

Artificial Intelligence for Clinical Decision Support in Acute Ischemic Stroke Care: A Systematic Review

Adam Hilbert (✉ adam.hilbert@charite.de)

Charité Lab for Artificial Intelligence in Medicine (CLAIM), Charité Universitätsmedizin Berlin <https://orcid.org/0000-0003-3447-5453>

Ela M. Akay

Charité Lab for Artificial Intelligence in Medicine (CLAIM), Charité Universitätsmedizin Berlin

Benjamin G. Carlisle

QUEST Center for Responsible Research, Berlin Institute of Health (BIH), Charité Universitätsmedizin Berlin <https://orcid.org/0000-0001-8975-0649>

Vince I. Madai

QUEST Center for Responsible Research, Berlin Institute of Health (BIH), Charité Universitätsmedizin Berlin <https://orcid.org/0000-0002-8552-6954>

Matthias A. Mutke

Heidelberg University Hospital

Dietmar Frey

Charité Lab for Artificial Intelligence in Medicine (CLAIM), Charité Universitätsmedizin Berlin <https://orcid.org/0000-0001-5407-2331>

Article

Keywords: Artificial Intelligence, Acute Ischemic Stroke, Machine Learning, Decision Support, Computer Aided

Posted Date: June 3rd, 2022

DOI: <https://doi.org/10.21203/rs.3.rs-1706474/v3>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Abstract

Established randomized-trial-based parameters for acute ischemic stroke care fail to consider individualized patient data, leading to attempts to support or automate treatment and diagnosis decisions using artificial intelligence methods. We review existing research, specifically regarding methodological robustness, thereby identifying constraints for clinical AI implementation. Our systematic review of clinical decision support systems (CDSS) includes full-text English language publications proposing AI-based methods for decision support in acute ischemic stroke cases in adult patients. We (a) describe data and outcomes used in those systems, (b) estimate the systems' benefits compared to traditional stroke diagnosis and treatment, and (c) report concordance with the MINIMAR checklist. 121 studies met our inclusion criteria. 65 were included for full extraction. In our sample, utilized data sources, methods, and reporting practices were highly heterogeneous; adherence to the MINIMAR checklist was low. Our results suggest significant validity threats, dissonance in reporting practices and challenges to clinical translation. We outline practical recommendations for successful implementation of AI in acute ischemic stroke treatment and diagnosis.

Ela M. Akay and Adam Hilbert contributed equally

Introduction

Ischemic stroke is a leading cause of death and disability worldwide and, without effective diagnostic and treatment strategies, its burden is expected to increase [1]. Evidence for treatment decisions in patients with acute ischemic stroke is based on large prospective randomized trials, in which, for example, time from symptom onset or a so-called imaging mismatch on magnetic resonance imaging have emerged established parameters. However, those parameters apply thresholds that are population-based and not individualized [2]–[4]. The rise of artificial intelligence (AI) methods and their application in other areas of medicine has inspired an attempt to revolutionize stroke care by the use of intelligent, individualized, data-driven decision aids.

AI can be used primarily to provide algorithm-based decision aids that can give additional information or guidance to the physician assessing a particular case [5]. AI based clinical decision support systems (CDSS) have already been developed for the diagnosis of ischemic stroke and are commercially available. Most of them aim to automate subtasks such as the calculation of the Alberta Stroke Program Early CT Score (ASPECTS) [6]–[9] or the identification of ischemic lesion biomarkers on imaging [10]–[12]. However, such algorithm-based, individualized solutions do not yet exist for more complex decisions such as treatment stratification or outcome prediction. These CDSSs are expected to lead to more efficient workflows and to ultimately improve outcomes in clinical practice [5] and could therefore offer great potential for both patients and clinicians.

In developing AI decision support tools, researchers face numerous challenges related to regulatory approval and prospective clinical validation requirements [13]–[15] reducing the number of available CDSSs in clinical use today. For this reason, there are few standards for research and reporting on these AI applications, and knowledge about research in this field is scattered and unstructured leading to a low evidence-level of the available literature. Some narrative reviews on AI-based solutions in ischemic stroke [2], [15]–[18] exist. However, these are limited to certain data modalities [19]–[23], target only diagnosis and specific cases of stroke [24], [25], or focus solely on interventions [26].

This systematic review analyzes the wide range of AI applied in stroke decision support, without restrictions to populations, interventions or data modalities. It is primarily aimed at clinicians and researchers in the field of stroke who are seeking an overview of CDSSs using AI. It aims to provide an assessment of the benefits and limitations of these methods and to identify potential for further research. Additionally, it seeks to enable the reader to assess the benefits of AI application compared to the clinical status quo and to evaluate which patients will profit most from these novel approaches. We also aim at informing scientists from technical fields developing AI methods for clinical use about the current methodological quality and clinical applicability of different methods in order to determine the level of standardization in this field and guide future research toward best practices.

Methods

Outcome measures

The primary aim of this study is to generate an overview of different decision support methods using artificial intelligence for acute ischemic stroke care. We wanted to define more closely what data and what outcome measures were used for the different models. Our secondary goals are to estimate the benefit of AI use in stroke care in comparison to conventional time and imaging-based decision-making, assess

which patients profit most from AI-based decision-making and assess the clinical applicability of methods. We were also interested in assessing the current methodological quality among different approaches as well as how many of the papers directly compare an AI method to a traditional clinical method.

Literature search, inclusion criteria

Trials were captured by searching Embase, Medline, ArXiv, BiorXiv, MedrXiv and Clinicaltrials.gov on the 17th and 19th of May 2021 for trials using the terms “stroke,” “cerebrovascular accident,” and variations on these terms and combined results with the terms “artificial intelligence”, “machine learning” and similar methods (Full search strategy available in Supplementary Materials online).

Publications were then screened by EA, AH and BC for the following inclusion criteria: full-text publication, English language, human research subjects, using an artificial intelligence method (e.g. machine learning, deep learning, support vector machines, etc), adult patients, acute ischemic stroke, proposing a method to be used in decision support in the acute setting. While meeting our inclusion criteria, papers proposing methods for automated stroke scoring and for stroke lesion segmentation were excluded from full extraction. These methods aim to automate the implementation of a theoretical concept, or diagnostic score usually completed by a human. This means their output does not imply specific decisions and does not directly support prognostication of treatment or clinical outcome. Consequently, the requirements for these systems differ substantially from those supporting decisions directly and thus their full text was not analyzed for information extraction. Data from each paper was extracted by two researchers. In case of disagreements the paper was discussed with all three investigators until a unanimous agreement was found.

Reporting guidelines

As a post-hoc analysis, we also report each trial's concordance with the MINimum Information for Medical AI Reporting (MINIMAR) checklist [27]. MINIMAR was designed to standardize reporting on artificial intelligence in medicine and thereby ensure generalizability as well as the documentation of potential biases. The full list of criteria can be found in Table 1. We made a concrete specification of the relevant criteria to clarify their representation in our data and deemed criteria that were not represented in our extraction out-of-scope. The cohort selection was defined as “Stroke subgroups” including elderly patients, first-time stroke patients, anterior circulation stroke patients and different treatment groups. The gold standard in our case was defined as “clinical comparator”, which was defined as a method used for the same classification task in a clinical setting e.g., a clinical score or a human rating. Model task was defined as decision support pertaining to the present (classification) or the future (prediction).

Extraction

We devised an extraction template (see Supplementary material 1) that captured variables in the following domains: AI technique, patient characteristics, dataset specifications, validation method, outcome endpoint, results, clinical comparator. Criteria for extraction were prespecified in a codebook, and coders underwent training before data collection.

Data synthesis

We performed data extraction with Numbat Systematic Review Manager v. 2.13 (RRID:SCR_019207).

This study was not subject to Institutional Review Board approval, as it relies on publicly available data; no research participants were involved. This study was prospectively registered on the Open Science Framework. The code and data sets used in preparation of this manuscript are available online (<https://osf.io/x5mb3/>).

Results

A total of 121 studies met our inclusion criteria. Among these, 65 were included for full extraction while 20 studies were identified as proposing a method for automated stroke scoring and 36 for stroke lesion segmentation in imaging. An overview of the papers proposing a

lesion segmentation method can be found in the supplementary material (see Supplementary material 2). Among the 20 papers on automated stroke scoring, 18 were automated calculations of the Alberta Stroke Program Early CT Score (ASPECTS) and 2 reported methods for automated collateral score calculation. The method most used among automated stroke scoring studies was e-ASPECTS. Details on the papers proposing a stroke score automation method can be found in the supplementary material (see Supplementary material 3).

The full table of the extracted information can be found in the supplementary material (see Supplementary material 5).

Reporting guidelines

Summary about the number of studies meeting the specific MINIMAR criteria can be found in Table 1. In general, there was moderate adherence to the included guidelines with some exceptions worth highlighting. 2 studies out of 65 (3%) reported race, ethnicity or socioeconomic status, 52 (80%) studies reported sex and 54 (83%) studies age. Within the reviewed articles, 17 (26%) works reported the use of a hyper-parameter tuning algorithm, namely Grid Search or Random Search algorithm with 13/17 (76%) and 4/17 (24%) applications respectively.

