)

Focus groups and semi-structured interviews will be conducted by an experienced qualitative researcher(20). We will recruit physicians via local, regional and national networks (clinical networks and the national specialist society, the British Association of Stroke Phycisians) and use maximum variance sampling to ensure inclusion of a range of relevant models of delivery and physician experiences such as: existing models of delivery, centres of excellence, type of hospital (regional centre or district hospital), any physician specialism (generalist, emergency care or stroke physician), and years in practice.

Three focus groups will be conducted in regional stroke centres in order to determine different clinician approaches and attitudes to the management of ischaemic stroke and thrombolysis practice(21). These groups will enable us to identify and pilot a range of visual displays and other methods of feedback from both pathway modelling and machine learning for use in individual interviews. We anticipate that these will involve: 1) national data, 2) regional data and 3) individual case data (or patient ‘vignettes’), to enable us to elicit perspectives and views about how best to present the feedback derived from machine learning.

30-40 interviews (face-to-face or telephone) will be undertaken with participant physicians, both career and training grades. Interviews will use a topic guide, based upon both the literature and expert opinion, and will ascertain which factors inform and influence clinical decision making and beliefs about the appropriate use of thrombolysis, or not. The second part of the interview will involve the introduction of

the data displays identified from the focus groups (and which can be sent ahead of any telephone interviews), and will identify which forms of data visualisation can best inform clinical practice. Patient and Public Involvement (PPI) representatives have enthusiastically endorsed the use of machine learning and qualitative methods and, reflecting earlier research(22,23) have suggested that our study materials must clearly emphasise why thrombolysis may, or may not be, of benefit to patients. All interviews will be audio recorded, transcribed verbatim and anonymised.

4.2 Sampling

SSNAP has near-complete coverage of all acute stroke admissions in the UK (outside Scotland). All hospitals admitting acute stroke participate in the audit, and year-on-year comparison with Hospital Episode Statistics confirms estimated case ascertainment of 95% of coded cases of acute stroke.

4.3 Setting/context

The model will use data from all English units registered as acutely admitting stroke units in SSNAP. Qualitative interviews will be held with stroke clinicians from acute stroke units in England.

4.4 Data collection and strorage

All patient-related data comes from SSNAP and is collected as part of routine care. We will access data through a single source managed by HQIP. Anonymised data will be handled in accordance with University of Exeter data protection policies. Qualitative research will be conducted in accordance with the General Data Protection Regulation and all relevant University Policies. See section 9 for details on ethics approvals for SSNAP and qualitative work.

4.5 Data analysis

Pathway simulation and machine learning

See Pilot Work and planned extensions for details of data analysis for ptahway modelling and machine learning.

Qualitative research

A thematic analysis of interviews will be conducted(24). The analysis will be iterative, moving between data collection and analysis to test emerging theories. This work will build upon the already identified barriers and facilitators to the use of thrombolysis, and will focus on the implementation of feedback from machine learning to optimise thrombolysis for ischaemic stroke management. We anticipate that this might involve the identification of examples of best practice that could inform the development of a future intervention to support quality improvement activities all along the pathway(25). Another possible output could be the development of a typology concerning the type of visualisation that might work best in a given scenario, e.g. whether they are organisational, clinician or patient factors.

A workshop will be conducted three months prior to project end, and up to 30 stakeholders will be invited to participate, which will include people who have had a stroke, carers, health professionals (physicians and members of the wider stroke team), NHS managers. The workshop will follow a structured format where participants engage in focused discussions, interspersed with brief presentations by the research team. We will present early findings from the various aspects of the research, including both the

simulation and machine learning, and the focus groups and interviews. Our PPI collaborators will also inform workshop participants about their contribution to the overall research project. Contingent on the nature of preliminary findings, activities will initially be in separate groups of professionals, patients, and others, followed by mixed groups of people from different backgrounds. This process and the multiple perspectives will be recorded in several ways, 1) by note-takers within sessions (JF, KL), 2) the researchers' participant observation notes made immediately after the stakeholder workshop and 3) participants' flipchart notes and summaries made during certain sessions. With consent, discussions will be audio-recorded, although we do not anticipate transcribing the whole event, rather recording will provide an aide memoire of the breadth of discussions as opposed to the attribution of data at an individual level. Thematic analysis of this data will augment and triangulate preliminary findings, to inform the development of an intervention to support quality improvement in thrombolysis practice. Previous experience with this method suggests that it will contribute concerns and issues that might otherwise be by those conducting the research and also offer additional interpretations and suggestions for implementation(26,27).

Following feedback from the review board we aim to use the qualitative work to also explore views on what unintended consequences may come about from changing the acute stroke pathway, and how might such adverse effects be detected and mitigated. We will produce a summary of these points in the form of a Failure Mode Effects Analysis (FMEA) as used in engineering when trying to anticipate what may go wrong with a product.

5. Pilot work

Pathway simulation and machine learning methods, has been piloted using SSNAP data from seven acute stroke units.

5.1 Data

SSNAP data was obtained from 7 hospitals in England, for patients with a confirmed stroke over a period of two years (2013-2014) for each hospital. These data were anonymised secondary data, collected during routine care. No patient identifiable information was obtained. For the pathway simulation model, the dataset contained 7,864 patient records with complete data for 12 parameters regarding their characteristics and time stamped pathway location. These data represent out-of-hospital onset of stroke (which account for 94% of all cases of acute stroke in the SSNAP data used). For machine learning, only those patients with a completed National Institutes of Health Stroke Scale (NIHSS) and who had at least 30 minutes left to give thrombolysis were used (1,862 patients).

