

UCLA

UCLA Previously Published Works

Title

Using artificial intelligence for improving stroke diagnosis in emergency departments: a practical framework.

Permalink

<https://escholarship.org/uc/item/4bt7w3th>

Authors

Abedi, Vida

Khan, Ayesha

Chaudhary, Durgesh

et al.

Publication Date

2020

DOI

10.1177/1756286420938962

Peer reviewed

Using artificial intelligence for improving stroke diagnosis in emergency departments: a practical framework

Vida Abedi, Ayesha Khan, Durgesh Chaudhary, Debdipto Misra, Venkatesh Avula, Dhruv Mathrawala, Chadd Kraus, Kyle A. Marshall, Nayan Chaudhary, Xiao Li, Clemens M. Schirmer, Fabien Scalzo, Jiang Li and Ramin Zand 

Abstract: Stroke is the fifth leading cause of death in the United States and a major cause of severe disability worldwide. Yet, recognizing the signs of stroke in an acute setting is still challenging and leads to loss of opportunity to intervene, given the narrow therapeutic window. A decision support system using artificial intelligence (AI) and clinical data from electronic health records combined with patients' presenting symptoms can be designed to support emergency department providers in stroke diagnosis and subsequently reduce the treatment delay. In this article, we present a practical framework to develop a decision support system using AI by reflecting on the various stages, which could eventually improve patient care and outcome. We also discuss the technical, operational, and ethical challenges of the process.

Keywords: acute stroke, artificial intelligence, cerebrovascular disease/stroke, computer aided diagnosis, ischemic stroke, machine learning, stroke diagnosis, stroke in emergency department

Received: 21 May 2020; revised manuscript accepted: 2 June 2020.

Introduction

Stroke is the fifth leading cause of death in the United States and a significant cause of severe disability in adults.¹ Each year, around 800,000 Americans experience a new or recurrent stroke.² Rapid diagnosis and treatment of stroke is crucial and leads to improved outcomes and prognosis among patients treated within the 'Golden Hour'.^{3,4}

However, strokes, especially posterior circulation strokes, are associated with significant (>10%) diagnostic error.⁵ The latter could be due to (1) some patients with acute stroke present with non-focal symptoms such as dizziness, diplopia, dysarthria, or ataxia,⁶ which may not trigger a neurology consult or a need for a more detailed neurological examination; (2) stroke is commonly misdiagnosed in younger patients^{7,8}; and (3) the emergency department (ED) is a challenging

environment for providers, especially with the multiplicity of care protocols, and the dynamic nature of patient care.^{8,9} Triage, consultations, admissions, discharge, and other steps in emergency care are time-sensitive, complex, and always changing to further improve efficacy and quality of care. Therefore, identifying potential stroke symptoms can be challenging,^{10–12} especially when the providers are in training.^{13,14} Besides, the risk of misdiagnosis can be higher among walk-in patients,¹⁵ when the providers do not receive a pre-arrival notification from emergency medical services,¹⁶ or when a neurologist is not readily available for an urgent consultation.^{17–19} Scoring systems for the diagnosis of stroke and recurrent stroke do not have a high sensitivity to diagnose the posterior circulation stroke.^{20,21} Furthermore, these tools are also not automatic, and require that the physicians suspect stroke as a differential diagnosis to apply the scoring system.

Ther Adv Neurol Disord

2020, Vol. 13: 1–8

DOI: 10.1177/
1756286420938962

© The Author(s), 2020.
Article reuse guidelines:
sagepub.com/journals-
permissions

Correspondence to:

Ramin Zand
Neuroscience Institute,
Geisinger Health
System, Stroke Program,
Geisinger Northeast
Region, GRA Stroke Task
Force, American Heart
Association, Department
of Neurosciences, 100 N
Academy Ave, Danville, PA
17822-2101, USA
rzand@geisinger.edu;
ramin.zand@gmail.com

Vida Abedi
Department of Molecular
and Functional Genomics,
Geisinger Health System,
Danville, PA, USA