Patient cohorts

All extracted studies provided information on the patient population used. The median sample size in our sample was 220 patients, with a range from 4 patients whose imaging was processed voxel-wise [28] to 2604 patients [29]. We recognized a trend of a growing number of patients included in the studies throughout the years. An overview of the number of patients included in studies per year can be seen in Fig. 3. Numbers shown reflect mean and median numbers considering the number of studies from the specific years. 50 (77%) of the studies specified stroke subgroups (See Fig2 A). The most common groups specified were treatment groups with 20 studies (31%) including patients treated with mechanical thrombectomy and 9 (14%) with tissue-plasminogen activator. 19 (29%) of the studies included only patients with stroke within the anterior circulation (anterior and middle cerebral artery) and 7 (11%) included patients with strokes within the territory of the middle cerebral artery only.

We also extracted the demographic data reported. 8 studies (12%) reported no demographic data. Most studies only include information on age (54/65, 83%) and sex (52/65, 80%). Race was reported by 2 (3%) out of 65 studies while ethnicity and socioeconomic status were not reported by any of the included studies.

Data used in AI models

For the extraction of the data used as features for the artificial intelligence model we differentiated between MRI, CT, other imaging, and clinical data. 23 (35%) of 65 studies used both image information and clinical data for their proposed method. Raw imaging data was used by 22 (34%) for MRI and 10 (15%) for CT while 10 (15%) studies relied only on clinical data. The clinical data used varies greatly within the dataset of extracted studies. Some used few clinical features in combination with an imaging technique while others relied on a high volume of these data points. The highest number of features was used by Kappelhof et al. with 63 clinical features in combination with CT imaging [30]. The most common clinical features used were "Age" (32/65, 49%), "National Institutes of Health Stroke Scale" (NIHSS, 31/65, 48%), and "Sex" (23/65, 35%). 2 papers (3%) gave no further specification of what patient characteristics were used for the model.

Model technique

The reported outcome endpoints can be seen in Fig2 B. Almost a third of the papers (21/65, 32%) proposed methods for the dichotomized prediction of the modified Rankin Scale at 90 days after stroke. Final infarct prediction (17/65, 26%) and infarct core mapping (17/65, 26%) were the second most common outcome endpoint. Successful treatment by mechanical thrombectomy or thrombolysis were predicted by 2 studies (3%) each. A clinical comparator was reported in 24 studies (37%) of which 20 were automated comparators and four compared the proposed method to a human reading of the data. An overview of artificial intelligence methods used can be seen in Fig2 C. Most commonly used techniques were Convolutional Neural Networks (CNNs, 17/65, 26%) and Random Forest algorithms (11/65, 17%). Data splitting into training set for initial model fitting and test set for independent evaluation was reported for slightly more than half of the studies extracted (38/65, 58%).

Optimization, validation and outcome measures

We found considerable variance and lack of reporting of standard machine learning methods concerning optimization, validation and outcome measures. Internal model validation was reported by 61 (94%) papers. External model validation with an independent or hold-out test set was reported by 38 (58%) studies. The most used performance measurement was AUC, used in 54 (83%) of our samples. A clinical

comparator as a gold standard was reported by 24 (37%) papers. The comparator was outperformed by the model in 18/24 cases (75%), 2/24 (8%) reported a worse performance by the model and in 4/24 (17%) a direct comparison was not obviously determinable. Individual results for each study as an overview can be found in Tables 2, 3 and 4. For details see the supplementary material (see Supplementary material 4).

Criterion	Specification	studies that report this criterion	
1. Study population and setting			
Population	Number of patients	65/65	100%
Study setting	Out-of-scope	----	----
Data source	Out-of-scope	----	----
Cohort selection	Stroke subgroups	50/65	77%
2. Patient demographic characteristics			
Age	+	54/65	83%
Sex	+	52/65	80%
Race	+	2/65	3%
Ethnicity	+	0/65	----
Socioeconomic status	+	0/65	----
3. Model architecture			
Model output	+	65/65	100%
Target user	Out-of-scope (Clinician)	----	----
Data splitting	Internal/External model validation	----	----
Gold standard	Clinical comparator	24/65	37%
Model task	Classification/prediction (future or present decision support)	---	----
Model architecture	+	64/65	98%
Features	+	65/65	100%
Missingness	Out-of-scope	----	----
4. Model evaluation			
Optimization	+	17/65	26%
Internal model validation	+	61/65	94%
External model validation	+	38/65	58%
Transparency	Out-of-scope	----	----

Table 1. MINIMAR criteria and number of studies that meet them

Citation	Outcome Endpoint	Patient Subgroup	Number of patients	AI Technique	Clinical	Train/Test Reporting	AUC	Dice	Comparator
[31]	discharge mortality	Not Specified	229	SVM	6	Absolute n	x	0.5	None
[32]	90 day mRS	tPA	1984	Logistic Regression	5	Absolute n	0.786	x	Auto
[33]	90 day mRS	tPA	425	Random Forest	49	Not reported	0.808	x	Auto
[29]	90 day mRS	No recanalization therapy (not specified)	2604	Multilayer Perceptrons	38	Not reported	0.888	x	Auto
[34]	sICH, 90-day-mortality	tPA	331	Multilayer Perceptrons	5/6	Not reported	siCH: 0.941 Mort: 0.976	x	Auto
[35, p. 202]	Successful thrombolysis, 90 day mRS	Elderly, <3h	80	Multilayer Perceptrons	9	Not reported	0.974	x	None
[36]	sICH	tPA	2237	Multilayer Perceptrons	5	Ratio	0.82	x	None
[37]	Post-stroke pneumonia	Not Specified	3160	XGBoost	6	Ratio	0.841	x	Auto
[38]	90 day mRS	Not Specified	1121	SVM	14	Absolute n	0.71	x	None
[39]	90 day mRS	Supratentorial	314	Multilayer Perceptrons	7	Ratio	0.83	x	None
[40]	Six-month mRS	Not Specified	1735	Random Forest	21	Ratio	0.874	x	Auto

Table 2. Overview of studies using clinical data (AUC=Area Under the Curve, mRS=modified Rankin Scale, sICH=symptomatic intracranial hemorrhage, tPA=tissue Plasminogen Activator, SVM=Support Vector Machine, Absolute n= Absolute number of patients, Auto=Automated)

Citation	Outcome Endpoint	Patient Sub-group	N° of patients	AI Technique	Imaging	Train/ Test Reporting	AUC	Dice	Comparator
[41]	Final infarct	Not Specified	12	Multilayer Perceptrons	MRI: T1, T2, DWI, Proton-density WI	Not reported	0.89	x	None
[42]	Onset time	MCA	105	Stepwise Multi-linear Regression	MRI: DWI, ADC, FLAIR, PWI	Not reported	0.683	0.765	None
[43]	Final infarct w or w/o reperfusion	MCA	80	Random Forest	MRI: DWI, ADC, GRE, DSC-enhanced perfusion MRI, T2, Gd-MRA, T1C, TOF angiography	Absolute n	Positive: 0.94 Negative: 0.96	x	Auto
[44]	sICH	< 6h	155	Kernel spectral regression	MRI: PWI, DWI	Not reported	0.837	0.717	None
[45]	Final infarct	Not Specified	170	XGBoost	MRI: DSC-PWI, T2-FLAIR, DWI	Not reported	0.92	x	None
[46]	Final infarct	tPA	222	CNN	MRI: PWI, T2-FLAIR, DWI	Ratio, absolute n	0.88	x	Auto
[28]	Final infarct	Anterior	4	SVM	MRI: PWI	Not reported	x	x	None
[47]	Final infarct	Thrombectomy	29	CNN	CTP	Ratio	x	0.43	None
[48]	Onset time	MCA	131	Logistic Regression	MRI: DWI, ADC, FLAIR, PWI perfusion parameters	Ratio	0.765	0.788	Human
[49]	Final infarct	Anterior, <6/12h, tPA/conservative treatment	55	Adaptive boosting	MRI: T2 FLAIR, DWI, ADC, PWI, CBV, CBF, TTP, MTT, TMAX	Absolute n	0.88	0.28	None
[50]	Final infarct	MCA	48	CNN	MRI: FLAIR, PWI	Not reported	0.871	0.347	None
[51]	LVO + Infarct Core mapping	Anterior	224	CNN	CTA	Not reported	30mL: 0.88 50mL: 0.90 LVO: 0.844	x	None
[52]	Successful thrombolysis	tPA, ICA/M1 MCA	67	SVM	NCCT, CTA, manually extracted thrombus	Not reported	0.85	x	None
[53]	Successful recanalization, 90 day mRS	Thrombectomy	1301	CNN	CTA	Not reported	mTICI: 0.65 mRS: 0.71	x	None
[54]	LVO	Not Specified	42	Unknown	Headpulse from cranial accelerometer, electrocardiogram outputs	Not reported	0.79	x	None
[55]	Final infarct	Not Specified	284	Random Forest	Multiphase CTA	Absolute n	x	0.447	Auto
[56]	LVO	Not Specified	540	CNN	Multiphase CTA	Ratio	0.89	x	None
[57]	LVO	Anterior	584	CNN	NCCT, CTA, 4D-CTA, derived perfusion maps	Absolute n	0.98	x	None
[58]	Penumbral Tissue Mapping	Thrombectomy, Anterior	149	CNN	MRI: 3D pCASL, DWI	Ratio	0.959	0.47	None
[59]	Final infarct	Thrombectomy, Anterior	182	CNN	MRI: DWI, PWI, ADC, Tmax, CBF,	Ratio	0.92	0.53	Auto

