STROKE-IMPACT (RfPB) STAGE 1

PLAIN ENGLISH SUMMARY

AIM

Our aim is to improve the benefit that can be achieved from the use of the medical stroke treatment, 'thrombolysis', which breaks down clots causing most strokes.

BACKGROUND

Stroke is one of the leading causes of death and disability. Most strokes are caused by a clot in the brain. For these patients there is the possibility to break down the clot by use of thrombolysis, avoiding or reducing the disability that would have been caused. Thrombolysis use and speed varies a lot between hospitals, and the overall use is only about half of the NHS target.

Our modelling has already identified potential effects of the variation in treatment. One observation is that for people suffering a mild stroke, early use of thrombolysis appears to reduce the risk of death, but late use can increase the risk of death. This observation may be used to support a change in clinical practice. However, our current models are not the right type for us to state any cause-and-effect conclusions. To be able to confirm that early use of thrombolysis is actually causing the observed reduction in death we would use 'causal inference' analysis. By extending our current work to use these new methods, we will strengthen our findings by ruling out any current observation being due to a coincidence.

DESIGN AND METHODS

Our models are based on the Sentinel Stroke National Audit Programme stroke patient dataset. It contains a record for each patient that has had a stroke. No modelling method will be perfect for confirming a cause-and-effect relationship from routinely collected data. But by using a combination of methods we can test our conclusions as thoroughly as possible. We will hold stakeholder workshops to help us develop the work.

PATIENT AND PUBLIC INVOLVEMENT

We regularly discuss our work with an engaged group of patients and carers. This process has previously helped us to communicate the models and the results clearly, which in turn has improved our knowledge of the models. The patients and carers also guided our work to look more closely at what enhances patient outcomes, and not just at meeting arbitrary treatment targets.

LINK TO PATIENT BENEFIT

We have strong direct links for how our work will benefit patients. NHS-England will use our results to support specific hospitals with their improvements. The national stroke audit will incorporate our findings from spring 2024. The knowledge developed here will become part of our standard way of working.

DISSEMINATION

In addition to working with the national stroke audit and NHS-England, we will also publish our work in leading stroke journals, and present it at stroke conferences. All of our detailed work is published online for anyone to freely access and use.

1 WHAT IS THE PROBLEM?

Stroke remains one of the leading causes of death and disability, globally and in the UK. Thrombolysis with recombinant tissue plasminogen activator, can significantly reduce disability after ischaemic stroke, so long as it is given in the first few hours after stroke onset [1]. Despite thrombolysis being of proven benefit in ischaemic stroke, use of thrombolysis varies significantly both between and within European countries [2]. In England and Wales the national stroke audit reported that in 2021/22, thrombolysis rates for emergency stroke admissions varied from just 1% to 28% between hospitals. Overall, about 10% of patients receiving thrombolysis, against a NHS England long term plan that 20% of patients of emergency stroke admissions should be receiving thrombolysis.

Our research team uses various modelling techniques, including explainable machine learning to identify the variation that occurs across different hospitals during the first few hours of stroke care. We use these models to understand the the source and impact of that variation on patients [3;4]. We work together with the Sentinel Stroke National Audit Programme and NHS-England on reducing between hospital-variation in emergency stroke care.

2 WHY IS THIS RESEARCH IMPORTANT?

Variation in use of thrombolysis in emergency stroke care is hampering the ability to meet NHS targets on use of this disability-saving treatment, and is causing inequality of access to this treatment. Thrombolysis rates are, however, only half the story. As thrombolysis also carries a small risk of serious injury, it is important that thrombolysis use is increased in the right group of patients - those where thrombolysis is likely to improve the chance of a good outcome. We also look at patient outcomes (death and disability). But there is a risk we may be misled by observations in the data and in our models; what we think is a real relationship between two variables may be caused by another factor. The planned work seeks to add causal inference methodologies to our methods - these strengthen our understanding of cause-and-effect relationships and build trust that changes suggested would indeed lead to improved patient outcomes.

3 REVIEW OF EXISTING EVIDENCE

Studies have shown that reasons for low and varying thrombolysis rates are multi-factorial. Reasons include late presentation [2], lack of expertise [2] or lack of clear protocols or training [5], delayed access to specialists [6], and poor triage by ambulance or emergency department staff [5]. For many factors, the establishment of primary stroke centres has been suggested to improve the emergency care of patients with stroke and reduce barriers to thrombolysis [5], with a centralised model of primary stroke centres leading to increased likelihood of thrombolysis [7;8;9].