Biocomplexity Institute,
Virginia Tech, Blacksburg,
VA, USA

Ayesha Khan
Durgesh Chaudhary
Clemens M. Schirmer
Neuroscience Institute,
Geisinger Health System,
Danville, PA, USA

Debdipto Misra
Dhruv Mathrawala
Division of Informatics,
Geisinger Health System,
Danville, PA, USA

Venkatesh Avula
Jiang Li
Department of Molecular
and Functional Genomics,
Geisinger Health System,
Danville, PA, USA

Chadd Kraus
Kyle A. Marshall
Department of Emergency
Medicine, Geisinger Health
System, Danville, PA, USA

Nayan Chaudhary
Xiao Li
Genentech/Roche inc.,
South San Francisco, CA,
USA

Fabien Scalzo
Department of Neurology,
University of California,
Los Angeles, CA, USA
Department of Computer
Science, University of
California, Los Angeles,
CA, USA

Artificial intelligence (AI), a computational framework meaning to emulate human insight, is one of the most transformative technologies.^{22,23} The era of augmented intelligence in healthcare is driven by the notion that intelligent algorithms can support providers in diagnosis, treatment, and outcome prediction, especially with growing digital and connected patient data and advances in computational abilities.^{24–26} The augmented-diagnostic model for stroke may be particularly helpful in low volume or non-stroke centers' ED, where emergency providers have limited daily exposure to stroke. An automated, computer-assisted screening tool that can be seamlessly integrated into clinical workflow to quickly analyze patient symptoms and clinical data and suggest a diagnosis of stroke ("StrokeAlert" pop-up) in an ED setting could be valuable. Such a system will also help bring access and timely diagnosis for patients who choose to self-present to an ED. In this paper, we present a practical framework and summarize the stages needed to create a machine learning (ML)-enabled clinical decision support system for the screening of stroke patients in ED using data from electronic health records (EHRs) combined with the patient's presenting symptoms at the point of care. We have assembled a team of experts and are leading such effort at Geisinger. Figure 1 summarizes the key steps of such a system.

Building the training and testing cohorts

ML-enabled clinical decision support systems, with providers-in-the-loop, are essential advances in ensuring that decisions at the bedside are timely and data-driven. However, transparent reporting of prediction models is key to building confidence and improving reproducibility regardless of the study design or findings.²⁷

Case/control design

The initial phase is to create representative examples for model training. The inclusion and exclusion criteria should be restrictive enough to ensure that cases and controls have clear separation and are aligned with clinical pipelines. For instance, the case (confirmed stroke and transient ischemic event) cohort should have at least these conditions: (1) patient encounters should be of a minimum duration to ensure the severity of the condition and to remove noises from repetitive coding (e.g. >24h); (2) discharge codes should

be focused on primary diagnosis; (3) patient should have had confirmatory neuroimaging [e.g., brain magnetic resonance imaging (MRI)]. The same level of detail should be given to the design of inclusion and exclusion criteria to create a labeled dataset for the control group(s). The goal of creating control groups is to capture stroke mimics, stroke misdiagnoses, while capturing some level of diversity by including patients with similar general presentations as stroke, including, for instance, patients with a discharge diagnosis of migraine headache, seizure, and peripheral neuropathy.

Data extraction/processing

Maintaining data integrity during data aggregation is key. The past medical history should be defined carefully and not include new information from the index encounter. Quality control methods should (1) ensure variables have the same units and correct if needed; (2) remove clinically implausible values according to expert knowledge and aligned with the available literature; (3) apply filters to capture the relevant data elements within the desired timeframe; (4) identify and remove extreme outliers for longitudinal data and where multiple values are available; (5) use median as opposed to mean if multiple values are available; and (6) use imputation techniques apply to the training and testing dataset separately. If a laboratory value is missing, an estimated value should be generated to be used in the model. Nonetheless, in an operational stroke prediction model, laboratory values might not be essential, since, in real life, the diagnosis of stroke is made before any laboratory results are available. However, for model building, the results might be important to confirm other diagnoses. Some imputation methods include using median or mean values, K-nearest neighbors, multivariate imputation by chained equation (MICE),²⁸ or imputation designed for EHR.²⁹ Extraction and processing of clinical notes, especially triage/ED provider notes with information about the patient's symptoms, are critical for a technology designed for acute conditions. During natural language processing (NLP), a set of positive and negative keywords such as dizziness, vertigo, headache, confusion, etc., are extracted from a subset of cases and controls, are used to build a custom dictionary for the NLP. These keywords are expanded to medical concepts using the Unified Medical Language System (UMLS)