					CBV, MTT				
[60]	Infarct core mapping	Thrombectomy, successful reca	25	SVM	DSA (a.p., lateral)	Absolute n	0.904	x	None
[61]	First time recanalization, num passages	Thrombectomy	136	SVM	NCCT, CTA	Absolute n	1st pass: 0.88	x	None
[62]	90 day mRS	Thrombectomy, Anterior	324	CNN	MRI: DWI	Absolute n	0.73	x	Auto
[63]	Final infarct	First time, Anterior, NIHSS > 4, <12hrs	99	XGBoost	MRI: DWI, PWI	Not reported	0.893	0.387	None
[64]	Onset time	Not Specified	355	Random Forest	MRI: DWI, FLAIR, infarct segmentation	Not reported	0.851	x	Human
[65]	Final infarct w/ and w/o successful recanalization	MCA, Thrombectomy	92	Random Forest	MRI: DWI, PWI, ADC	Absolute n	x	0.49	None
[66]	90 day mRS	First time	1840	CNN	MRI Radiology Reports	Ratio	0.805	x	None
[67]	Final infarct	Anterior, <12h	99	Logistic Regression	MRI: DWI, PWI	Not reported	0.872	0.348	None
[68]	Onset time	Not Specified	422	CNN	MRI: FLAIR, DWI, ADC, T2	Ratio	0.74	x	Human
[69]	Final infarct	Thrombectomy	75	Restricted Boltzmann Machines, CNN	MRI: ADC, MTT, TTP, Tmax, rCBF, rCBV	Absolute n	x	0.38	None
[70]	Final infarct	Thrombectomy	109	CNN	MRI: DWI, FLAIR, PWI, CBF, CBV, MTT, Tmax, TTP	Not reported	Reperfused: 0.87 non-reperf: 0.81	Reperfused:0.43 non-reperf: 0.44	Auto
[71]	Infarct Core Mapping	Thrombectomy, Anterior	103	CNN	Dynamic CTP, perfusion maps (RAPID)	Absolute n	x	0.51	Auto
[72]	Tissue at risk, ischemic core	Not Specified	237	CNN	MRI: DWI, ADC, Tmax, MTT, CBF, CBV, thresholded masks	Ratio	TaR: 0.92 Core: 0.94	TaR: 0.60 core: 0.57	Auto

Table 3. Overview of studies using image information(AUC=Area Under the Curve, mRS=modified Rankin Scale, sICH=symptomatic intracranial hemorrhage, LVO=large vessel occlusion, tPA=tissue Plasminogen Activator, SVM=Support Vector Machine, CNN=Convolutional Neural Network, MRI=Magnetic Resonance Imaging, CT=Computer Tomography, CTA=CT Angiography, NCCT=Non-Contrast-CT, CTP=CT perfusion, DWI=Diffusion-weighted imaging, ADC=apparent diffusion coefficient, FLAIR= Fluid-attenuated inversion recovery, PWI=Perfusion weighted imaging, CBF=Cerebral Blood Flow, CBV=Cerebral Blood Volume, MTT=Mean Transit Time, Tmax=Time to maximum, pcASL= pseudocontinuous Arterial Spin Labeling, DSA=Digital Substraction Angiography, GRE=Gradient Echo Imaging, WI=Weighted Imaging, DSC= Dynamic susceptibility contrast imaging, Gd=Gadolinium, TOF=Time-of-flight, Absolute n= Absolute number of patients, mTICI= modified treatment in cerebral infarction score, TaR=Tissue at Risk, Auto=Automated)

Citation	Outcome Endpoint	Patient Subgroup	N° of patients	AI Technique	Imaging	Clinical	Train/ Test Reporting	AUC	Dice	Comparator
[73]	sICH	tPA	194	Probabilistic NN	CT findings, ASPECTS	27	Ratio	0.788	0.522	None
[74]	sICH	tPA	116	SVM	NCCT	1	Not reported	0.744	x	Auto
[75]	Successful recanalization, 90 day mRS	Thrombectomy, Anterior	1383	Random Forest	NCCT, ASPECTSCTA	27	Ratio	mTICI: 0.55 mRS: 0.79	x	None
[76]	Infarct core mapping	Anterior, <8h	128	Multilayer Perceptrons	CTP	4	Ratio	0.87	0.43	None
[77]	Final infarct	Thrombectomy, tPA, distal ICA/M1 of MCA	100	Random Forest	MRI: PWI, DWI	4	Not reported	x	0.464	None
[78]	90 day mRS	Anterior	512	Gradient Boosting Machines	NCCT, CTA, CTP, ASPECTS	3	Not reported	0.748	x	None
[79]	90 day mRS	Thrombectomy, elderly	146	M5P - regression decision tree	ASPECTS	9	Absolute n	x	x	None
[80]	LVO	Not Specified	300	XGBoost	NCCT	Not specified	Absolute n	0.847	0.804	None
[81]	Final infarct	Thrombectomy, Anterior, <=6h	188	CNN	CTP, manual AIF	1	Not reported	0.54 (PR)	0.47	None
[82]	Edema	MCA	116	Random Forest	NCCT	3	Not reported	0.96	0.91	None
[83]	90 day mRS, NIHSS 24h	Not Specified	204	CNN	NCCT	Not specified	Ratio	mRS: 0.75 NIHSS: 0.70	mRS: 0.69 NIHSS: 0.74	Auto
[84]	90 day mRS	Thrombectomy, Anterior	246	Gradient Boosting	NCCT, CTA, CTP	13	Not reported	0.747	x	None
[85]	90 day mRS	Thrombectomy, Anterior	1526	Multilayer Perceptrons	NCCT, CTA, ASPECTS	32	Ratio	0.81	x	None
[86]	90 day mRS	Thrombectomy, Anterior	502	Regularized Logistic Regression	ASPECTS	15	Absolute n	0.90	x	Auto
[87]	90 day mRS, >=8 point NIHSS improvement at 24h	tPA, age 18-80, NIHSS 4-25	196	Multilayer Perceptrons	MRI/CT	10	Absolute n	x	x	None
[88]	Worsening of NIHSS within 3 days	NIHSS => 3	739	tree boosting	MRI/CT	17	Not reported	0.934	0.8	None
[89]	90 day mRS, in-hospital mortality	Not Specified	3445	Gradient boosting	CT/MRI	49	Ratio	mRS: 0.92 Mort: 0.84	x	Auto
[90]	First time recanalization	Thrombectomy	220	Random Forest	CT/MRI	20	Ratio	0.659	x	None
[30]	90 day mRS, mRS after recanalization	Thrombectomy	1363	Fuzzy Decision Tree	NCCT, CTA, ASPECTS	63	Not reported	x	x	None
[91]	90 day mRS	MCA, Thrombectomy, M1, <6 hrs	222	Random Forest	MRI: PWI, manual ROIs	12	Ratio	0.684	x	Auto
[92]	90 day mRS	Anterior	1431	XGBoost	CTA, CTP	3	Not reported	0.80	x	Auto

Table 4. Overview of studies using clinical and image information (AUC=Area Under the Curve, mRS=modified Rankin Scale, sICH=symptomatic intracranial hemorrhage, LVO=large vessel occlusion, tPA=tissue Plasminogen Activator, SVM=Support Vector Machine, CNN=Convolutional Neural Network, MRI=Magnetic Resonance Imaging, CT=Computer Tomography, CTA=CT Angiography, NCCT=Non-Contrast-CT, CTP=CT perfusion, DWI=Diffusion-weighted imaging, PWI=Perfusion weighted imaging, ROI=Region of Interest, AIF=Arterial Input Function, Absolute n= Absolute number of patients, mTICI= modified treatment in cerebral infarction score, Acc=Accuracy, NPV=Negative Predictive Value, PPV=Positive Predictive Value, Auto=Automated), TaR=Tissue at Risk)

Discussion

While there has been undeniable progress in the performance of AI models, our results suggest significant potential validity threats, dissonance in reporting practices and challenges to clinical translation across the studies reviewed.