In addition to organisational factors, clinicians can have varying attitudes to which patients are suitable candidates for thrombolysis. In a discrete choice experiment [10], 138 clinicians considered hypothetical patient vignettes, and responded as to whether they would give the patients thrombolysis. The authors concluded that there was considerable heterogeneity among respondents in their thrombolysis decision-making. Areas of difference were around whether to give thrombolysis to mild strokes, to older patients beyond 3 hours from stroke onset, and when there was pre-existing disability.

Based on national audit data from three years of emergency stroke admissions, we have previously built models of the emergency stroke pathway using clinical pathway simulation to examine the potential scale of the effect of changing two aspects of the stroke pathway performance (1. the in-hospital process speeds, and 2. the proportion of patients with a determined stroke onset time), and using machine learning to examine the effect of replicating clinical decision-making around thrombolysis from higher thrombolysing hospitals to lower thrombolysing hospitals[3;4]. The machine learning model, which has 85% accuracy (ROC-AUC = 0.92), learned whether any particular patient would receive thrombolysis in any particular emergency stroke centre. Using these models we found that it would be credible to target

an increase in average thrombolysis in England and Wales, from 11% to 18%, but that each hospital should have its own target, reflecting differences in local populations. We found that the largest increase in thrombolysis use would come from replicating thrombolysis decision-making practice from higher to lower thrombolysing hospitals. Two other important factors influencing thrombolysis rates were determination of stroke onset time in some hospitals, and improving the speed of the in-hospital thrombolysis pathway.

We have extended the model to including explainability [11] (see figure 1). This revealed that lower thrombolysing hospitals, compared with higher thrombolysing hospitals, were particularly less likely to give thrombolysis to patients with 1) mild stroke, 2) imprecisely known onset time, 3) patient with prior disability, 4) any combination of those.

We are developing a web app (available at https://stroke-predictions.streamlit.app/ that allows teams to see how changing pathway process speeds, or changing decision-making, may change thrombolysis use and outcomes.Current NIHR-funded work (https://fundingawards.nihr.ac.uk/award/NIHR134326) is focussing on building explainable machine learning models to predict death and disability-level outcomes with and without thrombolysis (and with differing time to thrombolysis). The observational data set we are using has over 40,000 patients receiving thrombolysis, compared to 6,756 in the seminal meta-analysis conducted by Emberson et al. [1]. Use of observational data also allows the investigation of the effect of thrombolysis as use is extended outside of clinical trial inclusion criteria.

Some patterns are appearing in the analysis which may help guide clinicians on use of thrombolysis. In mild stroke (which was not generally included in the clinical trials), we observe that early thrombolysis is associated with a reduced chance of death, whereas later thrombolysis is associated with an increased chance of death (fig 2). It is possible, however, that this effect is co-incidental rather than causal. We therefore wish to strengthen this type of analysis, by combining explainable machine learning with a range of causal inference methods.

4 RESEARCH AIMS

The key aim of the proposed research is to incorporate causal inference methodology into our clinical pathway simulation and explainable machine learning analysis of emergency stoke care. We consider this a key methodological development, alongside explainable machine learning, to test, and build trust in, how changes in use of thrombolysis will effect patient outcomes.

4.1 SPECIFIC RESEARCH QUESTIONS

We are interested in incorporating causal inference analysis into a wide range of our work, but he have specific research questions to pin this development to. These research questions will help ensure advice given to stroke teams has the strongest foundations for improving outcomes.

* Is there good evidence that the hospital is causing the variation in thrombolysis use in subgroups of patients where we see significant between-hospital variation in use of thrombolysis (rather than this effect coinciding with other factors)? Example patient subgroups are 1) Mild stroke (at presentation), 2) Presence of prior disability, and 3) Imprecisely known stroke onset time.

* In subgroups of patients where we see significant changes in outcome, is there good evidence that the main feature of interest in the patient subgroup is causing the variation in outcome in these groups (rather than this effect coinciding with other factors)? Example patient subgroups include early vs. late thrombolysis in mild stroke (where early thrombolysis appears to reduce the odds of death, but late thrombolysis increases the odds of death)?