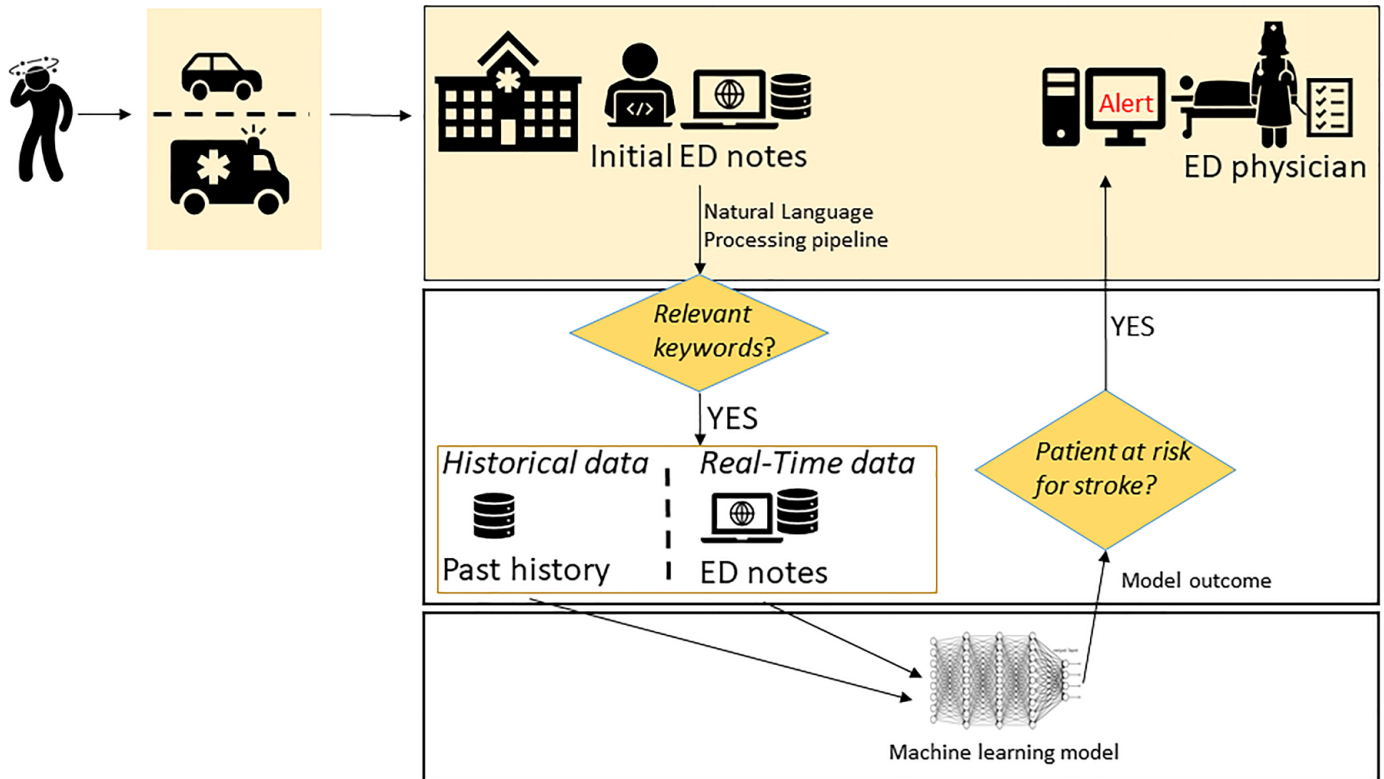


Figure 1. Key steps for a stroke ML-enabled decision support system for EDs. ED, emergency department; ML, machine learning.

dictionary,³⁰ allowing for efficient translation and interoperability. The NLP pipeline will generate added insights, such as the polarity of the words and context.

