

Machine Learning Application: A Bibliometric Analysis From a Half-Century of Research on Stroke

Review began 08/01/2023

Review ended 08/20/2023

Published 08/26/2023

© Copyright 2023

Che Nawi et al. This is an open access article distributed under the terms of the Creative Commons Attribution License CC-BY 4.0., which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Che Muhammad Nur Hidayat Che Nawi ¹, Suhaily Mohd Hairon ¹, Wan Nur Nafisah Wan Yahya ², Wan Asyraf Wan Zaidi ², Mohd Rohaizat Hassan ³, Kamarul Imran Musa ¹

1. Department of Community Medicine, School of Medical Sciences, Universiti Sains Malaysia, Kubang Kerian, MYS 2. Department of Internal Medicine/ Neurology, Universiti Kebangsaan Malaysia Medical Centre, Kuala Lumpur, MYS 3. Department of Community Health, Faculty of Medicine, National University of Malaysia, Kuala Lumpur, MYS

Corresponding author: Kamarul Imran Musa, drkamarul@usm.my

Abstract

The quick advancement of digital technology through artificial intelligence has made it possible to deploy machine learning to predict stroke outcomes. Our aim is to examine the trend of machine learning applications in stroke-related research over the past 50 years.

We used search terms stroke and machine learning to search for English versions of original and review articles and conference proceedings published over the past 50 years in Scopus and Web of Science databases. The Biblioshiny web application was utilized for the analysis. The trend of publication and prominent authors and journals were analyzed and identified. The collaborative network between countries was mapped, and a thematic map was used to monitor the authors' trending keywords. In total, 10,535 publications authored by 44,990 authors from 2,212 sources were retrieved. Two distinct clusters of collaborative network nodes were observed, with the United States serving as a connecting node. Three terms - deep learning, algorithms, and neural networks - are observed in the early stages of the emerging theme. Overall, international research collaborations, the establishment of global research initiatives, the development of computational science, and the availability of big data have facilitated the pervasive use of machine learning techniques in stroke research.

Categories: Neurology, Epidemiology/Public Health

Keywords: stroke research, neural networks, algorithm, deep learning, bibliometric analyses, stroke, machine learning

Introduction And Background

Globally, stroke is the main cause of death and disability. The absolute number of incident strokes increased by 70.0% (67.0-73.0), prevalent strokes increased by 85.0% (83.0-88.0), deaths from stroke increased by 43.0% (31.0-55.0), and disability-adjusted life years (DALYs) attributable to stroke increased by 32.0% (22.0-42.0) between 1990 and 2019. [1]. Besides, 63% and 80% of stroke cases happen among people less than 70 years old and have a low to moderate cardiovascular disease absolute risk, respectively. Furthermore, nearly 90% of the global stroke deaths and disability combined reside in low-to-middle-income countries [2].

Nowadays, stroke research with structured data applying machine learning (ML) algorithms for outcome prediction has gained more popularity [3]. The complexity of a condition and the availability of routinely collected datasets like stroke may lend themselves well to the application of ML methods, which can incorporate a large number of variables and observations into a single predictive framework. There were a number of studies that incorporated various machine-learning algorithms to predict stroke outcomes in various global populations [4-6]. Mortality, functional outcome, duration of hospital stay, and neurologic deterioration were the post-stroke outcomes most frequently predicted. In addition, artificial neural network (ANN), support vector machine (SVM), decision tree (DT), and random forest (RF) were the most popular machine learning algorithms [3].

Both stroke and machine learning have been the subject of extensive research in the past. However, no investigation has been conducted into how these two distinct topics have been studied together. As the focal point of scientific evaluation, it is necessary to evaluate the growth, advancement, and impact of global research on these overlapping issues. Bibliometric analysis is a newly developed method for revealing patterns in published research. It evaluates an exhaustive literature search that demonstrates advancements within the field. Bibliometric analysis is an effective method for obtaining quantitative data on particular subjects [7]. Additionally, recognizing influential researchers and publications, countries to take into account for collaboration, and popular keywords on the application of ML in stroke research provide relevant references to the stakeholders working on stroke research to address current issues like a high stroke burden and research limitations.

How to cite this article

Che Nawi C, Mohd Hairon S, Wan Yahya W, et al. (August 26, 2023) Machine Learning Application: A Bibliometric Analysis From a Half-Century of Research on Stroke. Cureus 15(8): e44142. DOI 10.7759/cureus.44142

In light of this, the purpose of this study was to examine the trend of machine learning applications in stroke research. Specifically, we examined its trend over the past 50 years and identified distinctive collaborative network clusters and emerging themes.

Review

Methodology

Source of Data and Eligibility Criteria

The Scopus and Web of Science databases were selected for this bibliometric analysis. Scopus and Web of Science were databased by Elsevier and Clarivate, respectively. Only original English-language articles, reviews, and conference proceedings published before August 1, 2022, were eligible. Other publications, including editorials and letters, were excluded.

Search Strategy

The terms "stroke" and "machine learning" were searched in Scopus and Web of Science on August 1, 2022, for the article title, abstract, and keywords using the advanced search option in the respective databases and an appropriate combination of Boolean search operators. The flowchart of our investigation is depicted in Figure 1.