5.2 Pathway simulation

Hospital	1	2	3	4	5	6	7
Actual thrombolysis use	13.7%	8.0%	8.5%	7.2%	7.7%	14.3%	9.5%
Aged 80+	50.2%	51.3%	41.3%	49.8%	45.8%	51.1%	48.9%
Onset time known[1]	67.8%	53.7%	53.3%	43.4%	44.6%	72.7%	56.2%
Known arrival within 4 hours of onset[2]	68.4%	69.8%	69.3%	67.7%	69.7%	70.5%	70.3%
Average arrival time after onset (min)[3]	96	87	95	100	82	93	87
% Patients scanned within 4 hours[3]	84.4%	78.3%	94.7%	97.9%	83.7%	84.9%	91.3%
Average arrival to scan (min)[4]	38.3	30.6	43.6	11.9	40.5	28.6	22.1
Proportion ischaemic stroke[5]	86.4%	88.5%	85.2%	85.8%	89.2%	85.7%	88.2%
Proportion ischaemic patients receiving thrombolysis[5]	55.7%	38.6%	39.2%	34.4%	44.5%	49.5%	34.7%
Average scan to treatment (min)	22	36	44	40	35	35	36

[1]The known onset time may be recorded as precise or best estimate.
[2]of those with known time of onset
[3]for those arriving within 4 hours of known onset
[4]for those arriving within 4 hours of known onset and having a scan within 4 hours of arrival
[5]for those scanned with 30 minutes left to administer thrombolysis

Table 1. Key characteristics for the seven hospital pathways modelled. The number of good outcomes (modified Rankin Scale 0-1, no significant disability) without any use of thrombolysis was estimated at being 238-260 per 1,000 stroke patients admitted depending on the hospital (the differences being due to differences in age profile of patient populations).

Table 1 shows the key characteristics for the seven hospital pathways modelled. The model was validated by comparing modelled (predicted) use of thrombolysis with actual use. Predicted use was based on modelling a one year period, with replicates of 100 runs (each with different random number seeds to sample a different value from each distribution for each of the 100 runs with different random seeds) in order to determine expected year-to-year variation. The model showed excellent agreement (R-square 0.981) between actual and predicted values though the model slightly under-predicted actual thrombolysis use by an average of 0.84% (figure 2).

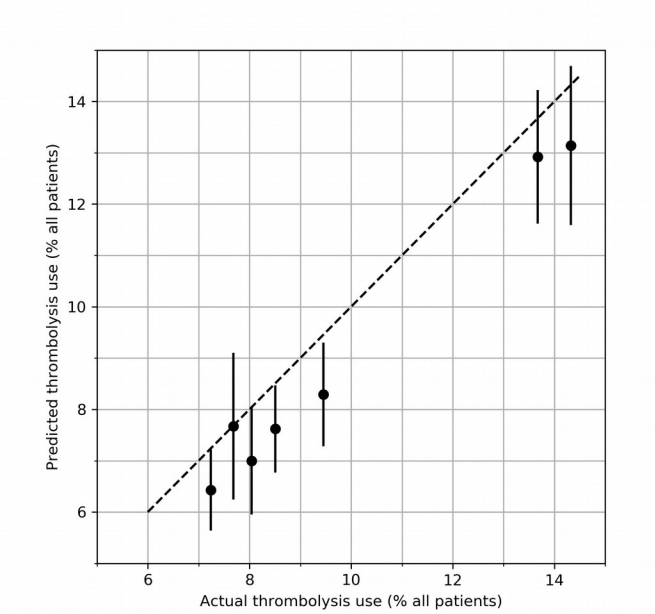


Figure 2. Validation of the pathway simulation model, comparing actual to modelled (predicted) thrombolysis use at seven hospital. Points show mean predicted thrombolysis use for all confirmed stroke patients arriving at hospital, with bars showing 5th and 95th percentiles from 100 runs, with each run modelling one year.

The model was run with various 'what-if?' scenarios for each of the seven hospitals (Table 2):

1. Base case: Model based on parameters derived from current hospital-specific performance
2. Scenario A: Door-to-needle time fixed at 30 minutes (by fixing arrival-to-scan time and scan-to-thrombolysis both at 15 minutes (with no variation in either time)
3. Scenario B: Judged to be eligible for thrombolysis fixed at 60% of ischaemic strokes. An analysis of ECASS-3/IST-3 results concluded that 591 out of 992 (59.6%) of ischaemic stroke patients arriving within 4 hours of stroke onset were ultimately considered suitable for thrombolysis(16).
4. Scenario C: Onset time known fixed at 77% (national SSNAP upper quartile for year 2015/16(2))
5. Combination of all of the above

In order to achieve the greatest improvement in thrombolysis use in each of the seven hospitals, for three hospitals (hospitals 1, 2 and 6) it would be best to improve the speed of the pathway, for two hospitals (hospitals 4 and 5) it would be best to improve determination of stroke onset time, and for two (hospitals 3 and 7) it would be best to judge more patients as eligible for thrombolysis for those scanned with time left to treat. If a priority is to maximise clinical outcome then for five hospitals (hospitals 1, 2, 3, 5 and 6) it would be best to improve the speed of the pathway, for one hospital (hospital 4) it would be best to improve the determination of stroke onset time and for one (hospital 7) it would be best to judge more patients as eligible for thrombolysis for those scanned with time left to treat. Combining all changes in the model could produce thrombolysis rates up to 22-26%, and 23-26 additional non-disabled outcomes per 1,000 stroke patients admitted.