Designing the ML-enabled diagnostic tool

The ML development process will consist of various iterative steps for training, testing, and predicting the probability of patients presenting to ED with stroke.

Exploratory data analysis will help modelers understand feature distribution, multicollinearity among features, missing data, data quality. Features that are irrelevant or partially relevant (such as procedure codes that are no longer in active use), or highly sparse can be removed or merged. Feature selection will help reduce overfitting, training time, and improve the model accuracy. Feature engineering, also an important step, can be used to construct robust high-level feature representations from complex and high-dimensional concepts (e.g. diagnoses, medications). During the model

development, typically, an 80–20 split is performed to test the model performance; that is 20% of the data is marked as unseen and is used for model testing, while 80% is used for model development. Furthermore, as the number of cases is likely to be significantly less than the number of controls, it is crucial to address the class imbalance. Standard techniques to address class imbalance include up-sampling (by use of SMOTE algorithm, etc.) the minority class and/or down-sampling the majority class.³¹

Training predictive models are done by using the training dataset to identify the best performing candidate model. Models like logistic regression, decision trees, random forest, autoencoders, and neural networks can be used in the training process. Typically, an interpretable framework such as logistic regression is used for benchmarking. Finally, nested K-fold cross-validation applied to 80% of the data should be performed for hyperparameter tuning and making modeling framework choices – to avoid the underfitting and overfitting. Various model metrics such as the

area under the receiver operating characteristics curve, sensitivity, specificity, F-score, positive predictive value (PPV), negative predictive value (NPV), and missed classification rates can be used to identify the final candidate model for implementation and testing, prospectively. These metrics should be used for model selection only after careful evaluation to better understand the clinical needs in actual care settings. In the case of stroke, the cost of misdiagnosis is asymmetric, meaning that underdiagnosis (labeling true stroke patients as non-stroke) might have a higher consequence than overdiagnosis. Therefore, the system should enjoy a high sensitivity and NPV while keeping the specificity in a reasonable range. Finally, depending on the model used, knowing the most influential driving features helping the stroke prediction for each patient would be helpful and can provide insights into the prediction and ultimately assist the physician in decision making.

Workflow and system implementation

In a typical ED setting, a patient arrives at the hospital, either by ambulance or by a private vehicle. Although our proposed ML-enabled clinical decision support system is designed to work for all patients regardless of the arrival mode, we believe such a system will be more beneficial among patients self-presenting with milder and atypical symptoms. These patients meet a nurse at the point-of-care desk where they are asked preliminary questions while their vitals are recorded. The patients' symptoms and any pertinent information are entered in the EHR and can be available to the NLP pipeline for modeling. The patient is then sent back to the waiting area. The ED providers rely on the clinical presentation and vitals to prioritize the patients and performing the physical examination (which may not include neurological examination), ordering labs, and, in some cases, imaging procedures.

The average time (in the United States) in EDs before admission is approximately 5.5 hours, and time until the patient is sent home is approximately 3 hours.³² The waiting time varies depending on the season and the physicians' workload. The imaging procedures ordered in ED also vary as per hospital protocol. Furthermore, if only brain computed tomography (CT) is ordered, it is still likely that the ischemic stroke is missed for a patient with atypical symptoms, as the head CT

cannot reveal a hyperacute stroke in the majority of cases, and it has reduced sensitivity for lacunar strokes.³³ Despite the rapid increase in the use of advanced neuroimaging, it may be challenging to reduce misdiagnosis of stroke as the use of urgent MRI to diagnose stroke in ED is still limited.³⁴ Nevertheless, in reality, a provider should consider the possible stroke diagnosis to order additional neuroimaging.

Patients' notes at the point-of-care contain vital information that, if combined with the patients' profile and medical history, can be ingested by an intelligent system to identify at-risk patients. Implementation of such a system must be seamless without affecting the ED workflow. If a patient has a relatively significant chance of stroke, a 'stroke alert' will be generated when the patient's chart is viewed by the next ED provider. The provider will then have a chance to act upon the alert, and, if needed, call stroke-alert, request an urgent neurological consult, or order confirmatory imaging.