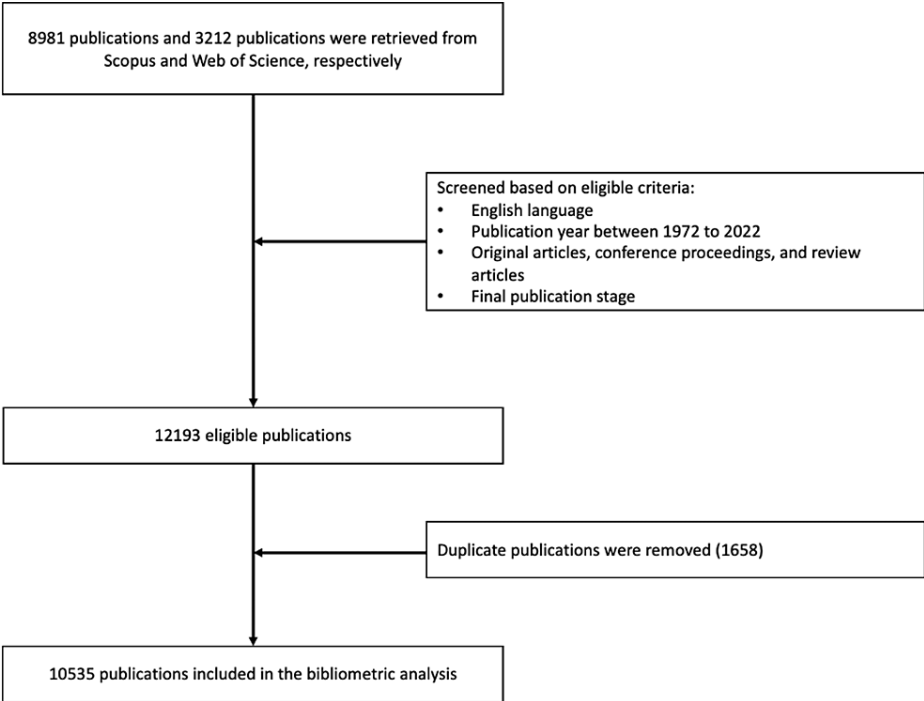


FIGURE 1: The flowchart of the research process for selecting publications in databases for the study

Appendix A contains an in-depth search strategy for the study.

Bibliometric Analyses

The BibTex format was used to retrieve and obtain the metadata consisting of all the primary information and characteristics of the publications included in the analysis. Data management and bibliometric analyses were performed utilizing the Bibliometrix package and Biblioshiny web applications under the R programming language (version 4.2.1, RStudio: Integrated Development for R. RStudio, PBC, Boston, MA, USA) [8,9]. The BibTex format was used to retrieve and obtain all the primary information and characteristics of the publications used in the analysis. The development of the publication trend spanned 50 years.

Based on the total number of publications and Hirsch (H)-indexed citations, we identified the top 10 most influential authors and sources. The total number of publications is defined as the number of publications produced within 50 years. Meanwhile, the H-index is computed by counting the number of publications for which an author has been cited at least that many times by other authors. The highest number of

publications and H-indexed citations will be in the top rank and followed by others. The influence of the top 10 sources was then tabulated in order to compare the total number of citations and H-indexed citations.

The publication by country was visualized, and a collaboration network was established between the top 43 countries that published the most articles during the time period. The frequency of the authors' keywords was plotted against the years to illustrate the dominant topics during the specified time period. Four themes were used to construct the 'thematic map'. They were (i) motor themes, (ii) niche themes, (iii) emerging or declining themes, and (iv) transversal and fundamental themes. On the map, co-occurrence networks were used to identify concentrations of keyword plus, which are keywords assigned by journals after extracting words and phrases from the titles of cited articles.

Results

Based on our study criteria, 8,981 and 3,212 publications were retrieved from Scopus and Web of Science, respectively. After screening for duplicate publications, we finally determined that 10,535 publications were eligible for analysis. The publications were from 2,212 sources and consisted of 9,653 original articles (91.6%), 216 review papers (2.1%), and 666 conference proceedings (6.3%). The annual publication growth rate is 13.8%, with average citations of 28.7 for each publication, as depicted in Figure 2.

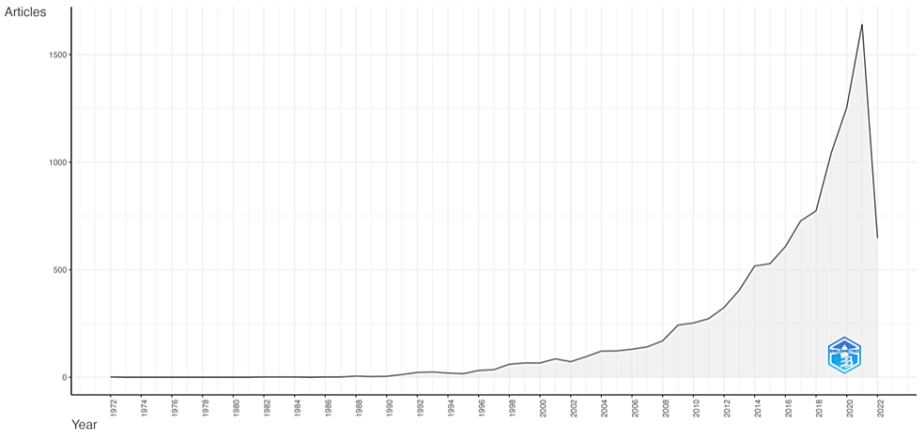


FIGURE 2: Annual scientific publications on stroke and machine learning from 1972 to 2022

Authors and Journals

Within the 50-year period, 44,990 authors were involved in publications regarding the application of ML to stroke research. Most publications (99%) were co-authored, with an average of 9.03 authors per publication. Table 1 lists the top 10 authors who published the most works over a 50-year span.