Potential validity threats

Transparent, responsible and valid research is particularly important for the impact of AI on clinical decision-making in stroke diagnosis and treatment. Because of this, the MINIMAR criteria have been developed to standardize these criteria among different methods [27]. As a post-hoc analysis, we tested adherence of the reviewed articles to the most relevant points of these reporting guidelines (See Table 1). First, we see a trend of systematically limited description of patient demographic information beyond variables used by the developed model. Only a marginal percentage of the studies reported race, ethnicity or socioeconomic status. This represents a major risk of bias and prevents even the assessment of the external validity of models to extend their use to other demographic groups.

Second, reproducibility is essential, when it comes to further prospective clinical validation or deployment of the proposed models since AI models parameterized and trained differently will realize different solutions for any learning task. Thus, model performance greatly depends on the use of an efficient search mechanism to find the best parameters across a well-defined parameter space; called hyper-parameter tuning. While description of model architecture is well-reported, only a quarter of the works report on hyper-parameter tuning practices. Even though the reported Grid Search and Random Search methods are frequently used and are simple to implement, they are considered the two most basic and naive approaches, potentially limiting performance.

Third, no AI model could be implemented in clinical practice without addressing concerns regarding technical deployment. Above all, model robustness and confidence can be tested *a priori* to simulate real world scenarios of varying quality, distribution and noise in data. The very first step to enable assessment of robustness is to reserve validation sets for model selection and parameter tuning, and report model performance on a hold-out test set. Validation results are naturally overfitted and often cherry-picked to the given parameter combination and optimization process. This means that even though these parameters deliver the best model, performance measured on the validation set does not necessarily reflect performance in a real-world setting. Almost half of the reviewed articles did not describe the use of an additional test set, thus presumably reported validation performances. Moreover, this process is ideally repeated once the test set is detached from model optimization to report average metrics while inducing stochasticity by random splitting of training and test data. Common frameworks in ML exist for such purposes e.g., nested cross-validation [93]. However, we saw only a marginal number of works testing model robustness on multiple, distinct test sets. This implies that the majority of AI models proposed for stroke CDSSs to date would need additional, rigorous testing to assess applicability in clinical practice.

Beyond the MINIMAR criteria, we found a large degree of heterogeneity in reporting practices in the reviewed articles, making comparison between study outcomes difficult. The most used performance measurement was AUC and while it is a robust metric and its use is widespread, reporting performance measured by multiple metrics is advised to facilitate more reliable interpretation. We saw minor alignment across the reviewed studies with many of them failing to report result metrics that allow for obvious interpretation and comparison to other methods. Moreover, due to the routine use of several image-derived biomarkers, studies including these variables as input for their models are prone to neglect the necessity for describing the specific image sequence to be acquired and the desired mechanism to extract the marker when this is indispensable for a model to be integrated into a clinical workflow.

The dependence of AI on sufficient labeled data to yield models with proper generalization and reliable results has been discussed frequently [14], [17], [94]. Due to data privacy concerns, this is a pain-point for all medical AI development. On a positive note, Figure 3 shows the upwards trend of the number of patients included in the studies per publishing year. Data used in the reviewed articles typically originates either from clinical trials or national stroke registries. The rise in use of extensive data collected by the latter holds great promise. While randomized clinical trials give a clear scope for developing models, these more comprehensive databases not only enable model development on more data points but also provide a close sample to general stroke populations.

Clinical translation

To fully exploit the potential impact of AI-based decision support systems, clinical feasibility must be prioritized, and efforts must be aimed at solving real challenges of the clinical workflow.

First, our results suggest that the definition of decision support in stroke literature is ambiguous, exact clinical utility is rarely discussed and is often not considered in the design phase of AI solutions. We identified three main categories: a) Solutions for automated extraction of biomarkers that aid the diagnostic process, b) a subset of these that are included in current operating procedures of treatment selection and c) models for predicting a future outcome. Studies of the first category were not considered for full text review since they merely improve accuracy and speed of image interpretation and have not been shown to augment the range of available information for decision making. However, it's worth noting their significant proportion among all included articles. There seems to be a higher motivation towards lowering the burden of expert interpretation in the current workflow rather than augmenting findings for more accurate and prompt prognosis. With respect to the second category, 14 (22%) papers proposed a model that classified patients according to onset time, the presence of large vessel occlusion or infarct core mapping. These biomarkers have shown ability to stratify patients with respect to likely benefits of treatment, thus serving valuable, in most cases complementary information in treatment decisions. Lastly, 51 (78%) papers focused on prediction of a future outcome or complication such as functional outcome measured by mRS, successful treatment, final infarct volume or intracranial hemorrhage. The prediction of successful treatment in particular could be used to improve specificity of patient stratification for mechanical thrombectomy which is one of the main goals of research in stroke treatment. However, we saw only 4 (6%) papers proposing a method on this specific task. Even though detailed extraction of this exceeded the scope of our study, we want to emphasize the importance of optimizing both sensitivity and specificity of treatment stratification models to enable better understanding and support of treatment allocation as well as opt-out.

Second, an important goal of AI in stroke diagnosis and treatment is to discover as-yet unknown predictors of outcomes in raw imaging. Deep Learning solutions seem particularly suitable for this task, due to their ability to perform feature extraction and classification at the same time. However, we found only a minority of articles that pursued prediction of treatment and functional outcomes using information from image data with most works predicting more indirect outcomes such as final infarct. In particular, there were only 2 (3%) articles that proposed an AI method to extract image information alone for outcome prediction. The best performing - and only - model predicting treatment outcome defined as the dichotomized mTICI score achieved an AUC of 0.65 by a CNN model using CTA imaging [53]. The highest performance of predicting functional outcome defined as the dichotomized mRS score was an AUC of 0.73 by a CNN model using DWI imaging [62], while a slightly lower AUC of 0.71 was achieved on a much bigger cohort using CTA imaging [53]. When combining AI extracted image information and clinical data an AUC of 0.75 was achieved by a combined CNN and neural network model using NCCT imaging, however the presented image only model achieved an AUC of 0.54 alone [83]. Considering the key role of imaging assessment in treatment selection, the performance of the above-mentioned models and exploitation of AI-driven image processing for outcome prognosis is rather limited. Also, reported predictive performance of clinical data and established biomarkers remain superior, implying that to date, Deep Learning methods have failed to meet their expectations in the field of stroke.

Third, to effectively support decision-making in an acute, time-pressured workflow, solutions need to prioritize usability. Tools using data or variables that are not routinely acquired will hamper efficiency as will models that rely on an extensive set of clinical variables or imaging. Close to half of the articles involving clinical parameters employed more than 10 variables, the most being 63 variables additionally to imaging [30]. Also, almost a quarter of papers with imaging involved, relied on the acquisition of at least 3 image sequences. This seems to indicate a lack of interdisciplinary cooperation to work towards standardized requirements of decision support in an acute scenario. We advise future research to focus on a minimal set of data optimized for both performance and utility at the same time. Moreover, interdisciplinary efforts should be made to improve clinical utility of future AI-based decision support systems.

Finally, even though the reviewed articles were mainly concerned with aspects of model development and not deployment, we anticipated more discussion of user interaction. Only a few articles touched on potential use cases for their developed models, and none elaborated on necessary points of user experience to enable their use in a clinical setting.

Implications and recommendations

In what follows, we outline several implications and recommendations based on our findings for the specific stakeholders involved in the application of AI research in acute stroke decision support to the clinic.

First, we caution researchers to better adhere to best practices in model development such as data splitting and hyper-parameter tuning. Evaluation of model performance on hold-out test sets not involved in model training, model selection and hyper-parameter tuning should be

warranted. Modern hyper-parameter tuning approaches such as Bayesian Optimization [95], Hyperband [96], Spectral Analysis [97] or Covariance Matrix Adaptation Evolution Strategy [98] should be evaluated for the given use case and when applicable, favored over classical Grid or Random Search [99] methods. Researchers should involve clinical practitioners in the design process from the early stages of model development to ensure feasibility in daily clinical practice. Researchers must also review, implement and report on relevant trustworthy and reliability considerations such as technical robustness and transparency prescribed by e.g., the EU Ethics Guidelines for Trustworthy AI [100].

Second, we call for journal editors and reviewers to demand adherence to stricter reporting requirements, e.g., MINIMAR. Describing a targeted use case and clinical decision to aid in publications of such work is also of great importance. Here, we recommend that researchers and reviewers in the field use the guidelines and standard operating procedures of the European Stroke Organization (ESO) for specific treatments of ischemic stroke [3], [4].

Third, we encourage funders to prioritize projects that focus on decision support tools with a clear outline of feasible integration into real-life clinical care, including angles of trustworthiness, usability, technical robustness and data governance [101]. Specific consideration of stroke care such as acute, timely predictions, inclusion of neuroimaging and an absolute user-centered approach to effectively unburden medical professionals should be taken into account and evaluated by a multidisciplinary team of AI scientists and engineers, clinicians and ethicists.