System adoption and evaluation

Promoting adoption

A methodical plan for promoting adoption should be part of a thoughtful implementation. In general, physicians have relatively positive attitudes toward the idea of decision support systems.^{35,36} However, many challenges including low specificity,^{37,38} work-flow interruption,³⁹⁻⁴¹ computer literacy and confusing interface,^{42,43} low confidence in the evidence,⁴⁴ awareness of the information,⁴⁵ requirement for a lot of data,^{41,46} interference with physician autonomy,^{35,47} or lack of relevance,⁴¹ limit the effective use and adoption of such systems in many healthcare systems. "Alert fatigue" can be caused by poorly designed and implemented clinical decision support system.^{35,48-50} Too many alerts will discourage adoption.

Improving coordination and capacity building

To identify barriers and facilitators in this process, it is essential to design a model based on the unified theory of acceptance and use of technology.^{50,51} A proposed model for user acceptance and adaptation should include the following stakeholders in the planning stages: (1) end-users (care providers); (2) hospital and service-line leaders; (3) EHR engineers, innovation team;

and (4) next-providers (inpatient/outpatient neurologists/hospitalists). The discussion objectives should focus on understanding the current workflow; how the intervention(s) may impact the people and processes; what the care providers' perceptions are; and how to achieve their buy-in while communicating the current gaps, clinical goals, and how the proposed system can help.

System evaluation

A mixture of targeted chart-review, systematic evaluation, and assessment of trends and rate of misidentification is important. The process should be designed to be agile and iterative. Prospective evaluation is critical and cannot be a one-time process. As long as a decision support system remains in operation, the model outcome should be re-assessed at regular intervals, and integration of variables with most and least promising relevance can be reassessed. Fine-tuning should be subtle and iterative to ensure smooth and clear improvements in both effectiveness and performance. Presenting the summary findings to the stakeholders to demonstrate transparency, gaps, areas of misclassifications, and the overall added value is important.

Challenges and opportunities

Technical challenges – tool or model-dependencies

Selection of the right mix of tools, techniques, and languages for productionalizing. Specific tools/languages (e.g., Spark) might be conducive to handling large data volumes; however, they might not have mature data science libraries such as the ones offered in Python to develop stable models. The proper selection of tools will cause downstream technical challenges, especially if the designed pipeline is not agnostic to the implementation language.

Model drifting. Models deteriorate in terms of their predictive power and clinical utility if not continuously adjusted. In healthcare, addressing model drift takes on a larger dimension since changing trends in population health manifests itself as both concept drift and data drift. Continuous domain adaptation is an active field of research to address such challenges.⁵² Healthcare systems that are in stable regions with low drop-out rates (such as Geisinger Health System) are

better equipped to incorporate continual learning in training the model for improved clinical utility.

Model generalizability. Developing generalizable models require a comprehensive and multi-level view of the patients with an added effort to ensure adequate patient representation to reduce algorithmic bias. The utilization of data from two or more centers will be important to develop generalizable models. Techniques such as transfer learning can be designed to evaluate model generalizability and transferability across health systems.⁵³ These solutions are critical for the development of models that can function well in smaller systems, to drive technological advances for the mainstream, in rural and urban areas alike.

Operational challenges

Implementation of an ML model to provide predictions and recommendations in real-time in EHR requires specialized programming expertise and purchase of specialized products from the vendor, which could be prohibitive for smaller healthcare systems with limited resources, especially since, for a real clinical utility, there will always be a need for maintenance with added financial burden. Operational challenges also entail usability and adoption. A less discussed challenge is the need for continuous learning. Feedback loop process creation can be a tool to facilitate creating an automated process to generate data used for continual learning of the model to improve performance; the latter needs commitments from users to add needed information into the system. Adding additional steps to the already busy schedules of the users is a challenge; however, the development of tools with clinical-experts-in-the-loop from the initial phase could provide opportunities for better adoption.