Rank	Element	Field of interest	Institution	Country	Total number of publications	H-indexed publications	G-indexed publications	M-indexed publications	Total citations	Start of publication year
1	Yanzhong Wang	Stroke epidemiology	King's College London,	United Kingdom	387	40	91	1.667	9695	2000
2	Kennedy R. Lees	Stroke, Internal medicine, cardiology	University of Glasgow	United Kingdom	62	34	62	1.214	5413	1996
3	Ralph L. Sacco	Neurology, public health	University of Miami	USA	77	34	69	0.971	4813	1989
4	Diederik Dippel	Neurology	Erasmus University Rotterdam	Netherlands	93	31	83	1.192	6903	1998
5	Stephen Davis	Neurology	University of Melbourne	Australia	48	30	48	1.364	3907	2002
6	Andrew Demchuk	Neurology	University of Calgary	Canada	67	30	65	1.2	4286	1999
7	Mitchell S. V. Elkind	Neurology	Columbia University	USA	67	30	54	1.25	3040	2000
8	Joon-Tae Kim	Neuroradiology	Chonnam National University	South Korea	150	30	51	1.364	3145	2002
9	Carlos Molina	Neurology	Hospital Universitari Vall d'Hebron	Spain	43	30	43	1.304	4654	2001
10	Chun-Fu Chen	Neurology	Shandong University	China	97	29	44	1.706	2145	2007

TABLE 1: Top 10 authors who have published research work on stroke and machine learning research over five decades

The author Yanzhong Wang from King’s College London, United Kingdom, was the most impactful author (387 publications, 3.67%), followed by Joon-Tae Kim from Chonnam National University, South Korea. Although the author Kennedy R. Lees produced less than half of Joon-Tae Kim’s publications, he was in the second rank with 34 H-indexed publications, followed by Ralph L. Sacco (34 H-indexed publications) and Diederik Dippel (31 H-indexed publications).

The top 10 most relevant journals publishing research on the application of ML in stroke research are listed in Table 2.

Rank	Element	Total number of publications	H-indexed publications	G-indexed publications	M-indexed publications	Total citations	Start of publication year
1	Stroke	1073	145	214	3.919	76449	1987
2	Neurology	152	51	85	1.545	7947	1991
3	Circulation	59	48	59	1.171	17734	1983
4	PLOS One	332	41	61	2.412	6648	2007
5	Journal of Vascular Surgery	130	37	62	1.028	4698	1988
6	Cerebrovascular Diseases	132	35	49	1.094	3436	1992
7	Journal of Thoracic and Cardiovascular Surgery	59	34	59	1.063	5130	1992
8	Journal of Neurology, Neurosurgery and Psychiatry	45	31	45	1	2944	1993
9	Archives of Neurology	33	29	33	0.935	3330	1993
10	American Journal of Neuroradiology	92	28	48	1.4	2666	2004

TABLE 2: The top 10 journals with the highest number of publications, journal impact, and the start year of publication on the topic

Table 2 includes the name of the journal, total number of publications, citation impact parameters, and start publication years. Since 1987, the Stroke Journal has been the most productive and impactful source, with over one thousand publications (145 H-indexed publications) published. As a comparison, the Neurology journal has the second highest number of H-indexed publications (51) followed by the Circulation journal (48 H-indexed publications) and the PLOS One journal (41 H-indexed publications). Although the start of the publication year for PLOS One and the American Journal of Neuroradiology journals is later than the other journals, these journals were ranked among the tenth most impactful journals with 332 publications (41 H-indexed publications) and 92 publications (28 H-indexed publications) within a shorter period. All 10 journals that are involved in publications regarding the application of ML to stroke research are specifically related to cardiovascular and cerebrovascular topics, except for PLOS One, which is an open-access and multidisciplinary journal.

Countries and Collaboration Network

Overall, authors from 107 countries have contributed to the publications on the application of ML to stroke research. The United States (2,484 publications, Single Country Production (SCP) of 2,439, Multiple Country Production (MCP) of 55) dominated the within and interstate production of stroke and machine learning research publications, followed by China (1,888 publications, SCP of 1,840, MCP of 48), the United Kingdom (517 publications, SCP of 505, MCP of 12), and Korea (476 publications, SCP of 467, MCP of 9). Table 3 provides a summary of the top 10 countries, along with their total publications, annual publication rate, single-country publications, multi-country publications, and their proportions.