Last, we would like to highlight the crucial role of high-quality training and validation data. Here, policymakers have a major role to provide a path for researchers to obtain the necessary plurality of data and to meet the requirements of robust model development and validation.

Limitations

Our study has several limitations. First, our study included only published research articles, making it susceptible to publication bias. Even though unpublished works might shed light on further, novel methods for supporting decisions, practical implementation of AI algorithms in clinical settings must rely on rigorously peer-reviewed solutions. Hence, we do not see unpublished works having an influential effect on the state-of-the-art of AI in stroke decision support. Second, this study was descriptive in nature, where we elaborated on trends and presented the distribution of contributions in the field from some specific angles. However, we did not formally test the translated impact and did not carry out targeted quantitative analyses to corroborate our claims. All the highlighted shortcomings and derived recommendations are based on theoretical interpretations and thus do not reflect the actual impact of practical implementations.

Conclusion

While there have been great advances in growing availability of research data to further medical AI development, the main stress-points in clinical decision-making in stroke have yet to be addressed. Furthermore, there is limited coordination in reporting of artificial intelligence techniques applied to the context of acute ischemic stroke care and best practices of AI model development should be better adopted. If correctly implemented, these approaches making use of the individualized data available while providing additional information to the physician could lead to better patient outcomes in acute ischemic stroke.

Declarations

Disclosures

Adam Hilbert reported receiving personal fees from ai4medicine outside the submitted work. Dr Madai reported receiving personal fees from ai4medicine outside the submitted work. Dr Frey reported receiving grants from the European Commission, reported receiving personal fees from and holding an equity interest in ai4medicine outside the submitted work.

References

- [1] GBD 2015 Neurological Disorders Collaborator Group, "Global, regional, and national burden of neurological disorders during 1990-2015: a systematic analysis for the Global Burden of Disease Study 2015," *Lancet Neurol.*, vol. 16, no. 11, pp. 877–897, Nov. 2017, doi: 10.1016/S1474-4422(17)30299-5.
- [2] K. H. C. Li, A. Jesuthasan, C. Kui, R. Davies, G. Tse, and G. Y. H. Lip, "Acute ischemic stroke management: concepts and controversies. A narrative review," *Expert Rev. Neurother.*, vol. 21, no. 1, pp. 65–79, Jan. 2021, doi: 10.1080/14737175.2021.1836963.

- [3] E. Berge *et al.*, "European Stroke Organisation (ESO) guidelines on intravenous thrombolysis for acute ischaemic stroke," *Eur. Stroke J.*, vol. 6, no. 1, p. I–LXII, Mar. 2021, doi: 10.1177/2396987321989865.
- [4] G. Turc *et al.*, "European Stroke Organisation (ESO) – European Society for Minimally Invasive Neurological Therapy (ESMINT) Guidelines on Mechanical Thrombectomy in Acute Ischaemic Stroke Endorsed by Stroke Alliance for Europe (SAFE)," *Eur. Stroke J.*, vol. 4, no. 1, pp. 6–12, Mar. 2019, doi: 10.1177/2396987319832140.
- [5] T. Jamieson and A. Goldfarb, "Clinical considerations when applying machine learning to decision-support tasks versus automation," *BMJ Qual. Saf.*, vol. 28, no. 10, pp. 778–781, Oct. 2019, doi: 10.1136/bmjqs-2019-009514.
- [6] S. Nagel *et al.*, "e-ASPECTS software is non-inferior to neuroradiologists in applying the ASPECT score to computed tomography scans of acute ischemic stroke patients," *Int. J. Stroke*, vol. 12, no. 6, pp. 615–622, Aug. 2017, doi: 10.1177/1747493016681020.
- [7] W. Brinjikji *et al.*, "e-ASPECTS software improves interobserver agreement and accuracy of interpretation of aspects score," *Interv. Neuroradiol.*, vol. 27, no. 6, pp. 781–787, Dec. 2021, doi: 10.1177/15910199211011861.
- [8] J. Pfaff *et al.*, "e-ASPECTS Correlates with and Is Predictive of Outcome after Mechanical Thrombectomy," *AJNR Am. J. Neuroradiol.*, vol. 38, no. 8, pp. 1594–1599, Aug. 2017, doi: 10.3174/ajnr.A5236.
- [9] C. Maegerlein *et al.*, "Automated Calculation of the Alberta Stroke Program Early CT Score: Feasibility and Reliability," *Radiology*, vol. 291, no. 1, pp. 141–148, Apr. 2019, doi: 10.1148/radiol.2019181228.
- [10] I. Q. Grunwald *et al.*, "Collateral Automation for Triage in Stroke: Evaluating Automated Scoring of Collaterals in Acute Stroke on Computed Tomography Scans," *Cerebrovasc. Dis.*, vol. 47, no. 5–6, pp. 217–222, 2019, doi: 10.1159/000500076.
- [11] E. Kellner *et al.*, "Automated Infarct Core Volumetry Within the Hypoperfused Tissue: Technical Implementation and Evaluation," *J. Comput. Assist. Tomogr.*, vol. 41, no. 4, pp. 515–520, Aug. 2017, doi: 10.1097/RCT.0000000000000570.
- [12] S. Dehkharghani *et al.*, "Performance and Predictive Value of a User-Independent Platform for CT Perfusion Analysis: Threshold-Derived Automated Systems Outperform Examiner-Driven Approaches in Outcome Prediction of Acute Ischemic Stroke," *AJNR Am. J. Neuroradiol.*, vol. 36, no. 8, pp. 1419–1425, Aug. 2015, doi: 10.3174/ajnr.A4363.
- [13] T. M. Leslie-Mazwi and M. H. Lev, "Towards artificial intelligence for clinical stroke care," *Nat. Rev. Neurol.*, vol. 16, no. 1, Art. no. 1, Jan. 2020, doi: 10.1038/s41582-019-0287-9.
- [14] L. Ding, C. Liu, Z. Li, and Y. Wang, "Incorporating Artificial Intelligence Into Stroke Care and Research," *Stroke*, vol. 51, no. 12, pp. e351–e354, Dec. 2020, doi: 10.1161/STROKEAHA.120.031295.
- [15] D. S. Liebeskind, "Artificial intelligence in stroke care: Deep learning or superficial insight?," *EBioMedicine*, vol. 35, pp. 14–15, Sep. 2018, doi: 10.1016/j.ebiom.2018.08.031.
- [16] M. S. Sirsat, E. Fermé, and J. Câmara, "Machine Learning for Brain Stroke: A Review," *J. Stroke Cerebrovasc. Dis.*, vol. 29, no. 10, p. 105162, Oct. 2020, doi: 10.1016/j.jstrokecerebrovasdis.2020.105162.
- [17] R. Feng, M. Badgeley, J. Mocco, and E. K. Oermann, "Deep learning guided stroke management: a review of clinical applications," *J. NeuroInterventional Surg.*, vol. 10, no. 4, pp. 358–362, Apr. 2018, doi: 10.1136/neurintsurg-2017-013355.
- [18] K. Mouridsen, P. Thurner, and G. Zaharchuk, "Artificial Intelligence Applications in Stroke," *Stroke*, vol. 51, no. 8, pp. 2573–2579, Aug. 2020, doi: 10.1161/STROKEAHA.119.027479.
- [19] A. Bivard, L. Churilov, and M. Parsons, "Artificial intelligence for decision support in acute stroke – current roles and potential," *Nat. Rev. Neurol.*, vol. 16, no. 10, pp. 575–585, Oct. 2020, doi: 10.1038/s41582-020-0390-y.
- [20] W. Wang *et al.*, "A systematic review of machine learning models for predicting outcomes of stroke with structured data," *PLOS ONE*, vol. 15, no. 6, p. e0234722, Jun. 2020, doi: 10.1371/journal.pone.0234722.
- [21] J. E. Soun *et al.*, "Artificial Intelligence and Acute Stroke Imaging," *Am. J. Neuroradiol.*, vol. 42, no. 1, pp. 2–11, Jan. 2021, doi: 10.3174/ajnr.A6883.