Ethical challenges

Defining how an “ethical” AI system should perform in this context is somewhat subjective and encompasses our experience in the field and perception about how the AI software operates. Overall, the generic goal would be to ensure fairness, efficiency, efficacy, and patient safety. In an ideal scenario, a system would benefit all identified groups equally; however, in practice, such a goal is often impossible, and it will be necessary to define an acceptable bias. Rigorous regulatory

institutions, such as the United States Food and Drug Administration (FDA), are starting to guide the development and maintenance of AI systems to ensure compliance and best practices for data-driven models. Nevertheless, the highest standard of AI-driven triage system for stroke will require to go a step further, and can be achieved only by a joint effort between care providers and modelers; the former will be able to define the needs, current limitations, and acceptable biases regarding stroke diagnostic, while the latter will provide insights about objective functions, computational fairness constraints, and estimation of projected model performance. A possible solution is to instantiate a Clinical-AI Review Board within institutions.

Our roadmap and framework targets ED; emergency medical services and telemedicine could also be a viable target for similar systems. The ML-enabled prediction model can seamlessly triage patients in real-time and alert the provider and the care team.

Acknowledgement

The authors would like to thank Sparkle Russell-Puleri for insightful discussion.

Author contributions

VA and RZ conceived of the presented idea. VA, AK, RZ wrote the manuscript. DPC, DMi, VeA, DMa, KAM, CMS, CK, NC, XL, SF, JL provided critical feedback and contributed to different sections of the manuscript. All authors reviewed and approved the final version of the manuscript.

Conflict of interest statement

The authors declare no competing interests. NC and XL are Genentech/Roche paid employees. Genentech/Roche had no role in study design, data collection, and interpretation, or the decision to submit the work for publication.

Funding

This work was sponsored in part by funds from the Geisinger Health Plan Quality Fund and National Institute of Health R56HL116832 (sub-award) to VA and RZ. The funders had no role in study design, data collection, and interpretation, or the decision to submit the work for publication.

ORCID iD

Ramin Zand  <https://orcid.org/0000-0002-9477-0094>

References

1. Murphy SL, Kochanek KD, Xu J, *et al.* Mortality in the United States, 2014. *NCHS Data Brief* 2015; 229: 1–8.
2. Benjamin EJ, Muntner P, Alonso A, *et al.*; American Heart Association Council on Epidemiology and Prevention Statistics Committee and Stroke Statistics Subcommittee. Heart disease and stroke statistics–2019 update: a report from the American heart association. *Circulation* 2019; 139: e56–e528.
3. Marler JR, Tilley BC, Lu M, *et al.* Early stroke treatment associated with better outcome: the NINDS rt-PA stroke study. *Neurology* 2000; 55: 1649–1655.
4. Advani R, Naess H and Kurz MW. The golden hour of acute ischemic stroke. *Scand J Trauma Resusc Emerg Med* 2017; 25: 54.
5. Tarnutzer AA, Lee SH, Robinson KA, *et al.* ED misdiagnosis of cerebrovascular events in the era of modern neuroimaging: a meta-analysis. *Neurology* 2017; 88: 1468–1477.
6. Gurley KL and Edlow JA. Avoiding misdiagnosis in patients with posterior circulation ischemia: a narrative review. *Acad Emerg Med* 2019; 26: 1273–1284.
7. Kuruvilla A, Bhattacharya P, Rajamani K, *et al.* Factors associated with misdiagnosis of acute stroke in young adults. *J Stroke Cerebrovasc Dis* 2011; 20: 523–527.
8. Newman-toker DE, Moy E, Valente E, *et al.* Missed diagnosis of stroke in the emergency department : a cross-sectional analysis of a large. *Diagnosis (Berl)* 2014; 1: 155–166.
9. Fordyce J, Blank FSJ, Pekow P, *et al.* Errors in a busy emergency department. *Ann Emerg Med* 2003; 42: 324–333.
10. Edlow JA, Newman-Toker DE and Savitz SI. Diagnosis and initial management of cerebellar infarction. *Lancet Neurol* 2008; 7: 951–964.
11. Caplan LR and Adelman L. Neurologic education. *West J Med* 1994; 161: 319–322.
12. Hansen CK, Fisher J, Joyce N, *et al.* Emergency department consultations for patients with neurological emergencies. *Eur J Neurol* 2011; 18: 1317–1322.

