



# Machine Learning in Stroke Medicine: Opportunities and Challenges for Risk Prediction and Prevention

# 5

Julia Amann

## 5.1 Introduction

“The essence of practicing medicine has been obtaining as much data about the patient’s health or disease as possible and making decisions based on that. Physicians have had to rely on their experience, judgement, and problem-solving skills while using rudimentary tools and limited resources.” [1]

Precision medicine aims to individualize prevention, diagnostics, and therapeutics by understanding differences in individuals’ genetics, lifestyle, and environment [2]. Over the past years, we have been witnessing an unprecedented push toward a more data-driven approach in healthcare that promises to take precision medicine to the next level, in part through artificial intelligence (AI). Simply put, AI can be understood as a set of sophisticated computational methods that seek to mimic human cognitive functions, including visual perception, speech recognition, and decision-making [3, 4]. AI uses certain machine learning (ML) algorithms to “learn” features from large datasets [3] and recognize patterns that are often invisible to the human eye [5–7]. Capitalizing on the availability of big data and ever-increasing computational power and storage capacities [1, 8], these novel tools seek to improve population health and well-being and to reduce healthcare costs.

A surge in scientific publications documents the potential to harness artificial intelligence in healthcare to prevent, diagnose, and treat diseases [9]. One of the pressing disease areas in focus for AI researchers is stroke, a leading cause of disability and mortality worldwide [3, 8]. Researchers aim to develop applications to optimize stroke diagnosis, treatment, and rehabilitation [10–12], and they also use AI to better understand risk. Several well-established risk prediction models have

---

J. Amann (✉)  
Health Ethics and Policy Lab, Department of Health Sciences and Technology,  
ETH Zurich, Zurich, Switzerland  
e-mail: [julia.amann@hest.ethz.ch](mailto:julia.amann@hest.ethz.ch)

© The Author(s) 2021  
F. Jotterand, M. Ienca (eds.), *Artificial Intelligence in Brain and Mental Health: Philosophical, Ethical & Policy Issues*, Advances in Neuroethics,  
[https://doi.org/10.1007/978-3-030-74188-4\\_5](https://doi.org/10.1007/978-3-030-74188-4_5)

been developed as tools for stroke prevention [13]. Prevention plays an instrumental role in reducing the global burden of stroke [14], and the strategic adoption and development of AI-driven prediction tools can contribute substantially to this mission [1, 13]. These new tools open welcome opportunities and introduce new questions for us, of course. We find ourselves only at the beginning of this exciting journey that will without a doubt confront us with novel ethical, societal, and regulatory challenges.

This chapter surveys the global burden of stroke and describes current practices for reducing stroke incidence and stroke mortality rates. In particular, the chapter reviews how ML applications are applied to stroke risk prediction and prevention and identifies important technological and methodological challenges for using AI in these contexts. The chapter concludes by drawing the readers' attention to some of the questions and ethical challenges that arise as clinicians widely adopt ML-based applications in practice.

---

## 5.2 Burden of Stroke

Stroke is one of the leading causes of disability and mortality worldwide [14–17]. Even though a decrease in stroke mortality and incident rates was observed from 1990 to 2016, absolute numbers show an increase in stroke-related mortality and disability [15, 16]. The absolute number of people affected by stroke almost doubled during this time [16] with incidence rates in low- to middle-income countries exceeding those observed in high-income countries [18]. Researchers estimate that in 2016, there were over 80 million people affected by stroke, many of them younger than 70 years of age [15, 16]. In 2017, Europe counted 1.5 million stroke diagnoses and nine million stroke survivors, with 1.2 million experiencing severe limitations in their activities of daily living [19]. That same year, 0.4 million people died because of stroke [19]. The increase in absolute numbers is largely attributed to population aging and growth [20, 21]. Yet, a noteworthy increase was also recorded in stroke incidence rates in younger age groups (15- to 49-year olds) [16].

The global increase in stroke incidents poses major challenges for healthcare systems, and these challenges extend beyond a patient's hospital stay. Patients who survive a stroke long to return to normality [22]. However, following hospital discharge, stroke survivors and their families must cope with the aftermath of stroke. People who suffered a stroke often experience more or less severe physical, cognitive, and emotional deficits that may limit their ability to perform certain activities in daily life [23, 24]. As a result, they remain at least partially dependent on an informal caregiver, usually a family member or partner [25]. Stroke survivors and informal caregivers commonly report physical, emotional, social, and financial challenges and concerns [26, 27]. They also face service deficiencies in health and social care, limited options for service offers outside of healthcare, and a paucity of options for continuity of care. All of this lays an additional burden on those affected by stroke, leaving them frustrated and under emotional strain [27].

In addition to the impact of stroke on individuals, societies are faced with the economic burden of stroke [28]. Healthcare utilization, informal care provision, and the loss of productivity in the workforce contribute to these rising costs [21, 29, 30]. A recent study analyzing stroke-related costs for 32 European countries estimates that total costs added up to €60 billion in 2017. This includes €27 billion (45%) incurred by healthcare systems, €5 billion (8%) incurred by social care systems, an estimated €16 billion (27%) for informal care costs, and €13 billion (20%) owed to lost productivity due to early death or absence from work [19]. While lower total costs to the healthcare system have been reported for the United States for 2014/2015 [31], per capita healthcare-related spending on stroke was higher in the USA compared to Europe [19]. Similar costs were reported for stroke-related healthcare costs per stroke survivor living in the USA and Europe [19].