- [22] J. Shen, X. Li, Y. Li, and B. Wu, "Comparative accuracy of CT perfusion in diagnosing acute ischemic stroke: A systematic review of 27 trials," *PLoS ONE*, vol. 12, no. 5, May 2017, doi: 10.1371/journal.pone.0176622.
- [23] Y. Mokli, J. Pfaff, D. P. dos Santos, C. Herweh, and S. Nagel, "Computer-aided imaging analysis in acute ischemic stroke – background and clinical applications," *Neurol. Res. Pract.*, vol. 1, no. 1, p. 23, Dec. 2019, doi: 10.1186/s42466-019-0028-y.
- [24] R. Karthik, R. Menaka, A. Johnson, and S. Anand, "Neuroimaging and deep learning for brain stroke detection - A review of recent advancements and future prospects," *Comput. Methods Programs Biomed.*, vol. 197, p. 105728, Dec. 2020, doi: 10.1016/j.cmpb.2020.105728.
- [25] N. M. Murray, M. Unberath, G. D. Hager, and F. K. Hui, "Artificial intelligence to diagnose ischemic stroke and identify large vessel occlusions: a systematic review," *J. NeuroInterventional Surg.*, vol. 12, no. 2, pp. 156–164, Feb. 2020, doi: 10.1136/neurintsurg-2019-015135.
- [26] E. Lotan, "Emerging Artificial Intelligence Imaging Applications for Stroke Interventions," *Am. J. Neuroradiol.*, Dec. 2020, doi: 10.3174/ajnr.A6902.
- [27] T. Hernandez-Boussard, S. Bozkurt, J. P. A. Ioannidis, and N. H. Shah, "MINIMAR (MINimum Information for Medical AI Reporting): Developing reporting standards for artificial intelligence in health care," *J. Am. Med. Inform. Assoc.*, vol. 27, no. 12, pp. 2011–2015, Dec. 2020, doi: 10.1093/jamia/ocaa088.
- [28] M. Giacalone *et al.*, "Local spatio-temporal encoding of raw perfusion MRI for the prediction of final lesion in stroke," *Med. Image Anal.*, vol. 50, pp. 117–126, Dec. 2018, doi: 10.1016/j.media.2018.08.008.
- [29] J. Heo, J. G. Yoon, H. Park, Y. D. Kim, H. S. Nam, and J. H. Heo, "Machine Learning–Based Model for Prediction of Outcomes in Acute Stroke," *Stroke*, vol. 50, no. 5, pp. 1263–1265, May 2019, doi: 10.1161/STROKEAHA.118.024293.
- [30] N. Kappelhof *et al.*, "Evolutionary algorithms and decision trees for predicting poor outcome after endovascular treatment for acute ischemic stroke," *Comput. Biol. Med.*, vol. 133, p. 104414, Jun. 2021, doi: 10.1016/j.combiomed.2021.104414.
- [31] K. C. Ho *et al.*, "Predicting Discharge Mortality after Acute Ischemic Stroke Using Balanced Data," *AMIA Annu. Symp. Proc.*, vol. 2014, pp. 1787–1796, Nov. 2014.
- [32] D. J. Seiffge *et al.*, "Simple variables predict miserable outcome after intravenous thrombolysis," *Eur. J. Neurol.*, vol. 21, no. 2, pp. 185–191, 2014, doi: <https://doi.org/10.1111/ene.12254>.
- [33] M. Monteiro *et al.*, "Using Machine Learning to Improve the Prediction of Functional Outcome in Ischemic Stroke Patients," *IEEE/ACM Trans. Comput. Biol. Bioinform.*, vol. 15, no. 6, pp. 1953–1959, Nov. 2018, doi: 10.1109/TCBB.2018.2811471.
- [34] C.-C. Chung, L. Chan, O. A. Bamodu, C.-T. Hong, and H.-W. Chiu, "Artificial neural network based prediction of postthrombolysis intracerebral hemorrhage and death," *Sci. Rep.*, vol. 10, no. 1, p. 20501, Dec. 2020, doi: 10.1038/s41598-020-77546-5.
- [35] C. C. Chung *et al.*, "Artificial neural network-based analysis of the safety and efficacy of thrombolysis for ischemic stroke in older adults in Taiwan," *Neurol. Asia*, vol. 25, no. 2, pp. 109–117, Jun. 2020.
- [36] F. Wang *et al.*, "Personalized risk prediction of symptomatic intracerebral hemorrhage after stroke thrombolysis using a machine-learning model," *Ther. Adv. Neurol. Disord.*, vol. 13, p. 175628642090235, Jan. 2020, doi: 10.1177/1756286420902358.
- [37] X. Li *et al.*, "Using machine learning to predict stroke-associated pneumonia in Chinese acute ischaemic stroke patients," *Eur. J. Neurol.*, vol. 27, no. 8, pp. 1656–1663, 2020, doi: 10.1111/ene.14295.
- [38] S. A. Alaka *et al.*, "Functional Outcome Prediction in Ischemic Stroke: A Comparison of Machine Learning Algorithms and Regression Models," *Front. Neurol.*, vol. 11, p. 889, Aug. 2020, doi: 10.3389/fneur.2020.00889.
- [39] E. Zihni *et al.*, "Opening the black box of artificial intelligence for clinical decision support: A study predicting stroke outcome," *PLOS ONE*, vol. 15, no. 4, p. e0231166, Apr. 2020, doi: 10.1371/journal.pone.0231166.
- [40] X. Li *et al.*, "Predicting 6-Month Unfavorable Outcome of Acute Ischemic Stroke Using Machine Learning," *Front. Neurol.*, vol. 11, p. 539509, Nov. 2020, doi: 10.3389/fneur.2020.539509.

- [41] H. Bagher-Ebadian *et al.*, "Predicting Final Extent of Ischemic Infarction Using Artificial Neural Network Analysis of Multi-Parametric MRI in Patients with Stroke," *PLoS ONE*, vol. 6, no. 8, p. e22626, Aug. 2011, doi: 10.1371/journal.pone.0022626.
- [42] K. C. Ho, W. Speier, S. El-Saden, and C. W. Arnold, "Classifying Acute Ischemic Stroke Onset Time using Deep Imaging Features," *AMIA Annu. Symp. Proc. AMIA Symp.*, vol. 2017, pp. 892–901, 2017.
- [43] R. McKinley *et al.*, "Fully automated stroke tissue estimation using random forest classifiers (FASTER)," *J. Cereb. Blood Flow Metab.*, vol. 37, no. 8, pp. 2728–2741, Aug. 2017, doi: 10.1177/0271678X16674221.
- [44] Y. Yu, D. Guo, M. Lou, D. Liebeskind, and F. Scalzo, "Prediction of Hemorrhagic Transformation Severity in Acute Stroke From Source Perfusion MRI," *IEEE Trans. Biomed. Eng.*, vol. 65, no. 9, pp. 2058–2065, Sep. 2018, doi: 10.1109/TBME.2017.2783241.
- [45] M. Livne, J. K. Boldsen, I. K. Mikkelsen, J. B. Fiebach, J. Sobesky, and K. Mouridsen, "Boosted Tree Model Reforms Multimodal Magnetic Resonance Imaging Infarct Prediction in Acute Stroke," *Stroke*, vol. 49, no. 4, pp. 912–918, Apr. 2018, doi: 10.1161/STROKEAHA.117.019440.
- [46] A. Nielsen, M. B. Hansen, A. Tietze, and K. Mouridsen, "Prediction of Tissue Outcome and Assessment of Treatment Effect in Acute Ischemic Stroke Using Deep Learning," *Stroke*, vol. 49, no. 6, pp. 1394–1401, Jun. 2018, doi: 10.1161/STROKEAHA.117.019740.
- [47] C. Lucas, A. Kemmling, N. Bouteldja, L. F. Aulmann, A. Madany Mamlouk, and M. P. Heinrich, "Learning to Predict Ischemic Stroke Growth on Acute CT Perfusion Data by Interpolating Low-Dimensional Shape Representations," *Front. Neurol.*, vol. 9, p. 989, Nov. 2018, doi: 10.3389/fneur.2018.00989.
- [48] K. C. Ho, W. Speier, H. Zhang, F. Scalzo, S. El-Saden, and C. W. Arnold, "A Machine Learning Approach for Classifying Ischemic Stroke Onset Time From Imaging," *IEEE Trans. Med. Imaging*, vol. 38, no. 7, pp. 1666–1676, Jul. 2019, doi: 10.1109/TMI.2019.2901445.
- [49] C. Tozlu *et al.*, "Comparison of classification methods for tissue outcome after ischaemic stroke," *Eur. J. Neurosci.*, vol. 50, no. 10, pp. 3590–3598, Nov. 2019, doi: 10.1111/ejn.14507.
- [50] K. C. Ho, "Predicting ischemic stroke tissue fate using a deep convolutional neural network on source magnetic resonance perfusion images," *J. Med. Imaging*, vol. 6, no. 02, p. 1, May 2019, doi: 10.1117/1.JMI.6.2.026001.
- [51] S. A. Sheth *et al.*, "Machine Learning-Enabled Automated Determination of Acute Ischemic Core From Computed Tomography Angiography," *Stroke*, vol. 50, no. 11, pp. 3093–3100, Nov. 2019, doi: 10.1161/STROKEAHA.119.026189.
- [52] W. Qiu *et al.*, "Radiomics-Based Intracranial Thrombus Features on CT and CTA Predict Recanalization with Intravenous Alteplase in Patients with Acute Ischemic Stroke," *Am. J. Neuroradiol.*, vol. 40, no. 1, pp. 39–44, Jan. 2019, doi: 10.3174/ajnr.A5918.
- [53] A. Hilbert *et al.*, "Data-efficient deep learning of radiological image data for outcome prediction after endovascular treatment of patients with acute ischemic stroke," *Comput. Biol. Med.*, vol. 115, p. 103516, Dec. 2019, doi: 10.1016/j.combiomed.2019.103516.
- [54] W. S. Smith, K. J. Keenan, and P. A. Lovoi, "A Unique Signature of Cardiac-Induced Cranial Forces During Acute Large Vessel Stroke and Development of a Predictive Model," *Neurocrit. Care*, vol. 33, no. 1, pp. 58–63, Aug. 2020, doi: 10.1007/s12028-019-00845-x.
- [55] W. Qiu *et al.*, "Machine Learning for Detecting Early Infarction in Acute Stroke with Non-Contrast-enhanced CT," *Radiology*, vol. 294, no. 3, pp. 638–644, Jan. 2020, doi: 10.1148/radiol.2020191193.
- [56] M. T. Stib *et al.*, "Detecting Large Vessel Occlusion at Multiphase CT Angiography by Using a Deep Convolutional Neural Network," *Radiology*, vol. 297, no. 3, pp. 640–649, Sep. 2020, doi: 10.1148/radiol.2020200334.
- [57] M. Meijs, F. J. A. Meijer, M. Prokop, B. van Ginneken, and R. Manniesing, "Image-level detection of arterial occlusions in 4D-CTA of acute stroke patients using deep learning," *Med. Image Anal.*, vol. 66, p. 101810, Dec. 2020, doi: 10.1016/j.media.2020.101810.
- [58] Wang Kai *et al.*, "Deep Learning Detection of Penumbra Tissue on Arterial Spin Labeling in Stroke," *Stroke*, vol. 51, no. 2, pp. 489–497, Feb. 2020, doi: 10.1161/STROKEAHA.119.027457.
- [59] Y. Yu *et al.*, "Use of Deep Learning to Predict Final Ischemic Stroke Lesions From Initial Magnetic Resonance Imaging," *JAMA Netw. Open*, vol. 3, no. 3, p. e200772, Mar. 2020, doi: 10.1001/jamanetworkopen.2020.0772.