13. Arch AE, Weisman DC, Coca S, *et al.* Missed ischemic stroke diagnosis in the emergency department by emergency medicine and neurology services. *Stroke* 2016; 47: 668–673.
14. Schrock JW, Glasenapp M, Victor A, *et al.* Variables associated with discordance between emergency physician and neurologist diagnoses of transient ischemic attacks in the emergency department. *Ann Emerg Med* 2012; 59: 19–26.
15. Mohammad YM. Mode of arrival to the emergency department of stroke patients in the United States. *J Vasc Interv Neurol* 2008; 1: 83–86.
16. Tennyson JC, Michael SS, Youngren MN, *et al.* Delayed recognition of acute stroke by emergency department staff following failure to activate stroke by emergency medical services. *West J Emerg Med* 2019; 20: 342–350.
17. Moulin T, Sablot D, Vidry E, *et al.* Impact of emergency room neurologists on patient management and outcome. *Eur Neurol* 2003; 50: 207–214.
18. Falco FA, Sterzi R, Toso V, *et al.* The neurologist in the emergency department. An Italian nationwide epidemiological survey. *Neurol Sci* 2008; 29: 67–75.
19. Morrison I, Jamdar R, Shah P, *et al.* Neurology liaison services in the acute medical receiving unit. *Scott Med J* 2013; 58: 234–236.
20. Antipova D, Eadie L, Macaden A, *et al.* Diagnostic accuracy of clinical tools for assessment of acute stroke: a systematic review. *BMC Emerg Med* 2019; 19: 49.
21. Chaudhary D, Abedi V, Li J, *et al.* Clinical risk score for predicting recurrence following a cerebral ischemic event. *Front Neurol* 2019; 10: 1106.
22. Krittanawong C, Zhang HJ, Wang Z, *et al.* Artificial intelligence in precision cardiovascular medicine. *J Am Coll Cardiol* 2017; 69: 2657–2664.
23. Noorbakhsh-Sabet N, Zand R, Zhang Y, *et al.* Artificial intelligence transforms the future of health care. *Am J Med* 2019; 132: 795–801.
24. Lee EJ, Kim YH, Kim N, *et al.* Deep into the brain: artificial intelligence in stroke imaging. *J Stroke* 2017; 19: 277–285.
25. Zuick S, Graustein A, Urbani R, *et al.* Can a computerized sepsis screening and alert system accurately diagnose sepsis in hospitalized floor patients and potentially provide opportunities for early intervention? A pilot study. *J Intensive Crit Care* 2016; 2.
26. Patel YR, Robbins JM, Kurgansky KE, *et al.* Development and validation of a heart failure with preserved ejection fraction cohort using electronic medical records. *BMC Cardiovasc Disord* 2018; 18: 128.
27. Collins GS, Reitsma JB, Altman DG, *et al.* Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. *BMJ* 2015; 350: g7594.
28. White IR, Royston P, Wood AM, *et al.* Multiple imputation using chained equations: issues and guidance for practice. *Stat Med* 2011; 30: 377–399.
29. Abedi V, Shivakumar MK, Lu P, *et al.* Latent-based imputation of laboratory measures from electronic health records: case for complex diseases. *BioRxiv* 2018; 275743.
30. National Library of Medicine. Unified medical language system (UMLS). <https://www.nlm.nih.gov/research/umls/index.html> (2019, accessed 10 February 2020).
31. Chawla NV, Bowyer KW, Hall LO, *et al.* SMOTE: synthetic minority over-sampling technique. *J Artif Intell Res* 2002; 16: 321–357.
32. ER Inspector. <https://projects.propublica.org/emergency/> (accessed 1 December 2019).
33. Kabra R, Robbie H and Connor SEJ. Diagnostic yield and impact of MRI for acute ischaemic stroke in patients presenting with dizziness and vertigo. *Clin Radiol* 2015; 70: 736–742.
34. Chaturvedi S, Ofner S, Baye F, *et al.* Have clinicians adopted the use of brain MRI for patients with TIA and minor stroke? *Neurology* 2017; 88: 237–244.
35. Varonen H, Kortteisto T and Kaila M. What may help or hinder the implementation of computerized decision support systems (CDSSs): a focus group study with physicians. *Fam Pract* 2008; 25: 162–167.
36. Bouaud J, Spano JP, Lefranc JP, *et al.* Physicians' attitudes towards the advice of a guideline-based decision support system: a case study with OncoDoc2 in the management of breast cancer patients. *Stud Health Technol Inform* 2015; 216: 264–269.
37. Van Der Sijs H, Mulder A, Van Gelder T, *et al.* Drug safety alert generation and overriding in a large Dutch university medical centre. *Pharmacoepidemiol Drug Saf* 2009; 18: 941–947.
38. Van Der Sijs H, Aarts J, Vulto A, *et al.* Overriding of drug safety alerts in computerized physician order entry. *J Am Med Informatics Assoc* 2006; 13: 138–147.