5. PROJECT PLAN

5.1 DATA

We use anonymous patient-level data from the Sentinel Stroke National Audit Programme (SSNAP, https://www. strokeaudit.org/). SSNAP collects data on essentially all emergency stroke units, from all stroke units in England, Wales, and Northern Ireland. The data includes a wide range of information such as data on times throughout the stroke pathway from stroke onset, pre-stroke disability (modified Rankin Scale, mRS), a breakdown of stroke symptoms using the NIHSS stroke scale, reperfusion treatment given (and timing), comorbidities, death in hospital, disability at discharge from inpatient care, and 6 month follow-up disability (the latter is complete for about 35% of discharged patients). Data is collected for about 85,000 patients each year, 10-11% of whom receive thrombolysis.

5.2 METHODS

When using observational data, no single method can provide proof of causality. Instead we will use a combination of methods, each of which will contribute to understand causality. These methods will be applied to 1) examining variation in apparent differences between hospitals in which patients receive thrombolysis, and 2) apparent differences in outcomes in patient subgroups. The methods we will investigate are:

* Explainable machine learning with SHAP [12] - We will continue to use SHAP analysis to provide machine learning models at a global level and for individual predictions.

* DAG (directed acyclic graphs) [13] - Articulating proposed causal relationships through use of DAGs help to make assumed or hypothesised causal and non-causal relationships clear. They are easily understood by non-technical audiences, and so form an excellent basis for discussions and workshops to explore proposed causal relationships with clinical experts. We will use these, alongside results from other methods, in co-production workshops.

* Target trial emulation [14] - This method involves mimicking the original thrombolysis clinical trials from observational data, using the original trial inclusion criteria, and then extending to groups that were not included in the original trial.

* Natural experiments [15] - Natural experiments compare outcomes when changes in services in any hospital are known to have occurred. We may, for example, look for shifts in thrombolysis use and pathway speed before/during/after COVID lockdowns.

* Covariate adjustment [16] - We will identify all of the potential confounding variables, and include all of them that we can in a predictive model. Confounders are features, such as stroke severity, that affect both the probability of receiving thrombolysis, and the patient outcome.

* Use of propensity score as an adjusting feature [17] - 'Propensity', the likelihood of a person receiving thrombolysis (estimated from a separate machine learning model), is added to the outcome prediction model, allowing the outcome model to be adjusted for patients' suitability for treatment. This helps to identify, and correct for, better outcomes occurring in patients suitable for thrombolysis, even if they did not receive thrombolysis.

* Matching - Creation of matched populations of control and test patients. Matching attempts to create cohorts of patients that are similar in both groups apart from the feature with proposed causal influence, allowing for an estimation of the effect of the causal feature under study. Matching may be performed using nearest-neighbour approaches [18] or by propensity scores [17].

* Stratification of results by propensity score [17] - allowing comparisons of groups of patients who have similar suitability for treatment with thrombolysis.

* Instrumental variable analysis [19] - This method compares outcomes depending on an 'instrument' that does not directly effect outcome, but effects the proposed causal feature. In our current model we isolate how a hospital affects the likelihood that a patient will be given thrombolysis; we will test this as an instrument to investigate whether there is a relationship between this feature and outcomes. This could provide a novel link between advanced explainable machine learning methods and causal inference work.

* Inverse propensity weighting [20] - Weighting the contribution of patients to the final outcome measurement based on their likelihood of receiving, or not-receiving, thrombolysis. This method gives most weight to patients who received thrombolysis, who are more like patients who usually do not receive thrombolysis (and vice versa). It also gives more weight to patients who are borderline in whether they would receive treatment or not.

* Double machine learning [21] - This method combines two machine learning models, one to predict use of treatment and the other to predict outcome from that treatment. The treatment model estimates the probability of treatment given the observed covariates, while the outcome model estimates the expected outcome given the observed covariates and treatment. The errors (residuals) in the models are used as surrogates for missing information in each models. This method isolates a features causal relationship with the outcome by fitting the residuals of the feature prediction to the residuals of the outcomes.

All work will be performed in Python using established libraries for each of the methods (especially the XGBoost, SHAP, DoWhy and EconML libraries). Work will be conducted in Jupyter notebooks, all of which will be published in an online Jupyter Book.