	Hospital	Base	A	B	C	ABC
Thrombolysis use (%)	1	12.7 (0.2)	17.4 (0.3)	13.7 (0.3)	14.6 (0.3)	21.1 (0.3)
	2	6.9 (0.2)	10.3 (0.3)	11.0 (0.3)	9.9 (0.2)	23.0 (0.3)
	3	7.6 (0.2)	10.1 (0.2)	12.2 (0.2)	11.1 (0.2)	22.2 (0.3)
	4	6.4 (0.2)	6.5 (0.2)	11.1 (0.3)	11.5 (0.2)	20.6 (0.3)
	5	7.4 (0.3)	10.2 (0.3)	10.2 (0.3)	13.1 (0.4)	23.8 (0.5)
	6	13.5 (0.3)	17.2 (0.4)	16.4 (0.4)	14.0 (0.4)	22.4 (0.4)
	7	8.2 (0.2)	9.6 (0.3)	14.3 (0.2)	11.2 (0.2)	22.8 (0.3)
Additional good outcomes per 1,000 admissions	1	11.1 (0.2)	16.9 (0.3)	12.0 (0.2)	12.7 (0.3)	20.6 (0.3)
	2	5.9 (0.2)	10.3 (0.3)	9.5 (0.3)	8.6 (0.2)	22.8 (0.4)
	3	6.3 (0.2)	10.3 (0.2)	10.0 (0.2)	9.1 (0.2)	22.6 (0.3)
	4	5.5 (0.2)	6.3 (0.2)	9.6 (0.2)	10.0(0.2)	19.9 (0.3)
	5	6.5 (0.3)	10.5 (0.4)	8.9 (0.3)	11.4 (0.4)	24.4 (0.5)
	6	11.5 (0.3)	16.8 (0.4)	14.1 (0.4)	11.8 (0.3)	21.8 (0.4)
	7	7.3 (0.2)	9.7 (0.3)	12.8 (0.2)	10.0 (0.2)	23.0 (0.3)

Table 2. Predicted thrombolysis use and clinical benefit across all modelled hospitals (1 to 7). Data shows: (Base) Model based on parameters derived from current performance; (A) Arrival-to-scan and scan-to-thrombolysis both fixed at 15 min (with no variation in either time); (B) Judged to be eligible for thrombolysis fixed at 60%; (C) Onset time known fixed at 77%; Combinations of the above. Results show mean and ±95% confidence limits (1,000 runs).

In the case of hospital 5, arrival-to-scan times could be slowed by an average of 30 minutes and clinical outcomes would still be greater if that hospital achieved a proportion of known stroke onset time equal to the national average.

5.3 Clinical decision (machine learning) model

Machine learning models were trained and tested on the subset of patients who have 30 minutes left to treat after scanning. Patient features used in the model are given in Appendix 3. Across the seven hospitals an average of 40% of these patients actually received thrombolysis, though use ranged from 31% to 52% in different hospitals.

Table 3 shows the performance of the various machine learning models. The models had 80-82% accuracy, and 88-89% ROC area under the curve. The Random Forest machine learning model was chosen to use from this point onwards.

	accuracy	sensitivity	specificity	ROC area
Logistic regression	0.807 (0.010)	0.788 (0.015)	0.819 (0.012)	0.882 (0.005)
Random Forest	0.819 (0.010)	0.792 (0.017)	0.837 (0.011)	0.889 (0.007)
SVM (linear)	0.807 (0.008)	0.820 (0.012)	0.798 (0.012)	0.880 (0.005)
SVM (rbf)	0.813 (0.006)	0.820 (0.013)	0.808 (0.008)	0.880 (0.006)
Neural Network	0.807 (0.010)	0.794 (0.017)	0.816 (0.012)	0.882 (0.005)

Table 3. *Accuracy, sensitivity, specificity, and ROC (Receiver-Operator Curve) area of five different machine learning models. Results show mean and standard error based on a 10-fold stratified validation/test split.*

A machine learning model may be trained on a subset of patients to investigate how the difference in thrombolysis use between hospitals may be proportionally attributed to either the hospital or the local patient population. Table 4 shows the predicted use of thrombolysis in a set of patients that attend one hospital, based on decisions made from training at another hospital. Taking hospital 7 as an example, between 31% to 45% of the patients that currently attend hospital 7 (with time left to receive

thrombolysis) might receive thrombolysis depending on which hospital decision making is used to train the model. Patient cohort also affects the predicted thrombolysis uses. Taking hospital 7 as an example again, if the model is trained on decisions made for patients attending hospital 7, and different hospital patient groups are then analysed in the model then thrombolysis use is to be between 23% and 50% depending on the hospital patient group analysed.

Actual thrombolysis use by hospital

1	2	3	4	5	6	7	All
52%	35%	48%	33%	49%	44%	31%	40%

Predicted thrombolysis use

		Hospital patients actually attended							
		1	2	3	4	5	6	7	All
Hospital used to train model	1	52%	42%	58%	50%	67%	57%	45%	51%
	2	48%	35%	55%	36%	46%	37%	29%	40%
	3	53%	38%	48%	46%	58%	41%	34%	44%
	4	40%	28%	48%	33%	52%	29%	26%	35%
	5	50%	36%	50%	40%	49%	45%	37%	43%
	6	49%	32%	55%	44%	59%	44%	39%	44%
	7	42%	23%	42%	31%	50%	36%	31%	35%

Table 4. Predicted thrombolysis use (for patients with patients scanned with time left to receive thrombolysis) if the decision to give thrombolysis is based on decisions made by a Random Forest model trained at different hospitals. The columns therefore represent the likely difference in thrombolysis use due to differences in decision making.