- [60] R. A. Rava *et al.*, "Performance of angiographic parametric imaging in locating infarct core in large vessel occlusion acute ischemic stroke patients," *J. Med. Imaging*, vol. 7, no. 1, p. 016001, Jan. 2020, doi: 10.1117/1.JMI.7.1.016001.
- [61] Hofmeister Jeremy *et al.*, "Clot-Based Radiomics Predict a Mechanical Thrombectomy Strategy for Successful Recanalization in Acute Ischemic Stroke," *Stroke*, vol. 51, no. 8, pp. 2488–2494, Aug. 2020, doi: 10.1161/STROKEAHA.120.030334.
- [62] H. Nishi *et al.*, "Deep Learning–Derived High-Level Neuroimaging Features Predict Clinical Outcomes for Large Vessel Occlusion," *Stroke*, vol. 51, no. 5, pp. 1484–1492, May 2020, doi: 10.1161/STROKEAHA.119.028101.
- [63] M. Grosser *et al.*, "Improved multi-parametric prediction of tissue outcome in acute ischemic stroke patients using spatial features," *PLOS ONE*, vol. 15, no. 1, p. e0228113, Jan. 2020, doi: 10.1371/journal.pone.0228113.
- [64] H. Lee *et al.*, "Machine Learning Approach to Identify Stroke Within 4.5 Hours," *Stroke*, vol. 51, no. 3, pp. 860–866, Mar. 2020, doi: 10.1161/STROKEAHA.119.027611.
- [65] Y.-C. Kim *et al.*, "Novel Estimation of Penumbra Zone Based on Infarct Growth Using Machine Learning Techniques in Acute Ischemic Stroke," *J. Clin. Med.*, vol. 9, no. 6, p. 1977, Jun. 2020, doi: 10.3390/jcm9061977.
- [66] T. S. Heo *et al.*, "Prediction of Stroke Outcome Using Natural Language Processing-Based Machine Learning of Radiology Report of Brain MRI," *J. Pers. Med.*, vol. 10, no. 4, p. 286, Dec. 2020, doi: 10.3390/jpm10040286.
- [67] M. Grosser *et al.*, "Localized prediction of tissue outcome in acute ischemic stroke patients using diffusion- and perfusion-weighted MRI datasets," *PLOS ONE*, vol. 15, no. 11, p. e0241917, Nov. 2020, doi: 10.1371/journal.pone.0241917.
- [68] H. Zhang *et al.*, "Intra-Domain Task-Adaptive Transfer Learning to Determine Acute Ischemic Stroke Onset Time," *Comput. Med. Imaging Graph.*, vol. 90, p. 101926, Jun. 2021, doi: 10.1016/j.compmedimag.2021.101926.
- [69] A. Pinto *et al.*, "Combining unsupervised and supervised learning for predicting the final stroke lesion," *Med. Image Anal.*, vol. 69, p. 101888, Apr. 2021, doi: 10.1016/j.media.2020.101888.
- [70] N. Debs *et al.*, "Impact of the reperfusion status for predicting the final stroke infarct using deep learning," *NeuroImage Clin.*, vol. 29, p. 102548, 2021, doi: 10.1016/j.nicl.2020.102548.
- [71] A. Hakim *et al.*, "Predicting Infarct Core From Computed Tomography Perfusion in Acute Ischemia With Machine Learning: Lessons From the ISLES Challenge," *Stroke*, vol. 52, no. 7, pp. 2328–2337, Jul. 2021, doi: 10.1161/STROKEAHA.120.030696.
- [72] Y. Yu *et al.*, "Tissue at Risk and Ischemic Core Estimation Using Deep Learning in Acute Stroke," *Am. J. Neuroradiol.*, vol. 42, no. 6, pp. 1030–1037, Jun. 2021, doi: 10.3174/ajnr.A7081.
- [73] P. Dharmasaroja and P. A. Dharmasaroja, "Prediction of intracerebral hemorrhage following thrombolytic therapy for acute ischemic stroke using multiple artificial neural networks," *Neurol. Res.*, vol. 34, no. 2, pp. 120–128, Mar. 2012, doi: 10.1179/1743132811Y.0000000067.
- [74] P. Bentley *et al.*, "Prediction of stroke thrombolysis outcome using CT brain machine learning," *NeuroImage Clin.*, vol. 4, pp. 635–640, 2014, doi: 10.1016/j.nicl.2014.02.003.
- [75] H. J. A. van Os *et al.*, "Predicting Outcome of Endovascular Treatment for Acute Ischemic Stroke: Potential Value of Machine Learning Algorithms," *Front. Neurol.*, vol. 9, p. 784, Sep. 2018, doi: 10.3389/fneur.2018.00784.
- [76] A. S. Kasasbeh, S. Christensen, M. W. Parsons, B. Campbell, G. W. Albers, and M. G. Lansberg, "Artificial Neural Network Computer Tomography Perfusion Prediction of Ischemic Core," *Stroke*, vol. 50, no. 6, pp. 1578–1581, Jun. 2019, doi: 10.1161/STROKEAHA.118.022649.
- [77] A. J. Winder, S. Siemonsen, F. Flottmann, G. Thomalla, J. Fiehler, and N. D. Forkert, "Technical considerations of multi-parametric tissue outcome prediction methods in acute ischemic stroke patients," *Sci. Rep.*, vol. 9, no. 1, p. 13208, Sep. 2019, doi: 10.1038/s41598-019-49460-y.
- [78] Y. Xie *et al.*, "Use of Gradient Boosting Machine Learning to Predict Patient Outcome in Acute Ischemic Stroke on the Basis of Imaging, Demographic, and Clinical Information," *Am. J. Roentgenol.*, vol. 212, no. 1, pp. 44–51, Jan. 2019, doi: 10.2214/AJR.18.20260.