Rank	Country	Number of publications	Frequency	SCP	MCP	MCP_SCP ratio
1	United States of America	2494	0.237	2439	55	0.022
2	China	1888	0.179	1840	48	0.025
3	United Kingdom	517	0.049	505	12	0.023
4	Korea	476	0.045	467	9	0.019
5	Japan	456	0.043	451	5	0.011
6	Germany	414	0.039	390	24	0.058
7	Netherlands	316	0.03	313	3	0.009
8	Canada	299	0.028	277	22	0.074
9	Spain	251	0.024	236	15	0.06
10	Australia	218	0.021	208	10	0.046

TABLE 3: The top 10 countries contributing publications relating to stroke and machine learning research

SCP: single country production; MCP: multiple country production

In the top 43 countries, two separate clusters of collaborative networks may be seen, as shown in Figure 3.

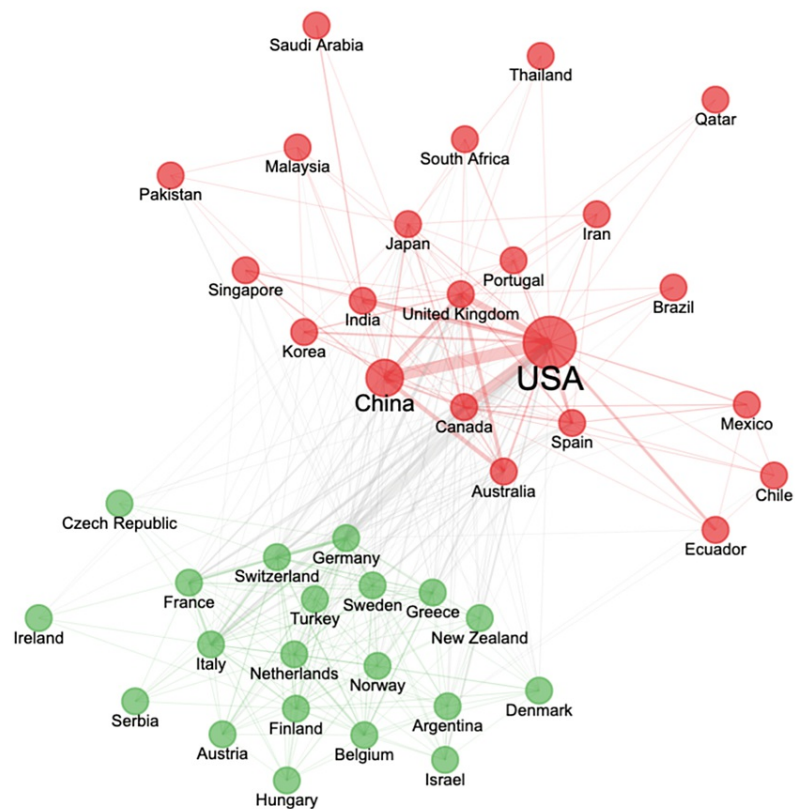


FIGURE 3: A collaborative network of countries producing research on stroke and machine learning

USA: United States of America

The bubble represents countries, while the connection lines between the bubbles represent their collaboration. The larger bubble is related to more publications, and the thicker line is related to their research collaboration activities. The first cluster (red bubbles) is made up of American and Asian countries, with the USA acting as the central node. In contrast, the second cluster (green bubbles) consists primarily of European countries, with Germany acting as the central node.

Trending Keywords and Thematic Map

Figure 4 shows the 63 trending author keywords that were found over the past 50 years of using ML in stroke research.

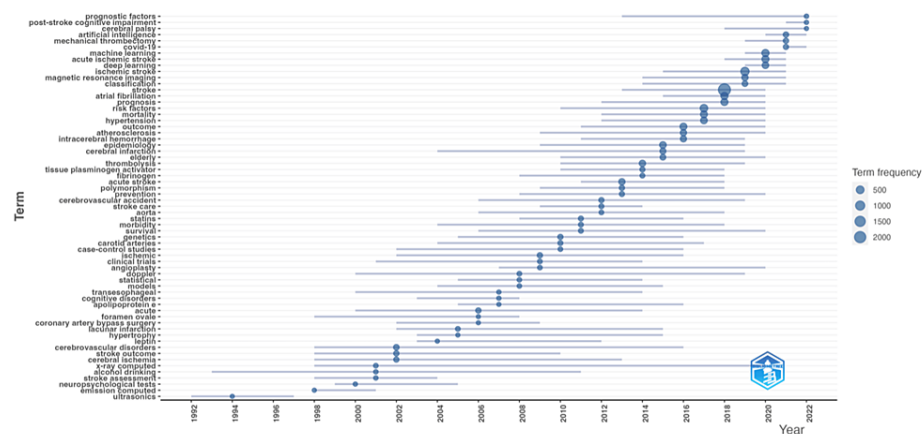


FIGURE 4: Trending topics and keywords for stroke and machine learning research from 1972 to 2022. The line shows the keyword's occurrence, while the bubble size shows its frequency throughout the year.