5.4 Combining pathway simulation and machine learning

The output from machine learning may be incorporated into the stroke pathway model by using the machine learning model to make the decision in the pathway model about whether a patient is 'judged to be eligible for thrombolysis (for patients scanned with 30 minutes left to administer thrombolysis)'. This should tailor the clinical decision to the local population, without being affected by any particular hospital's predisposition to use thrombolysis. The 'judged to be eligible for thrombolysis' parameter in the pathway model may take its value from a machine learning model trained using a reference set of hospitals. The clinical decision making from these reference hospitals may be used to predict which of the patients from the hospital under study are eligible for thrombolysis.

Table 5 compares base case hospital performance (predicted thrombolysis use and clinical benefit) with the performance obtainable by a new realistic 'alternative' practice which is in part informed by the Random Forest machine learning model: (1) the proportion of patients with a known stroke onset time is set at the national median (67%) unless a hospital is already higher, (2) the door-to-needle time is set to 40 minutes for 90% of patients (20 minutes arrival-to-scan, and 20 minutes scan-to-needle) with the other 10% of patients not receiving a scan within 4 hours of arrival, and (3) the clinical decision to administer thrombolysis for those patients scanned with 30 minutes left to treat is set by the machine learning model trained from a reference hospital (this example uses the hospital that has the highest use of thrombolysis for those patients scanned with time to treat). Resulting thrombolysis targets vary from 16-25% depending on the hospital (base case 6 to 13%).

	Thrombolysis use (%)		Clinical benefit: Additional good outcomes per 1,000 admissions	
Hospital	Current	Alternative	Current	Alternative
1	13.0 (0.3)	18.8 (0.3)	11.4 (0.3)	17.6 (0.3)
2	6.8 (0.2)	15.6 (0.3)	5.9 (0.2)	14.9 (0.3)
3	7.9 (0.2)	21.7 (0.3)	6.5 (0.1)	21.2 (0.3)
4	6.4 (0.2)	17.1 (0.3)	5.5 (0.2)	16.0 (0.3)
5	7.6 (0.3)	25.8 (0.5)	6.7 (0.3)	25.7 (0.5)
6	13.1 (0.3)	23.2 (0.4)	11.3 (0.3)	21.8 (0.4)
7	8.2 (0.2)	16.8 (0.3)	7.2 (0.2)	16.3 (0.3)

Table 5. Combining pathway simulation and machine learning. Predicted thrombolysis use and clinical benefit (additional good outcomes per 1,000 admitted patients) across all modelled hospitals (1 to 7) from the pathway simulation. Data shows: (Base) Model based on parameters derived from current performance; (Alternative) New realistic practice, fixing the proportion of known stroke onset times to the national SSNAP average (67% median) unless the hospital currently performs higher, fixing arrival-to-scan and scan-to-needle to 20 minutes each (with 10% of patients not scanned within 4 hours), and fixing the proportion of treatable patients (scanned with 30 minutes left to treat) according to the output of the machine learning model based on the hospital with the maximum predicted proportion given thrombolysis. Data shows mean and 95% confidence intervals.

5.5 Planned extension to pilot work

The pilot pathway simulation and machine learning work will be extended in the following ways:

- Use data for all acute stroke units in England
- Add analysis of time epochs to pathway simulation model (e.g. day/night/weekday/weekend) to test targeting of potential pathway improvements by time epoch.
- Optimise performance of current machine learning models (e.g. by refining selection of data used, testing of polynomial functions in inputs, testing of use of principal component analysis, optimising model meta-parameters).
- Apply additional machine learning techniques (e.g. k-nearest neighbours, Gaussian processes, novel Decision Tree methods, ‘deep-learning’ neural networks from PyTorch and Tensorflow).
- Apply ensemble techniques to combine multiple machine learning techniques to produce a single outcome.
- Incorporate confidence of decision into output.
- Produce machine learning code that will either work on all SSNAP data or on smaller-batch data to allow continual learning/update of the model with each run without having to process the whole data set each time.
- Structure simulation and machine learning models in such a way that they may extended in the future to include mechanical thrombectomy.
- Identify one or more sets of hospitals to use as benchmark hospitals for training the model. We are currently planning to use subgroups of hospitals identified by a range of other markers of clinical quality (e.g. .those with the highest rates of direct admission to an acute stroke unit within 4 hours), but we will also explore subgroups of hospitals of similar size and patient demographics to allow comparisons with a ‘similar 10’.

- Identify a reference group of patients (either through sequential or random sampling). These will be used as a final test and validation of the model
- Identify a second reference group of patients through random sampling (with or without stratification) to act as a representative sample that may be used to estimate the differences in use of thrombolysis between different hospitals given the same population group.
- Identify a subset (30 patients) of the representative patient group to give to small groups of stroke clinicians and ask them (independently) to make a decision whether they would or not give thrombolysis to these patients from the SSNAP data provided. We will compare cross-physician agreement on treatment decisions. We will also ask what additional information would have given greater confidence in the decision made.
- Produce a variety of visualisations of outputs from the pathway simulation and machine learning models for use in the qualitative research outlined above. Visualisations will be at 1) a national level, and 2) individual trust level.
- Run at least two pilots applying code as part of the national stroke audit (the first pilot will be used to test and refine the code for 'production' use, and the second pilot will generate output that will be given to all stroke units).
- Following feedback from a PPI meeting we also plan to perform some exploratory work at using machine learning method to predict outcomes (both benefit and risk of haemorrhage) at patient-level (the main model described performs clinical benefit analysis at a net population level). Decisions, if the tool were applied in a live decision-aid setting, could allow adjustment for patients acceptability of risk. We anticipate this work laying ground for further separately-funded work.