- [79] A. Alawieh, F. Zaraket, M. B. Alawieh, A. R. Chatterjee, and A. Spiotta, "Using machine learning to optimize selection of elderly patients for endovascular thrombectomy," *J. NeuroInterventional Surg.*, vol. 11, no. 8, pp. 847–851, Aug. 2019, doi: 10.1136/neurintsurg-2018-014381.
- [80] J. You *et al.*, "Automated Hierarchy Evaluation System of Large Vessel Occlusion in Acute Ischemia Stroke," *Front. Neuroinformatics*, vol. 14, p. 13, Mar. 2020, doi: 10.3389/fninf.2020.00013.
- [81] D. Robben *et al.*, "Prediction of final infarct volume from native CT perfusion and treatment parameters using deep learning," *Med. Image Anal.*, vol. 59, p. 101589, Jan. 2020, doi: 10.1016/j.media.2019.101589.
- [82] B. Fu *et al.*, "Image Patch-Based Net Water Uptake and Radiomics Models Predict Malignant Cerebral Edema After Ischemic Stroke," *Front. Neurol.*, vol. 11, p. 609747, Dec. 2020, doi: 10.3389/fneur.2020.609747.
- [83] S. Bacchi, T. Zerner, L. Oakden-Rayner, T. Kleinig, S. Patel, and J. Jannes, "Deep Learning in the Prediction of Ischaemic Stroke Thrombolysis Functional Outcomes," *Acad. Radiol.*, vol. 27, no. 2, pp. e19–e23, Feb. 2020, doi: 10.1016/j.acra.2019.03.015.
- [84] G. Brugnara *et al.*, "Multimodal Predictive Modeling of Endovascular Treatment Outcome for Acute Ischemic Stroke Using Machine-Learning," *Stroke*, vol. 51, no. 12, pp. 3541–3551, Dec. 2020, doi: 10.1161/STROKEAHA.120.030287.
- [85] L. A. Ramos *et al.*, "Predicting Poor Outcome Before Endovascular Treatment in Patients With Acute Ischemic Stroke," *Front. Neurol.*, vol. 11, p. 580957, Oct. 2020, doi: 10.3389/fneur.2020.580957.
- [86] H. Nishi *et al.*, "Predicting Clinical Outcomes of Large Vessel Occlusion Before Mechanical Thrombectomy Using Machine Learning," *Stroke*, vol. 50, no. 9, pp. 2379–2388, Sep. 2019, doi: 10.1161/STROKEAHA.119.025411.
- [87] C.-C. Chung *et al.*, "Predicting major neurologic improvement and long-term outcome after thrombolysis using artificial neural networks," *J. Neurol. Sci.*, vol. 410, p. 116667, Mar. 2020, doi: 10.1016/j.jns.2020.116667.
- [88] S. M. Sung *et al.*, "Prediction of early neurological deterioration in acute minor ischemic stroke by machine learning algorithms," *Clin. Neurol. Neurosurg.*, vol. 195, p. 105892, Aug. 2020, doi: 10.1016/j.clineuro.2020.105892.
- [89] K. Matsumoto, Y. Nohara, H. Soejima, T. Yonehara, N. Nakashima, and M. Kamouchi, "Stroke Prognostic Scores and Data-Driven Prediction of Clinical Outcomes After Acute Ischemic Stroke.," *Stroke*, vol. 51, no. 5, pp. 1477–1483, May 2020, doi: 10.1161/STROKEAHA.119.027300.
- [90] L. Velagapudi *et al.*, "A Machine Learning Approach to First Pass Reperfusion in Mechanical Thrombectomy: Prediction and Feature Analysis," *J. Stroke Cerebrovasc. Dis.*, vol. 30, no. 7, Jul. 2021, doi: 10.1016/j.jstrokecerebrovasdis.2021.105796.
- [91] J. Hamann *et al.*, "Machine-learning-based outcome prediction in stroke patients with middle cerebral artery-M1 occlusions and early thrombectomy," *Eur. J. Neurol.*, vol. 28, no. 4, pp. 1234–1243, Apr. 2021, doi: 10.1111/ene.14651.
- [92] B. Jiang *et al.*, "Prediction of Clinical Outcome in Patients with Large-Vessel Acute Ischemic Stroke: Performance of Machine Learning versus SPAN-100," *Am. J. Neuroradiol.*, vol. 42, no. 2, pp. 240–246, 2021.
- [93] G. Varoquaux, P. R. Raamana, D. A. Engemann, A. Hoyos-Idrobo, Y. Schwartz, and B. Thirion, "Assessing and tuning brain decoders: Cross-validation, caveats, and guidelines," *NeuroImage*, vol. 145, pp. 166–179, Jan. 2017, doi: 10.1016/j.neuroimage.2016.10.038.
- [94] F. Eitel, M.-A. Schulz, M. Seiler, H. Walter, and K. Ritter, "Promises and pitfalls of deep neural networks in neuroimaging-based psychiatric research," *Exp. Neurol.*, vol. 339, p. 113608, May 2021, doi: 10.1016/j.expneurol.2021.113608.
- [95] J. Snoek, H. Larochelle, and R. P. Adams, "Practical Bayesian Optimization of Machine Learning Algorithms," in *Advances in Neural Information Processing Systems*, 2012, vol. 25. Accessed: Mar. 02, 2022. [Online]. Available: <https://proceedings.neurips.cc/paper/2012/hash/05311655a15b75fab86956663e1819cd-Abstract.html>
- [96] L. Li, K. Jamieson, G. DeSalvo, A. Rostamizadeh, and A. Talwalkar, "Hyperband: A Novel Bandit-Based Approach to Hyperparameter Optimization," p. 52.
- [97] E. Hazan, A. Klivans, and Y. Yuan, "Hyperparameter Optimization: A Spectral Approach," *ArXiv170600764 Cs Math Stat*, Jan. 2018, Accessed: Mar. 02, 2022. [Online]. Available: <http://arxiv.org/abs/1706.00764>

[98] I. Loshchilov and F. Hutter, "CMA-ES for Hyperparameter Optimization of Deep Neural Networks," *ArXiv160407269 Cs*, Apr. 2016, Accessed: Mar. 02, 2022. [Online]. Available: <http://arxiv.org/abs/1604.07269>

[99] J. Bergstra, R. Bardenet, Y. Bengio, and B. Kégl, "Algorithms for Hyper-Parameter Optimization," in *Advances in Neural Information Processing Systems*, 2011, vol. 24. Accessed: May 18, 2022. [Online]. Available: <https://proceedings.neurips.cc/paper/2011/hash/86e8f7ab32cfd12577bc2619bc635690-Abstract.html>

[100] C. and T. (European C. Directorate-General for Communications Networks and Grupa ekspertów wysokiego szczebla ds. sztucznej inteligencji, *Ethics guidelines for trustworthy AI*. LU: Publications Office of the European Union, 2019. Accessed: Mar. 02, 2022. [Online]. Available: <https://data.europa.eu/doi/10.2759/346720>

[101] D. Higgins and V. I. Madai, "From Bit to Bedside: A Practical Framework for Artificial Intelligence Product Development in Healthcare," *Adv. Intell. Syst.*, vol. 2, no. 10, p. 2000052, 2020, doi: 10.1002/aisy.202000052.

Figures

Figure 1

PRISMA Flow chart

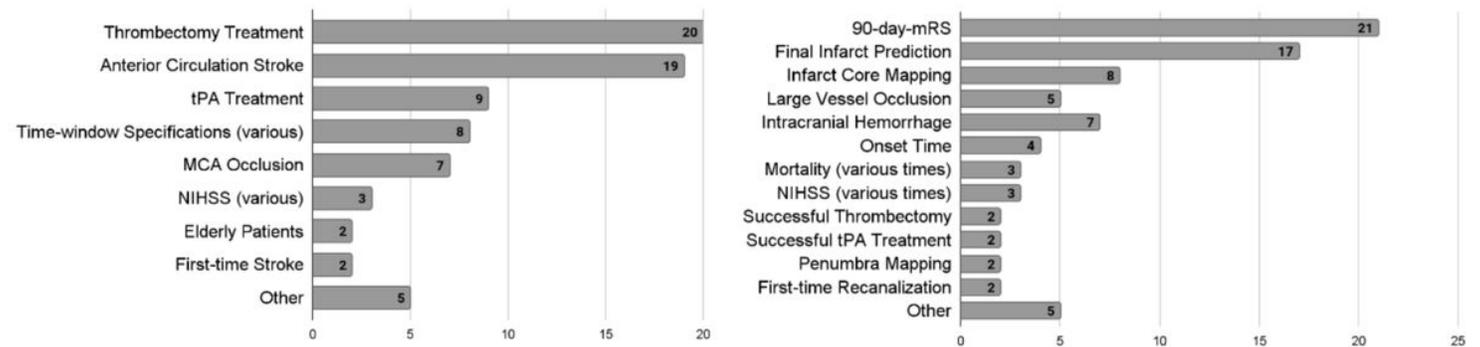


Fig. 2A Patient subgroups

Fig. 2B Outcome endpoint

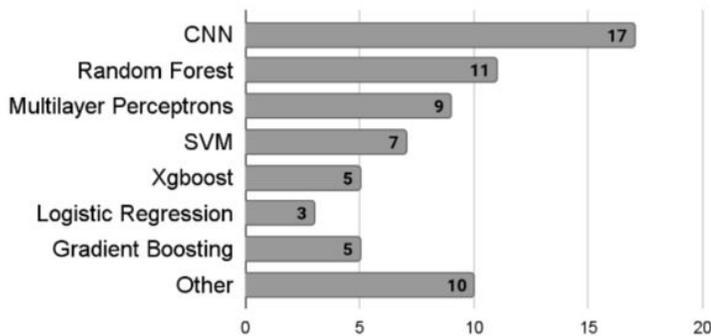


Fig. 2C Artificial intelligence method

Figure 2

See figure for legend.

Figure 3

See figure for legend.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [ArtificialIntelligenceforClinicalDecisionSupportinAcuteIschemicStrokeCareSUPPLEMENTAL.docx](#)
- [ArtificialIntelligenceforClinicalDecisionSupportinAcuteIschemicStrokeCareRESULTS.txt](#)