39. Bergman LG and Fors UGH. Computer-aided DSM-IV-diagnostics - Acceptance, use and perceived usefulness in relation to users' learning styles. *BMC Med Inform Decis Mak* 2005; 5: 1.
40. Curry L and Reed MH. Electronic decision support for diagnostic imaging in a primary care setting. *J Am Med Informatics Assoc* 2011; 18: 267–270.
41. Zheng K, Padman R, Johnson MP, *et al.* Understanding technology adoption in clinical care: clinician adoption behavior of a point-of-care reminder system. *Int J Med Inform* 2005; 74: 535–543.
42. Rousseau N, McColl E, Newton J, *et al.* Practice based, longitudinal, qualitative interview study of computerised evidence based guidelines in primary care. *Br Med J* 2003; 326: 314–318.
43. Johnson MP, Zheng K and Padman R. Modeling the longitudinality of user acceptance of technology with an evidence-adaptive clinical decision support system. *Decis Support Syst* 2014; 57: 444–453.
44. Sousa VEC, Lopez KD, Febretti A, *et al.* Use of simulation to study nurses' acceptance and nonacceptance of clinical decision support suggestions. *Comput Inform Nurs* 2015; 33: 465–472.
45. Terraz O, Wietlisbach V, Jeannot JG, *et al.* The EPAGE internet guideline as a decision support tool for determining the appropriateness of colonoscopy. *Digestion* 2005; 71: 72–77.
46. Gadd CS, Baskaran P and Lobach DF. Identification of design features to enhance utilization and acceptance of systems for internet-based decision support at the point of care. *Proc AMLA Symp* 1998; 91–95.
47. Khalifa M. Clinical decision support: strategies for success. *Procedia Comput Sci* 2014; 37: 422–427.
48. McCoy AB, Thomas EJ, Krousel-Wood M, *et al.* Clinical decision support alert appropriateness: a review and proposal for improvement. *Ochsner J* 2014; 14: 195–202.
49. Aakre CA, Dziadzko MA and Herasevich V. Towards automated calculation of evidence-based clinical scores. *World J Methodol* 2017; 7: 16–27.
50. Khairat S, Marc D, Crosby W, *et al.* Reasons for physicians not adopting clinical decision support systems: critical analysis. *JMIR Med Inform* 2018; 6: e24.
51. Venkatesh V, Morris MG, Davis GB, *et al.* User acceptance of information technology: toward a unified view. *MIS Q* 2003; 27: 425–478.
52. Lao Q, Jiang X, Havaei M, *et al.* Continuous domain adaptation with variational domain-agnostic feature replay, <http://arxiv.org/abs/2003.04382>. (2020, 9 March 2020)
53. Wiens J, Guttag J and Horvitz E. A study in transfer learning: leveraging data from multiple hospitals to enhance hospital-specific predictions. *J Am Med Informatics Assoc* 2014; 21: 699–706.