6. Dissemination, Outputs and anticipated Impact

6.1 What do you intend to produce from your research?

We have four production aims:

1) To produce code that will be routinely used by the national stroke audit SSNAP as part of their quarterly outputs to participating hospitals and commissioning groups, and in national reports. The code will have a structure for potential future extension to mechanical thrombectomy (an emerging alternative to thrombolysis for the most severe ischaemic strokes).

2) To publish the code in a public Open Source code repository (e.g. GutHub or GitLab).

3) To produce papers (in addition to the NIHR monologue)on at least:

- Application of machine learning to audit of thrombolysis use (technical machine learning publication)
- Application of combined simulation and machine learning to stroke thrombolysis audit (suitable for general/clinical readership)
- Qualitative publication of facilitators and barriers to use of computer simulation and machine learning in national audits.

4) Presentation of above paper themes at the UK stroke forum and an international stroke conference.

6.2 How will you inform and engage patients, NHS and the wider population about your work?

Working through SSNAP gives us a means of engaging with all stroke units as the models are implemented. We will also engage with clinicians through attendance and presentation at the national stroke forum. We also have support from the Stroke Association (see letter of support attached in appendices) to plan and implement wider engagement. With PPI involvement in the project we will will produce dissemination material suitable for public and patients.

6.3 How will your outputs enter our health and care system or society as a whole?

Although our principal aims relate primarily to increasing the effectiveness and impact of the national stroke audit SSNAP in increasing the uptake of treatments for acute stroke and reducing disability, we consider our work to have potential applications and transferability into other clinical areas, particularly those relating to other time-sensitive treatments and where considerable clinical variation persists. Our qualitative outputs may similarly transfer into the quality improvement field in clarifying methods to improve the reach and impact of comparative data in other national audits e.g. the Renal Registry, the National Diabetes In-patient Audit etc.

6.4 What further funding or support will be required if this research is successful (e.g. from NIHR, other Government departments, charity or industry)?

None to apply what we have done, as we will make our code able to run routinely. There are opportunities for further development. For example if machine learning proves accurate in predicting practice audit, a next step could be to format the models into a form that they could be used for clinical guidance in routine care (including the potential to predict outcome likelihoods given individual patient characteristics). We would also expect the model in future to be developed to include use of thrombectomy in addition to thrombolysis.

6.5 What are the possible barriers for further research, development, adoption and implementation?

We recognise that there is mistrust of methods of artificial intelligence as applied to complex clinical situations, and the qualitative component of our proposal seeks to directly address this issue. Adoption and implementation will be hampered without a thorough understanding of these barriers from our research, which may identify other, as yet unforseen, obstacles to wider acceptance. We are helped in this by the established credibility of SSNAP as a national comparative audit, and our proposal to pilot our dissemination methods leaves plenty of scope to explore these issues further, in order to mitigate them.

6.6 What do you think the impact of your research will be and for whom?

This work is intended to benefit stroke patients through improved stroke pathways and improved decision making regarding the use of thrombolysis. We believe there is a wider benefit of helping to establish simulation and machine learning in healthcare audit and practice.

7. Project / research timetable

Pre-work: Data access application to HQIP & access data from SSNAP. Ethics application for qualitative work through our local research ethics committee (REC) and via the Integrated Research Application System (IRAS).

Months 0-9: Primary focus is on development of simulation and machine learning models. Add epochs (day/night/weekday/weekend) to pathway model. Optimise current machine learning models for on national dataset. Apply additional machine learning methods and develop ensemble model. Identify a subset of hospitals to be used as a benchmark for thrombolysis decisions. Identify reference group of representative patients to use to compare decisions in different hospitals. Identify a smaller subset of patients to ask three clinicians to judge whether they would likely give thrombolysis or not. Produce a variety of visualisations of outputs from the pathway simulation and machine learning models (this will not require models to be finalised – development of models should refine accuracy rather than change the type of output produced). Preliminary visualisations will be produced in the first four months based on pilot regional work. This will allow an early start to qualitative interviews (which depend on example analysis); national level visualisations should begin to be available within the first nine months of the project.

Months 10-18: Qualitative phase of project to conduct 30-40 interviews with stroke clinicians on outputs of modelling. Continue to refine and expand machine learning models. Visualisations will be refined during this phase following feedback from clinicians. Provide code to SSNAP at 12 and 18 months (code will work on a standard CSV export from SSNAP database).

Months 19-24: Qualitative stakeholder workshop, writing of papers, refining code and making it of publication quality (e.g. well commented for other people to follow and amend).

(Patient involvement continues through project through membership of steering and project groups.

(PPI continues through project through membership of steering and project groups).

MILESTONES:

1) **End month 4:** Visualisations and summary of output from pilot regional work will be provided for beginning qualitative interviews. Qualitative work to being after these visualisations produced.

2) **End month 6**: Qualitative work to have begun (initially based on outputs from regional pilot work).

3) **End Month 9**: First outputs from national model will be generated with prototype visualisation of results that may be used for qualitative work. This will not be the final model, but should be advanced enough to form the basis of results that can be shared with clinicians, and a PPI group, through the qualitative work of the project.

4) **End of months 12 and 18**: provide code to SSNAP that will perform analysis and produce reports/visualations. This code will run on CSV file formats from SSNAP (providing an easy method to test and use code as a stand-alone module).

5) **End of month 18**: 1:1 qualitative interviews and focus groups complete.