COVID-19: coronavirus disease 2019

In the top keywords, we found four patterns. Firstly, the change in the stroke diagnostic imaging keyword from x-ray computed (1998-2020), which was gradually replaced by magnetic resonance imaging (2014-2021). Secondly, in terms of post-stroke assessment and outcome measures, the keywords have changed gradually from general (neuropsychological tests, stroke outcome, cognitive disorder, and stroke care within 1998-2014) to specific assessment and outcome measures of stroke (stroke prognosis, prognostic factors, COVID-19, and post-stroke cognitive disorder within 2012-2022). Thirdly, we noticed a shift in the types of risk factors under study. Among the risk factors, alcohol and genetics were trending from 1998 to 2016, which gradually shifted to the elderly, atherosclerosis, hypertension, and atrial fibrillation from 2009 to 2020. Finally, in terms of statistical analysis of stroke research, the keywords shifted from model-based regression analysis in 2003 to model-free analysis in 2019.

Thematic maps represent keyword clusters as bubbles based on Callon's centrality (x-axis) and density (y-axis) ranks. The bubble size reflects the cluster's word frequency. Centrality measures the relationship between network clusters in the same graph. It indicates a theme's importance to research advancement. The density measures cluster network internal strength and subject development. The three keywords (a) deep learning, b) algorithms, and c) neural networks are observed in the early development stages of emerging themes (Figure 5).

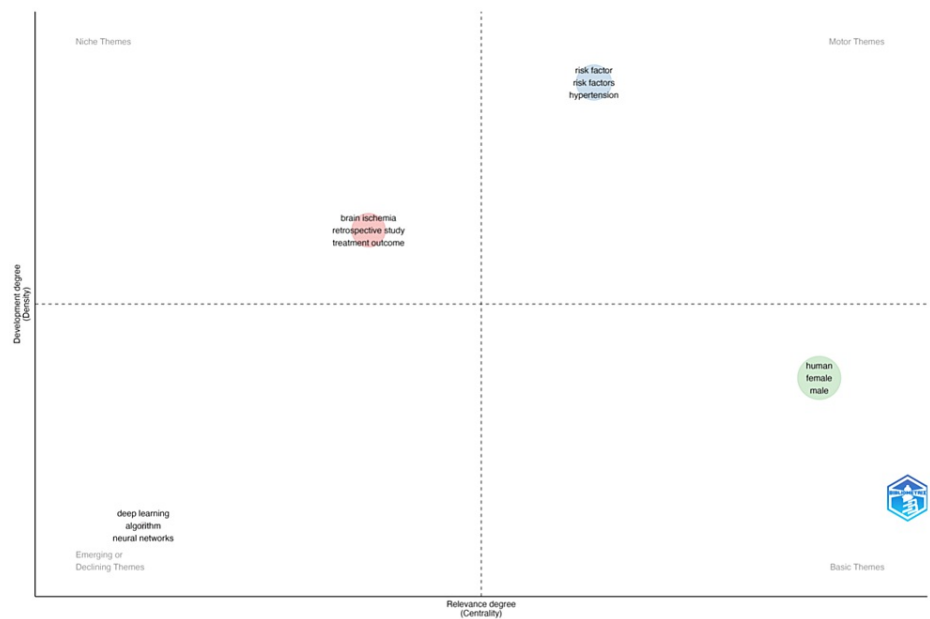


FIGURE 5: Thematic map of authors' keywords on stroke and machine learning research

Besides, stroke classification, research study method, and treatment outcome are observed in niche themes. Meanwhile, modifiable and non-modifiable stroke risk factors are observed in motor and basic themes, respectively.

Discussion

Bibliometric analysis using publication data from 1972-2022 shows that the publication on the application of ML in stroke research experienced a slow rise at first and a rapid rise thereafter. However, there is a notable decline in 2022, as the publication data were collected through August 2022. The steep decline does not accurately reflect the publication trend of stroke and machine learning in 2022. Besides, notable authors and institutions reside in the United Kingdom, Australia, and the United States. Furthermore, the United States served as the connecting node between two distinct clusters of collaborative networks. The terms deep learning, algorithm, and neural networks were classified in the early phase of an emerging theme.

The establishment of the World Stroke Organization (WSO) in 2006 has had a significant impact on stroke research [10]. The WSO was established to bring together professionals, organizations, and policymakers worldwide in the fight against stroke, aiming to reduce its global impact through prevention and improved care. Its membership encompasses 92 countries, fostering collaboration and the exchange of knowledge. The WSO contributes by increasing stroke awareness, providing evidence-based guidelines, promoting research, and advocating globally for improved stroke care and rehabilitation services [10-12]. Additionally, the availability of structured stroke data from multicenter stroke registries that are accessible for research, the development of computing capacity, and the availability of machine learning models all contribute to the research activity in this area [3].

All 10 prominent authors in stroke and machine learning research hail from diverse regions, including the United Kingdom, the Netherlands, Spain, China, South Korea, Canada, and the USA. Among them, Yanzhong Wang stands out as an exception, being a professor of statistics at King's College, London, focusing on stroke epidemiology. His expertise lies in clinical prediction models and the application of machine learning in stroke research [3,13]. Notably, in comparison to his clinician counterparts, Yanzhong Wang boasts a significantly higher number of publications, H-indexed citations, and other influential parameters. This data underscores the necessity for multidisciplinary collaboration in the realm of advanced machine learning stroke modeling, calling for expertise from statistics, clinical practice, and computer sciences.