5.3 CO-PRODUCTION WITH THE NHS, AND LINK TO PATIENT BENEFIT

During this project we will hold stakeholder workshops to obtain feedback on the work. This will use our existing stakeholder network across the SSNAP, Integrated Stroke Delivery Networks (ISDN) and Integrated Care Systems (ICS). As we are working alongside NHS-England and NHS-Elect with teams with low thrombolysis (see paragraph below), we will also use this experience to refine project outputs.

A pilot web portal is being made available from autumn 2023, providing analysis of thrombolyis use at each hospital, comparing decisions made to other hospitals, and providing modelling how changing process speeds or clinical decisions would likely effect thrombolysis use and outcomes (based on a mathematical model of outcome depending on time to thrombolysis).

The modelling work currently being performed (https://fundingawards.nihr.ac.uk/award/NIHR134326) is planned to be incorporated into the national stroke audit from Spring 2024. This will provide teams with realistic target thrombolysis use for their own patient population, and identify which area of the stroke pathway to focus on (for example pathway speed, ascertainment of stroke onset time, thrombolysis decision making). Additionally, NHS-Elect (who are working with NHS-England, the SAMueL team, and the national stroke audit) have an NHS improvement target to "Improve access to thrombolysis such that by the end of 2027/28, 20% of stroke patients will receive thrombolysis treatment". This collaborative will be working with 6 low-thrombolysing teams in the first instance, before extending work to all emergency stroke teams. As part of the quality improvement work, the SAMueL team will be providing modelling on use of thrombolysis, including on variation in decision-making between hospitals. We consider it important that this work should, if at all possible, include strengthened analysis of causality (the links between changes to pathway and decision making, through to thrombolysis use, and through to outcome) so that we have greater confidence that changes suggested (and potentially made) will lead to the expected improvement in thrombolysis use and, most importantly, in better outcomes and not 'just' increased thrombolysis use.

5.4 DISSEMINATION

In addition to working with the national stroke audit and stroke teams directly (see above), we will also publish our work in leading stroke journals, and present it at stroke conferences. All of our detailed work is published online (e.g. see https://samuel-book.github.io/samuel_shap_paper_1).

5.5 PROJECT TEAM

The project teams bring together experience of stroke care, machine learning, epidemiology, clinical trials, and causal inference work.

* Michael Allen (Co-PI, 25% FTE) co-leads our NIHR funded work on explainable machine learning (https://fundingawards.nihr.ac.uk/award/NIHR134326), with a focus on the technical aspects of the project. MA has extensive experience of health systems modelling and machine learning.

* Kerry Pearn (Co-PI, 50% FTE) has been the principal developer of our explainable machine learning methodology for prediction of thrombolysis use and outcome. KP will co-lead the proposed work, with a focus on detailed work planning and execution, bringing in advice from co-investigators.

* Martin James (Co-I, 5% FTE) is a stroke consultant and clinical director of the national stroke audit. MJ co-leads our NIHR funded work on explainable machine learning (https://fundingawards.nihr.ac.uk/award/NIHR134326), with a focus on clinical oversight of the project. MJ will provide clinical oversight for this proposed project.

* Mark Kelson (Co-I, 5% FTE) is a statistician working in the overlap between clinical trials, medical statistics, causal inference, reproducibility and data science. MK will advise on statistics and causal inference methods.

* Joao Delgado (Co-I, 10% FTE) is an epidemiologist focussing on management of conditions affecting cognitive function in old age, using traditional and modern epidemiology and statistical tools to analyse large datasets of electronic health records. JD will advise on epidemiological methodologies such as use of propensity scores.

* Lauren Asare (Researcher, 5%) is a research assistant in the University of Exeter PenARC Public and Patient involvement group. LA supports our stroke PCI team (which is chaired by a stroke survivor).

5.6 PATIENT AND CARER INVOLVEMENT

Our team has a dedicated PPI group who provide constant input into our work and have provided input to this bid. The PPI group have been a key voice in guiding our work to look at what most benefits patients, and not just meeting arbitrary targets on thrombolysis use. The PPI team is chaired by Leon Farmer, a stroke supervisor, with guidance from Lauren Asare of the University of Exeter PenARC PPI team.