6) **End of month 21**: Qualitative stakeholder workshop complete.

7) **End of month 24**: Papers to be complete (see dissemination), model code to be published (GitHub).

8. Project management

There will be an independent advisory board which will meet prior to project start, six months into project, and four months from the end. The membership of this advisory board is:

- Dr Thomas Monks (Operations Research, Southampton, Chair)
- Prof Anthony Rudd (National Clinical Director for Stroke)
- Prof Gary Ford (CEO Oxford AHSN, Stroke Physician, Prof Clinical Pharmacology, Oxford)
- Prof Nicky Britten (Professor of Applied Health Care Research, Exeter)

There will be a steering committee composed of Prof Ken Stein, Dr Martin James, Dr Benjamin Bray, Prof Richard Everson, two lay PPI members (separate from the project team) and one independent senior stroke physician/academic (to be approached). The steering committee will ensure that the project delivers on the stated aims and is kept in close alignment and contact with the National Stroke Audit team. The steering committee will meet every six months.

The project team will meet quarterly, and will consist of all project team members, including two patients and members. One carer member is a named collaborator to this application (PT).

The project team will meet monthly, and will consist of all project team members, including two patients and members. One carer member is a named collaborator to this application (PT).

9. Ethics / Regulatory Approvals

9.1 Access to SSNAP data for simulation and machine learning

All patient data will be from a single source: The Sentinel Stroke Audit Programme (SSNAP). No identifiable patient information will be requested or used. Explicit consent for the use of patient identifiable information is not required (although patients can choose to ‘opt-out’ from SSNAP at individual sites) as the audit has received exemption via section 251 of the NHS Act 2006, and so separate ethical approval for this work is not required. The section 251 approval comes from the Ethics and Confidentiality Committee of the National Information Governance Board (now superseded by the NHS Health Research Authority Confidentiality Advisory Group). The data controller is HQIP, and data access is managed through the the HQIP Data Access Request process, which attracts a £10,000 fee included in the application. The HQIP data access request group meets monthly, with outcomes communicated within 2 weeks of meetings. We plan to submit the HQIP request for access so that it is granted before planned project start (Feb 2018).

9.2 Qualitative research

For the qualitative work, we will seek the advice of our local research ethics committee (REC) on ethical matters as appropriate, and ethical approval will be sought via the Integrated Research Application System (IRAS).

Qualitative research will be conducted in accordance with the Data Protection Act and University Policy.

10. Patient and Public Involvement

This application was discussed with five members of the Peninsula Public Involvement Group. We This application was discussed with five members of the Peninsula Public Involvement Group. We discussed this with this group because they have experiences as in-patients or carers, and are interested in improving hospital care. They have also previously attended a workshop on simulation modelling in health services research and planning.

At the meeting they reviewed the Plain English Summary and gave feedback on its readability. They also discussed, with modellers Mike Allen and Kerry Pearn, the appropriateness and relevance of the proposed study and the research plan. In addition, they reviewed the qualitative research component in discussion with qualitative lead Julia Frost. Specific impact from these discussions were the following points, which have all been incorporated into the application:

- To frame the output of this study as a potential ‘decision-making tool’ for physicians
- To consider how, in the future, this decision-making tool might take patients’ views on risk/benefit balance into account
- To include some pilot modelling on outcomes at patient level
- To take different data displays to the qualitative interviews, to find out which forms of data might be most helpful
- To invite the Stroke Association to be a collaborator to the study

Overall the public advisors were positive to the study and see it as vital in improving understanding of when administration of thrombolysis is appropriate and when it’s not.

After these discussions the Plain English Summary was amended and the new version was reviewed by two people, with direct experience of stroke as next of kin to a patient, who were unable to attend the meeting. Finally, this section and the involvement plans described below were reviewed by members of the Peninsula Public Involvement Group.

PPI involvement throughout project

Support for involvement: This study is supported by the PenCLAHRC involvement team at the University of Exeter Medical School. The team has a track record of supporting people with complex needs to be research advisors. All involved patents and carers will have their travel fully reimbursed and their time recognised with a thank-you payment. Thank-you payments will be higher for patients on the steering committee, in recognition of their longer travel time. We have budgeted for an introductory training course for people new to involvement. Co-applicant Kristin Liabo will provide tailored support in advance of, during and after meetings.

Governance: Penny Thompson cares for her husband who has several health problems after a stroke. She is a collaborator on the study. Her husband has decided that he will contribute knowledge when he feels it is appropriate. PT also cared for her father after he had a stroke. As collaborator PT will attend quarterly project team meetings. We will recruit another stroke survivor to attend with her. Two patient or carer members will be invited to sit on the study steering committee. They will be asked through a national organisation so we have input from outside the South West.

Involvement group: We will set up a group of stroke patients and carer advisors to the study. The group will consist of 6-8 members and will meet three times during to discuss: research findings (meeting 1),

how the decision-making tool might take patients' views on risk/benefit balance into account (meeting 2), planning for the stakeholder engagement event described in the qualitative research section of this proposal (meeting 3). Patients and carers who are collaborators, and those who sit on the study steering committee, will be invited to join this group. If they prefer, they can input to the group remotely.