The three most active countries in producing publications related to the application of machine learning in stroke research are the USA, China, and the United Kingdom. These countries collectively contribute to nearly half of the global stroke publications. This higher publication volume is attributed to the increased degree of collaboration among them. Notable examples illustrating this principle are initiatives such as Stroke Therapy Academic Industry Roundtable (STAIR) and Stem Cell Therapies as an Emerging Paradigm in Stroke (STEPS), which play a pivotal role in guiding and supporting translational stroke research [14]. These initiatives exemplify effective collaborative networks between stroke experts, particularly from the USA and

other nations within the first cluster (depicted by red bubbles).

Furthermore, the progress of stroke research is furthered by the support provided by the Leducq Foundation, which fosters laboratories to pursue mechanisms and targets related to stroke and cardiovascular diseases [15]. The Leducq Foundation has facilitated numerous research collaborations encompassing countries from both the first and second clusters, including the USA, Canada, the United Kingdom, Germany, France, Belgium, Sweden, and Austria.

Moreover, Safe Implementation of Treatments in Stroke (SITS) and the Virtual International Stroke Trials Archive (VISTA) stand as pivotal organizations that offer platforms and databases for the international analysis and advancement of stroke therapies [16,17]. Datasets from SITS and VISTA have been harnessed in a variety of studies to predict stroke outcomes through machine learning techniques [5,18,19]. Ultimately, all endeavors in advancing stroke research converge on the overarching goal of saving lives and enriching the quality of life for those affected by stroke.

From 1972 to 2022, the trend of publications supported the idea that researchers' interest in applying ML to stroke research has grown. This can be depicted by the change of author keywords from model-based to model-free analysis and the emergence of keywords related to machine learning in Figures 4-5 of our analyses. The model-based analysis uses standard regression methods such as logistic regression and Cox proportional hazard regression. Meanwhile, the model-free analysis uses machine learning algorithms such as decision trees, artificial neural networks, and random survival forests. The improvement of computational capacity and the availability of model-free algorithms further catalyze the use of machine learning in stroke research. To predict the outcomes of a stroke, machine learning employs vast, frequently collected datasets to generate accurate, individualized prognoses. These ML models-artificial neural networks (ANN), Nave Bayes, support vector machines (SVM), decision trees (DT), and random forests (RF)-have been utilized to predict stroke mortality [3].

There are a few limitations in our analysis. We cannot retrieve the actual number of publications on the ML application on stroke because some publications were written in languages other than English. Moreover, financial limitations and other various reasons might hinder the submission of manuscripts to the journals. These will result in not all publications being available in the Web of Science and Scopus databases. Thus, future studies may compare our findings to data from more databases with a wider diversity of language styles.

Conclusions

In summary, the bibliometric analysis of the application of ML in stroke research from 1972 to 2022 reveals a significant growth in the application of machine learning (ML) in this field. The United Kingdom, Australia, and the United States are prominent contributors to this field, and international collaborations, supported by initiatives and organizations, have played a crucial role in advancing stroke research. The trend indicates a shift towards ML analysis and the emergence of ML-related keywords, reflecting the increasing interest and utilization of ML algorithms in the multifaceted management of stroke care. Besides, computational advancements and the availability of big data have facilitated the widespread application of machine learning techniques in stroke research. These findings highlight the potential of ML to enhance stroke prevention, treatment, and patient care. Greater emphasis should be placed on augmenting research endeavors to stimulate increased publication output, particularly within the Southeast Asian region.