5.7 PEOPLE DEVELOPMENT

In addition to development of the methodology, this project will be used to especially develop both the technical and project management skills and capabilities of Kerry Pearn (Co-PI) who will co-lead the project with Michael Allen (Co-PI). The stroke modelling and data science team is a stable team, funded by research grants and NHS-contracted work, and the techniques and skills developed here will find long-term use and benefit.

5.8 CONTRIBUTION BEYOND THIS PROJECT

We work on a range of projects that would benefit from this method development in our work. For example, we are working on stroke projects focussing on the pre-hospital pathway(https://fundingawards.nihr.ac.uk/award/NIHR202361) and the use of mobile stroke units

(https://fundingawards.nihr.ac.uk/award/NIHR153982). The methodological development above will add value to both of these, and also to other similar, projects. For example, it has previously been assumed that outcome from hemorrhagic stroke (the 20% of strokes that are caused by a bleed rather than a clot) is independent of ambulance travel time, but some recent clinical trials on taking stroke patients further, to a hospital with more capabilities for stokes caused by clots, have suggested this may not be true. Using methods such as clinical trial emulation and our large database of stroke data we will have the potential to enhance our pre-hospital stroke care model to better model outcomes of haemorrhagic strokes with alternative pre-hospital pathways.

6 INTELLECTUAL PROPERTY

We work on a principal that public funding should give public benefit, and so all our detailed work (code) is shared with an Open Licence, allowing anyone to use or develop it.

7 REFERENCES

[1] Emberson, J. 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 384, 1929–1935 (2014).
[2] Aguiar de Sousa, D. et al. Access to and delivery of acute ischaemic stroke treatments: A survey of national scientific societies and stroke experts in 44 European countries. European Stroke Journal 4, 13–28 (2019). URL https://www.ncbi.nlm.nih.gov/pmc/artic les/PMC6533860/.
[3] Allen, M. et al. Using simulation and machine learning to maximise the benefit of intravenous thrombolysis in acute stroke in England and Wales: the SAMueL modelling and qualitative study. Health and Social Care Delivery Research 10, 1–148 (2022). URL https://www.journalslibrary.nihr.ac.uk/hsdr/GVZL5699/.
[4] Allen, M. et al. Use of Clinical Pathway Simulation and Machine Learning to Identify Key Levers for Maximizing the Benefit of Intravenous Thrombolysis in Acute Stroke. Stroke 53, 2758–2767 (2022). URL https://www.ahajournals.org/doi/10.1161/STROKEAH A.121.038454. Publisher: American Heart Association.
[5] Carter-Jones, C. R. Stroke thrombolysis: Barriers to implementation. International Emergency Nursing 19, 53–57 (2011).
[6] Kamal, N. et al. Delays in Door-to-Needle Times and Their Impact on Treatment Time and Outcomes in Get with the Guidelines. Stroke 48, 946–954 (2017). Number: 4.
[7] Lahr, M. M. H., Luijckx, G.-J., Vroomen, P. C. A. J., van der Zee, D.-J. & Buskens, E. Proportion of patients treated with thrombolysis in a centralized versus a decentralized acute stroke care setting. Stroke 43, 1336–1340 (2012).
[8] Morris, S. et al. Impact of centralising acute stroke services in English metropolitan areas on mortality and length of hospital stay: difference-in-differences analysis. Bmj 349, g4757–g4757 (2014). URL http://www.bmj.com/cgi/doi/10.1136/bmj.g4757.
[9] Hunter, R. M. et al. Impact on clinical and cost outcomes of a centralized approach to acute stroke care in London: a comparative effectiveness before and after model. PloS One 8, e70420 (2013).
[10] De Bru´n, A. et al. Factors that influence clinicians' decisions to offer intravenous alteplase in acute ischemic stroke patients with uncertain treatment indication: Results of a discrete choice experiment. International Journal of Stroke 13, 74–82 (2018). URL https://doi.org/10.1177/1747493017690755. Publisher: SAGE Publications.
[11] Pearn, K. et al. What would other emergency stroke teams do? Using explainable machine learning to understand variation in thrombolysis practice. European Stroke Journal 23969873231189040 (2023). URL https://doi.org/10.1177/23969873231189040..
[12] Aas, K., Jullum, M. & Løland, A. Explaining individual predictions when features are dependent: More accurate approximations to Shapley values (2020). URL http://arxiv.org/abs/1903.10464. Number: arXiv:1903.10464 arXiv:1903.10464 [cs, stat].