A workshop will be organised with up to 30 stakeholders, including people who have had a stroke, carers, health professionals and NHS managers. The workshop will follow a structured format where participants are presented with early findings from the research, including the pathway simulation and machine learning models, the focus groups and the interviews. Our patient and carer collaborators will also present on their contribution to the study. Contingent on the nature of preliminary findings, activities will first be in separate stakeholder groups, followed by mixed groups of people from different backgrounds. This process and the multiple perspectives will be recorded in several ways: 1) by note-takers within sessions (JF, KL); 2) the researchers' participant observation notes made immediately after the stakeholder workshop and 3) participants' flipchart notes and summaries made during certain sessions. With consent, discussions will be audio-recorded, although we do not anticipate transcribing the whole event, rather recording will provide an aide memoir of the breadth of discussions as opposed to the attribution of data at an individual level. Thematic analysis of this data will augment and triangulate preliminary findings, to inform the development of an intervention to support quality improvement in thrombolysis practice. This method helps contribute to the study with stakeholders' concerns and offer additional interpretations and suggestions for implementation based on the study's findings.

References

1. Feigin VL, Forouzanfar MH, Krishnamurthi R, Mensah GA, Connor M, Bennett DA, et al. Global and regional burden of stroke during 1990-2010: Findings from the Global Burden of Disease Study 2010. The Lancet. 2014;383(9913):245–255.

2. Sentinel Stroke Audit Programme (SSNAP). Apr2015Mar2016-AnnualResultsPortfolio. 2016.

3. Newton JN, Briggs ADM, Murray CJL, Dicker D, Foreman KJ, Wang H, et al. Changes in health in England, with analysis by English regions and areas of deprivation, 1990-2013: A systematic analysis for the Global Burden of Disease Study 2013. The Lancet. 2015;386(10010):2257–2274.

4. Emberson J, Lees KR, Lyden P, Blackwell L, Albers G, Bluhmki E, et al. Effect of treatment delay, age, and stroke severity on the effects of intravenous thrombolysis with alteplase for acute ischaemic stroke: A meta-analysis of individual patient data from randomised trials. The Lancet. 2014;384(9958):1929–1935.

5. Meretoja A, Keshtkaran M, Saver JL, Tatlisumak T, Parsons MW, Kaste M, et al. Stroke thrombolysis: Save a minute, save a day. Stroke. 2014;45(4):1053–1058.

6. Saver JL. Time is brain - Quantified. Stroke. 2006;37(1):263–266.

7. Lyden P. Thrombolytic Therapy for Acute Stroke — Not a Moment to Lose. N Engl J Med. 2008;359(13):1393–1395.

8. Eissa A, Krass I, Bajorek BV. Barriers to the utilization of thrombolysis for acute ischaemic stroke. J Clin Pharm Ther. 2012;37(4):399–409.

9. Carter-Jones CR. Stroke thrombolysis: Barriers to implementation. Vol. 19. 2011. 53–57 p.

10. Engelter ST, Gostynski M, Papa S, Ajdacic-Gross V, Lyrer PA. Barriers to stroke thrombolysis in a geographically defined population. Cerebrovasc Dis. 2007;23(2–3):211–215.

11. Lahr MMH, Luijckx GJ, Vroomen PCAJ, Van Der Zee DJ, Buskens E. The chain of care enabling tPA treatment in acute ischemic stroke: A comprehensive review of organisational models. Vol. 260. 2013. 960–968 p.

12. Monks T, Pitt M, Stein K, James M. Maximizing the population benefit from thrombolysis in acute ischemic stroke: A modeling study of in-hospital delays. Stroke. 2012;43(10):2706–2711.

13. Lahr M, van der Zee D-J, Luijckx G, Vroomen P, Buskens E. A simulation-based approach for improving utilization of thrombolysis in acute brain infarction. Med Care. 2013;51(12):1101–1105.

14. Monks T, Pitt M, Stein K, James M a. Hyperacute stroke care and NHS England's business plan. BMJ. 2014;348(May 2014):g3049.

15. Kotsiantis SB, Zaharakis I, Pintelas P. Supervised machine learning: A review of classification techniques. Informatica. 2007;31:249–268.

16. Bembenek J, Kobayashi A, Sandercock P, Czlonkowska A. How many patients might receive thrombolytic therapy in the light of the ECASS-3 and IST-3 data? Vol. 5. 2010. 430 p.

17. Morris S, Hunter RM, Ramsay AIG, Boaden R, McKevitt C, Perry C, et al. Impact of centralising acute stroke services in English metropolitan areas on mortality and length of hospital stay: difference-in-differences analysis. Bmj. 2014;349(aug04 4):g4757–g4757.

18. Rokach L. Ensemble-based classifiers. Artif Intell Rev. 2010;

19. Markatou M, Tian H, Biswas S, Hripcsak G. Analysis of variance of cross-validation estimators of the generalization error. J Mach Learn Res. 2005;6:1127–1168.

20. Pope C, Mays N. Qualitative Research in Health Care. 2006.

21. Green J, Thorogood N. Qualitative Methods for Health Research. 2009.

22. Hangar HC, Fogarty B, Wilkinson TJ, Sainsbury R. Stoke patients' views on stroke outcomes: death versus disability. Clin Rehabil. 2000;14:417–424.

23. Flatharta TO, Khan A, Walsh T, O'Donnell M, O'Keefe ST. Advance preferences regarding thrombolysis in patients at risk for stroke: a cross-sectional study. QJM Mon J Assoc Physicians. 2015;

24. Miles MB, Huberman MA, Saldaña J. Qualitative Data Analysis. A Methods Sourcebook. Z Für Pers. 2014;28(4):485–487.

25. Healthcare Quality Improvement Partnership. Rising to the Challenge: The Fourth SSNAP Annual Report. 2017;

26. Raynor DKT, Blenkinsopp A, Knapp P, Grime J, Nicolson DJ, Pollock K, et al. A systematic review of quantitative and qualitative research on the role and effectiveness of written information available to patients about individual medicines. Health Technol Assess. 2007;

27. Frost J, Garside R, Cooper C, Britten N. A qualitative synthesis of diabetes self-management strategies for long term medical outcomes and quality of life in the UK. BMC Health Serv Res. 2014;

11. Appendices

11.1 Hospital performance parameters used in the pathway simulation model

The parameters used in the thrombolysis pathway simulation model are shown in table A1.