Appendices

Appendix A

Search strategy in Scopus and Web of Science

Databases	Search terms
Scopus	TITLE-ABS ("Strokes" OR "Cerebrovascular Accident" OR "Cerebrovascular Accidents" OR "CVA (Cerebrovascular Accident)" OR "CVAs (Cerebrovascular Accident)" OR "Cerebrovascular Apoplexy" OR "Apoplexy, Cerebrovascular" OR "Vascular Accident, Brain" OR "Brain Vascular Accident" OR "Brain Vascular Accidents" OR "Vascular Accidents, Brain" OR "Cerebrovascular Stroke" OR "Cerebrovascular Strokes" OR "Stroke, Cerebrovascular" OR "Strokes, Cerebrovascular" OR "Cerebral Stroke" OR "Cerebral Strokes" OR "Stroke, Cerebral" OR "Strokes, Cerebral" OR "Stroke, Acute" OR "Acute Stroke" OR "Acute Strokes" OR "Strokes, Acute" OR "Cerebrovascular Accident, Acute" OR "Acute Cerebrovascular Accident" OR "Acute Cerebrovascular Accidents" OR "Cerebrovascular Accidents, Acute") AND TITLE-ABS ("machine learning" OR "machine learnings" OR "deep learning" OR "Decision tree" OR "Classification tree" OR "Tree-based" OR "Support vector" OR "Transfer learning" OR "neural network" OR "neural networks" OR "Logistic regression" OR "discriminant analysis" OR "k-nearest neighbor" OR "k-nearest neighbour" OR "nearest neighbor" OR "nearest neighbour" OR bayes OR "random forest") AND (LIMIT-TO (OA , "all")) AND (LIMIT-TO (DOCTYPE , "ar") OR LIMIT-TO (DOCTYPE , "cp") OR LIMIT-TO (DOCTYPE , "re")) (TITLE-ABS=((cerebrovascular accident \$ OR apoplexy OR brain vascular accident \$ OR cva\$ OR cerebral stroke \$ OR Cerebral Infarction OR cerebrovascular stroke \$ OR stroke\$) OR (ischemi* AND (stroke \$ OR post Stroke OR poststroke)) OR ((transient AND ischemi*) OR tia\$) OR ((h\$emorrhag* OR (carotid AND artery) OR (carotid AND stenosis) OR (carotid AND ulcer\$)) AND (Stroke\$ OR poststroke OR post stroke)))) AND (TITLE-ABS=(machine learning OR Artificial Intelligence OR Naive Bayes OR bayesian learning OR Neural network\$ OR Natural language processing OR support vector machine\$ OR random forest\$ OR boosting OR deep learning OR machine intelligence OR computational intelligence OR computer reasoning))
Web of Science	#1: (AB=("Strokes" OR "Cerebrovascular Accident" OR "Cerebrovascular Accidents" OR "CVA (Cerebrovascular Accident)" OR "CVAs (Cerebrovascular Accident)" OR "Cerebrovascular Apoplexy" OR "Apoplexy, Cerebrovascular" OR "Vascular Accident, Brain" OR "Brain Vascular Accident" OR "Brain Vascular Accidents" OR "Vascular Accidents, Brain" OR "Cerebrovascular Stroke" OR "Cerebrovascular Strokes" OR "Stroke, Cerebrovascular" OR "Strokes, Cerebrovascular" OR "Cerebral Stroke" OR "Cerebral Strokes" OR "Stroke, Cerebral" OR "Strokes, Cerebral" OR "Stroke, Acute" OR "Acute Stroke" OR "Acute Strokes" OR "Strokes, Acute" OR "Cerebrovascular Accident, Acute" OR "Acute Cerebrovascular Accident" OR "Acute Cerebrovascular Accidents" OR "Cerebrovascular Accidents, Acute")) AND AB=("machine learning" OR "machine learnings" OR "deep learning" OR "Decision tree" OR "Classification tree" OR "Tree-based" OR "Support vector" OR "Transfer learning" OR "neural network" OR "neural networks" OR "Logistic regression" OR "discriminant analysis" OR "k-nearest neighbor" OR "k-nearest neighbour" OR "nearest neighbor" OR "nearest neighbour" OR bayes OR "random forest") #2: (TI=("Strokes" OR "Cerebrovascular Accident" OR "Cerebrovascular Accidents" OR "CVA (Cerebrovascular Accident)" OR "CVAs (Cerebrovascular Accident)" OR "Cerebrovascular Apoplexy" OR "Apoplexy, Cerebrovascular" OR "Vascular Accident, Brain" OR "Brain Vascular Accident" OR "Brain Vascular Accidents" OR "Vascular Accidents, Brain" OR "Cerebrovascular Stroke" OR "Cerebrovascular Strokes" OR "Stroke, Cerebrovascular" OR "Strokes, Cerebrovascular" OR "Cerebral Stroke" OR "Cerebral Strokes" OR "Stroke, Cerebral" OR "Strokes, Cerebral" OR "Stroke, Acute" OR "Acute Stroke" OR "Acute Strokes" OR "Strokes, Acute" OR "Cerebrovascular Accident, Acute" OR "Acute Cerebrovascular Accident" OR "Acute Cerebrovascular Accidents" OR "Cerebrovascular Accidents, Acute")) AND TI=("machine learning" OR "machine learnings" OR "deep learning" OR "Decision tree" OR "Classification tree" OR "Tree-based" OR "Support vector" OR "Transfer learning" OR "neural network" OR "neural networks" OR "Logistic regression" OR "discriminant analysis" OR "k-nearest neighbor" OR "k-nearest neighbour" OR "nearest neighbor" OR "nearest neighbour" OR bayes OR "random forest") #1 OR #2

TABLE 4: Database search terms for Scopus and Web of Science

Additional Information

Disclosures

Conflicts of interest: In compliance with the ICMJE uniform disclosure form, all authors declare the following: **Payment/services info:** All authors have declared that no financial support was received from any organization for the submitted work. **Financial relationships:** All authors have declared that they have no financial relationships at present or within the previous three years with any organizations that might have an interest in the submitted work. **Other relationships:** All authors have declared that there are no other relationships or activities that could appear to have influenced the submitted work.

Acknowledgements

Author contributions: C.M.N.H.C.N., S.M.H., W.N.N.W.Y., W.A.W.Z., M.R.H., and K.I.M. contributed to the conception and design of the study. C.M.N.H.C.N. contributed to the acquisition of the data. C.M.N.H.C.N., S.M.H., and K.I.M. contributed to the formal analysis and interpretation of the findings. C.M.N.H.C.N. drafted the manuscript. All authors have read and agreed to the published version of the manuscript.