[13] Tennant, P. W. G. et al. Use of directed acyclic graphs (DAGs) to identify confounders in applied health research: review and recommendations. International Journal of Epidemiology 50, 620–632 (2021). URL https://doi.org/10.1093/ije/dyaa213.

[14] Bigirumurame, T. et al. Current practices in studies applying the target trial emulation framework: a protocol for a systematic review. BMJ Open 13, e070963 (2023). URL https://bmjopen.bmj.com/content/13/6/e070963. Publisher: British Medical Journal Publishing Group Section: Research methods.

[15] Craig, P., Katikireddi, S. V., Leyland, A. & Popham, F. Natural Experiments: An Overview of Methods, Approaches, and Contributions to Public Health Intervention Research. Annual Review of Public Health 38, 39–56 (2017).

[16] Igelstr¨om, E. et al. Causal inference and effect estimation using observational data. J Epidemiol Community Health 76, 960–966 (2022). URL https://jech.bmj.com/content/76/11/960. Publisher: BMJ Publishing Group Ltd Section: Glossary.

[17] Rosenbaum, P. R. & Rubin, D. B. The Central Role of the Propensity Score in Observational Studies for Causal Effects. Biometrika 70, 41–55 (1983). URL https://www.jstor.org/stable/2335942. Publisher: [Oxford University Press, Biometrika Trust].

[18] Stuart, E. A. Matching methods for causal inference: A review and a look forward. Statistical science : a review journal of the Institute of Mathematical Statistics 25, 1–21 (2010). URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2943670/.

[19] Stel, V. S., Dekker, F. W., Zoccali, C. & Jager, K. J. Instrumental variable analysis. Nephrology Dialysis Transplantation 28, 1694–1699 (2013). URL https://doi.org/10.1093/ndt/gfs310.

[20] Glynn, A. N. & Quinn, K. M. An Introduction to the Augmented Inverse Propensity Weighted Estimator. Political Analysis 18, 36–56 (2010). URL https://www.cambridge.org/core/journals/political-analysis/article/an-introduction-to-the-augmented-inv erse-propensity-weighted-estimator/4B1B8301E46F4432C4DCC91FE20780DB. Publisher: Cambridge University Press.