Index	Parameter	Distribution	Range used (base case)
1	Allowed onset-to-treatment time for age up to 80 (mins)	Fixed	270
2	Allowed onset-to-treatment time for age 80 + (mins)	Fixed	180
3	Age 80+	Bernoulli	0.413-0.513
4	Arrivals per year	fixed	300-800
5	Onset time known	Bernoulli	0.434-0.727
6	Onset-to-arrival <4hrs	Bernoulli	0.677-0.705
7	Onset-to-arrival time (mins)	Log (ln) normal	Mean (ln): 4.411-4.609 StdDev (ln): 0.475-0.588
8	Arrival-to-scan <4hrs	Bernoulli	0.78-0.97
9	Arrival-to-scan time (mins)	Log (ln) normal	Mean (ln): 2.477-3.776 StdDev (ln): 0.750-1.271
10	Ischaemic stroke	Bernoulli	0.85-0.89
11	Eligible for thrombolysis	Bernoulli	0.344-0.571
12	Scan-to-thrombolysis (mins)	Log (ln) normal	Mean (ln): 2.970-3.791 StdDev (ln): 0.578-0.849

Table A1: parameters used in the stroke pathway simulation model.

Notes:

1. *The allowed onset-to-treatment time for patients aged up to 80 is 4.5 hours.*
2. *The allowed onset-to-treatment time for patients aged 80+ is 3 hours.*
3. *The proportion of patients aged 80+ varies between hospitals.*
4. *The number of arrivals is the number of confirmed stroke patients in SSNAP per year*
5. *'Onset time known' is binary yes/no. Known onset time may be recorded as precise or estimated in SSNAP. Those without a known onset time cannot be treated with thrombolysis in the model.*
6. *'Onset to arrival <4hrs' is the proportion of patients with known stroke onset that arrive within 4 hours of onset. Those arriving more than 4 hours after onset cannot be treated with thrombolysis in the model.*
7. *'Onset to arrival time' is a log-normal distribution. It is applied only to those patients arriving within 4 hours of known stroke onset.*
8. *'Arrival to scan time <4hrs' is the proportion of patients (those with known stroke onset time and arriving within 4 hours of onset) that receive a CT head scan within 4 hours of arrival.*
9. *'Arrival to scan time' is a log-normal distribution. It is applied only to those patients arriving within 4 hours of known stroke onset, and receiving a scan within 4 hours of arrival.*
10. *'Ischaemic stroke' is the proportion of patients with ischaemic (rather than haemorrhagic stroke). It is applied only to those patients arriving within 4 hours of known stroke onset, and receiving a scan within 4 hours of arrival.*
11. *'Eligible for thrombolysis' is the proportion of ischaemic stroke patients (arriving within 4 hours of known stroke onset, and scanned within 4 hours of arrival) who are considered clinically eligible for thrombolysis.*

This figure is obtained by examining the proportion of ischaemic stroke patients who are scanned with at least 30 minutes left to thrombolyse who are given thrombolysis.

12. *'Scan to thrombolysis' is a log-normal distribution. It is applied only to those patients arriving within 4 hours of known stroke onset, receiving a scan within 4 hours of arrival, and are ischaemic strokes considered eligible for thrombolysis.*

11.2 SSNAP features used in pilot machine learning

The following factors were used in the pilot study:

Thrombolysis given	Anticoag before stroke_1
Hosp_1	Anticoag before stroke_NK
Hosp_2	Stroke severity group_1. No stroke symtpoms
Hosp_3	Stroke severity group_2. Minor
Hosp_4	Stroke severity group_3. Moderate
Hosp_5	Stroke severity group_4. Moderate to severe
Hosp_6	Stroke severity group_5. Severe
Hosp_7	Stroke Type_I
Male	Stroke Type_PIH
Age	S2RankinBeforeStroke
80+	S2NihssArrival
Onset Time Known Type_BE	S2NihssArrivalLocQuestions
Onset Time Known Type_NK	S2NihssArrivalLocCommands
Onset Time Known Type_P	S2NihssArrivalBestGaze
# Comorbidities	S2NihssArrivalVisual
2+ comorbidotes	S2NihssArrivalFacialPalsy
Hypertension	S2NihssArrivalMotorArmLeft
Atrial Fib	S2NihssArrivalMotorArmRight
Diabetes	S2NihssArrivalMotorLegLeft
TIA	S2NihssArrivalMotorLegRight
Co-mordity	S2NihssArrivalLimbAtaxia
Antiplatelet_0	S2NihssArrivalSensory
Antiplatelet_1	S2NihssArrivalBestLanguage
Antiplatelet_NK	S2NihssArrivalDysarthria
Anticoag before stroke_0	S2NihssArrivalExtinctionInattention

There are additional SSNAP items which we would request for the proposed study including (but not limited to):

- Has it been decided in the first 72 hours that the patient is for palliative care?
- What was the initial brain imaging modality?
- Assessment of ischaemic penumbra by perfusion imaging
- Did the patient have any complications from the thrombolysis? (With further detailed fields)
- What was the patient's NIHSS score at 24 hours after thrombolysis / intra-arterial intervention?
- Has the patient had a TIA within the last month?
- What was the patient's Barthel score before the stroke?