References

1. GBD 2019 Stroke Collaborators: Global, regional, and national burden of stroke and its risk factors, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019. *Lancet Neurol.* 2021, 20:795-820. [10.1016/S1474-4422\(21\)00252-0](https://doi.org/10.1016/S1474-4422(21)00252-0)
2. World Stroke Organization: Global Fact Sheet 2022. (2022). Accessed: 2022: <https://www.dropbox.com/scl/fi/f9c3n8jlrq85ohts6t8m/WSO-Global-Stroke-Factsheet-09-02.2022.ppt?dl=0&rlkey=grvocqyp5i....>
3. Wang W, Kiik M, Peek N, et al.: A systematic review of machine learning models for predicting outcomes of stroke with structured data. *PLoS One.* 2020, 15:e0234722. [10.1371/journal.pone.0234722](https://doi.org/10.1371/journal.pone.0234722)
4. Fernandez-Lozano C, Hervella P, Mato-Abad V, et al.: Random forest-based prediction of stroke outcome. *Sci Rep.* 2021, 11:10071. [10.1038/s41598-021-89434-7](https://doi.org/10.1038/s41598-021-89434-7)
5. Monteiro M, Fonseca AC, Freitas AT, Pinho E Melo T, Francisco AP, Ferro JM, Oliveira AL: Using machine learning to improve the prediction of functional outcome in ischemic stroke patients. *IEEE/ACM Trans Comput Biol Bioinform.* 2018, 15:1953-9. [10.1109/TCBB.2018.2811471](https://doi.org/10.1109/TCBB.2018.2811471)
6. van Os HJ, Ramos LA, Hilbert A, et al.: Predicting outcome of endovascular treatment for acute ischemic stroke: potential value of machine learning algorithms. *Front Neurol.* 2018, 9:784. [10.3389/fneur.2018.00784](https://doi.org/10.3389/fneur.2018.00784)
7. Van Houten BA, Phelps J, Barnes M, Suk WA: Evaluating scientific impact. *Environ Health Perspect.* 2000, 108:A392-3. [10.1289/ehp.108-a392](https://doi.org/10.1289/ehp.108-a392)
8. Aria M, Cuccurullo C: bibliometrix: an R-tool for comprehensive science mapping analysis. *J Informetr.* 2017, 11:959-75. [10.1016/j.joi.2017.08.007](https://doi.org/10.1016/j.joi.2017.08.007)
9. RStudio: integrated development for R. RStudio. (2020). Accessed: 2023: <http://www.rstudio.com/>.
10. Feigin VL, Brainin M, Norrving B, et al.: World Stroke Organization (WSO): global stroke fact sheet 2022. *Int J Stroke.* 2022, 17:18-29. [10.1177/17474930211065917](https://doi.org/10.1177/17474930211065917)
11. Mead GE, Sposato LA, Sampaio Silva G, et al.: A systematic review and synthesis of global stroke guidelines on behalf of the World Stroke Organization. *Int J Stroke.* 2023, 18:499-531. [10.1177/17474930231156753](https://doi.org/10.1177/17474930231156753)
12. Lindsay P, Furie KL, Davis SM, Donnan GA, Norrving B: World Stroke Organization global stroke services guidelines and action plan. *Int J Stroke.* 2014, 9 Suppl A100:4-13. [10.1111/ijis.12371](https://doi.org/10.1111/ijis.12371)
13. Wang W, Rudd AG, Wang Y, Curcin V, Wolfe CD, Peek N, Bray B: Correction: risk prediction of 30-day mortality after stroke using machine learning: a nationwide registry-based cohort study. *BMC Neurol.* 2022, 22:319. [10.1186/s12883-022-02840-w](https://doi.org/10.1186/s12883-022-02840-w)
14. Lapchak PA, Zhang JH, Noble-Haeusslein LJ: RIGOR guidelines: escalating STAIR and STEPS for effective translational research. *Transl Stroke Res.* 2013, 4:279-85. [10.1007/s12975-012-0209-2](https://doi.org/10.1007/s12975-012-0209-2)
15. Tancredi D, Braunwald E: Fondation leducq. *Circ Res.* 2014, 115:419-22. [10.1161/CIRCRESAHA.115.304626](https://doi.org/10.1161/CIRCRESAHA.115.304626)
16. Weimar C, Ali M, Lees KR, Bluhmki E, Donnan GA, Diener HC: The Virtual International Stroke Trials Archive (VISTA): results and impact on future stroke trials and management of stroke patients. *Int J Stroke.* 2010, 5:103-9. [10.1111/j.1747-4949.2010.00414.x](https://doi.org/10.1111/j.1747-4949.2010.00414.x)
17. Ahmed N, Lees KR, von Kummer R, et al.: The sits open study: a prospective, open label blinded evaluation study of thrombectomy in clinical practice. *Stroke.* 2021, 52:792-801. [10.1161/STROKEAHA.120.031031](https://doi.org/10.1161/STROKEAHA.120.031031)
18. Liu J, Chen X, Guo X, Xu R, Wang Y, Liu M: Machine learning prediction of symptomatic intracerebral hemorrhage after stroke thrombolysis: a cross-cultural validation in Caucasian and Han Chinese cohort. *Ther Adv Neurol Disord.* 2022, 15:17562864221129380. [10.1177/17562864221129380](https://doi.org/10.1177/17562864221129380)
19. Quandt F, Flottmann F, Madai VI, et al.: Machine learning-based identification of target groups for thrombectomy in acute stroke. *Transl Stroke Res.* 2023, 14:311-21. [10.1007/s12975-022-01040-5](https://doi.org/10.1007/s12975-022-01040-5)