[21] Chernozhukov, V. et al. Double/Debiased Machine Learning for Treatment and Causal Parameters (2017). URL http://arxiv.org/abs/1608.00060. ArXiv:1608.00060 [econ, stat].
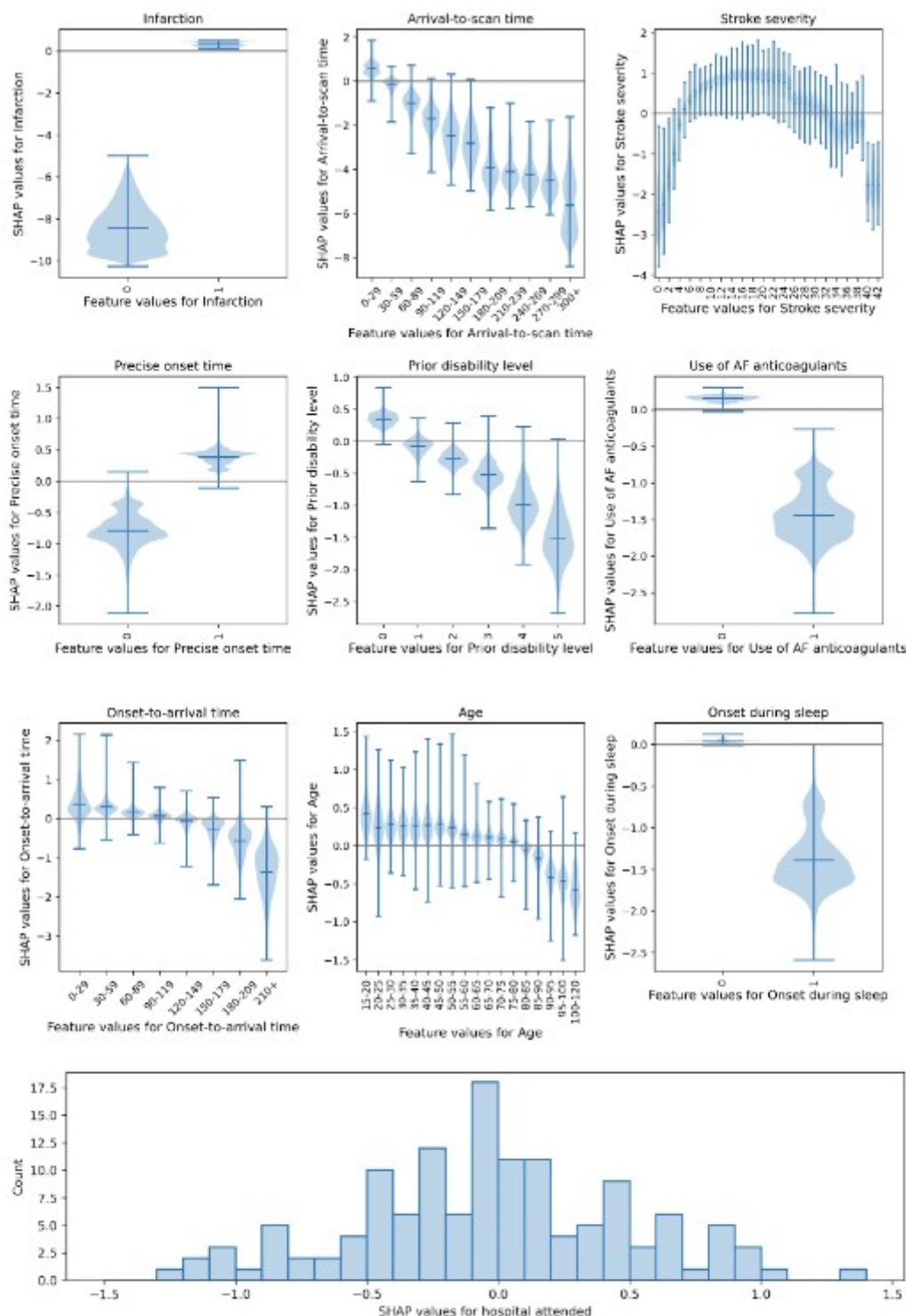
Figure 1: Plots showing the relationship between feature values and how the log-odds of receiving thrombolysis are effected (the *SHAP* value). A SHAP value above zero represents an increased likelihood to receive thrombolysis, and a SHAP value below zero represents a decreased likelihood to receive thrombolysis. Top: Violin plots showing the relationship between SHAP values and feature values. The horizontal blue line shows the median SHAP value. The black horizontal line marks the SHAP value of zero. The plots are ordered in ranked feature importance (using the mean absolute SHAP value across all instances). Bottom: Histogram showing the frequency of SHAP values for the hospital attended.
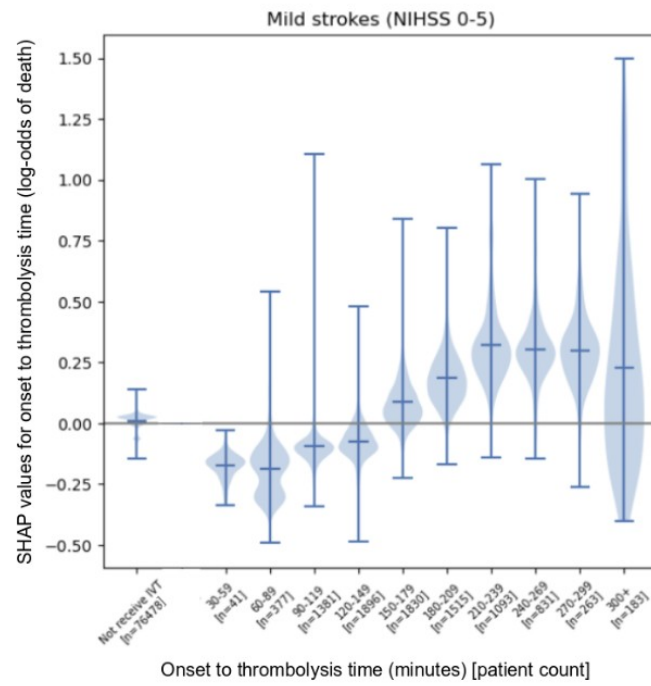
Figure 2: Violin plot showing the relationship between time to thrombolysis and and how the log-odds of death are effected (the *SHAP* value) in mild strokes (NIHSS 0-5).