---

### 5.3 Stroke Prevention: A Public Health Priority

As the global stroke burden increases, researchers and policymakers call for more efficient stroke prevention and management strategies and improved access to stroke services [16, 17, 32, 33]. In 2006, the World Health Organization (WHO) highlighted neurological disorders, including stroke, as a public health priority [34]. With its Global Status Report on Noncommunicable Diseases 2014, WHO aimed to unite and support nations in the fight against stroke and vascular diseases [32, 33].

There is common agreement that prevention is one, if not the most, promising strategy to reduce the burden of stroke [16, 35–37]. It is well established that there are non-modifiable (e.g., sex, gender, genetics) and modifiable (e.g., smoking cessation, physical inactivity) risk factors for stroke [38, 39]. Modifiable risk factors are the obvious targets of stroke prevention efforts. In an international case-control study, researchers found that ten risk factors (history of hypertension, current smoking, waist-to-hip ratio, diet risk score, regular physical activity, diabetes mellitus, binge alcohol consumption, psychosocial stress and depression, cardiac diseases, and ratio of apolipoproteins B to A1) were associated to 90% of the risk of stroke [39]. The authors concluded that lifestyle interventions targeting blood pressure reduction, smoking cessation, and the promotion of physical activity and a healthy diet could help to significantly reduce the burden of stroke.

There are two main approaches in stroke prevention [40]: population-wide prevention strategies and prevention strategies that target high-risk individuals. Population-wide strategies aim at modifying behavioral and lifestyle risk factors in the entire population to promote health maintenance [41]. In doing so, they can also contribute to preventing other diseases and chronic conditions (e.g., hypertension and diabetes mellitus) that constitute known stroke risk factors [14]. Recent advances in our ability to accurately assess individual risk for cardiovascular diseases have motivated some countries to prioritize risk-based screening approaches to identify individuals at risk [42, 43].

Despite a formal distinction between these two approaches, it is important to note that stroke risk is a continuum with no determined threshold at which certain

interventions are automatically indicated. Therefore, it may not be appropriate to categorize individuals into low-, moderate-, and high-risk groups when communicating absolute cardiovascular risk [44]. To effectively reduce stroke incidence and mortality rates, efforts must be undertaken to educate the general population about known behavioral risk factors [14, 43]. In addition, inexpensive screening strategies should be adopted to assist clinicians in identifying and protecting high-risk individuals [14, 43].

---

## 5.4 The Advent of Data-Driven Risk Prediction Models

Early prediction of stroke risk is the cornerstone of stroke prevention [45]. Identifying individuals who could benefit most from specific therapeutics or interventions helps them get the care they need and simultaneously helps avoid unnecessary treatments for others [10, 46, 47]. To date, several well-established statistically derived risk prediction models have been developed to provide long-term risk prediction [42, 45, 48, 49]. Clinicians commonly rely on these models to assess long-term risk, because the models provide parameters that are easy to interpret, such as odds ratios, relative risks, and hazard ratios [50]. However, these traditional models are subject to several limitations. They can, for example, only include a small number of risk factors (predictors) and generally do not include image-based morphological characteristics [13, 50, 51] nor behavioral risk factors (except smoking) or independent genetic factors [43]. Moreover, traditional approaches rely on certain assumptions of linearity, thus forcing models to behave in a certain way [51]. Often, traditional models are not generalizable across different populations due to the specific characteristics of the cohorts they were derived from [13]. This may lead clinicians to over- or underestimate risk for their patients [52].

Researchers are now trying to use ML in cardiovascular diseases and stroke risk assessment to overcome some of the challenges associated with traditional risk prediction models. ML methods use computational algorithms to relate all or some predictor variables of a given set to an outcome variable [50]. Classification and regression are the two primary tasks performed by ML-based algorithms [13]. Put simply, classification tasks categorize input data into predefined labels or outcomes (e.g., event or no event), whereas regression tasks predict some real-valued output (e.g., real-valued percentage risk between 0% and 100%). Despite various commonalities, ML differs from traditional statistical approaches in some aspects [53–55]. Contrary to classical statistics, ML is a data-driven approach that does not rely on a predefined model and assumption of data normality [53, 56]. Moreover, unlike traditional statistics which are focused on the “typical patient,” ML is capable of making inferences at the individual level, taking into account individual differences in the data [53]. ML is also inherently a multivariate approach that can be used to analyze complex and heterogeneous kinds of data and incorporate them into risk prediction models, making it a promising solution for stroke risk prediction [53, 54, 57].

Studies investigating the use of these techniques in cardiovascular diseases and stroke prediction indicate that ML-based approaches can boost prediction accuracy. A recently published review found that the most common ML-based algorithms used in cardiovascular risk assessment are support vector machines, artificial neural networks, linear and logistic regression, and tree-based algorithms, such as random forests and gradient tree boosting [13]. In their review, Jamthikar et al. further showed that ML-based algorithms performed better compared to traditional regression-based methods for risk assessment, and that including both image-based features and conventional cardiovascular risk factors drives prediction accuracy. Indeed, imaging plays a pivotal role in cardiovascular and stroke risk detection. Ultrasound, in particular carotid ultrasound screening, can also easily be performed in routine clinical practice—unlike other non-invasive techniques, such as computed tomography or magnetic resonance imaging [47]—making ultrasound an invaluable tool for stroke prevention. In line with these findings, Ambale-Venkatesh et al. [58] emphasized the importance of subclinical disease markers obtained from imaging, electrocardiography, and blood tests. The authors found that ML in conjunction with deep phenotyping (i.e., multiple evaluations of different aspects of a specific disease process) enhanced prediction accuracy of cardiovascular events compared to traditional risk scores.

Several other studies provide similar evidence. In a prospective cohort study using routine clinical data, for example, researchers compared four machine-learning algorithms (random forest, logistic regression, gradient boosting machines, neural networks) to an established algorithm (American College of Cardiology guidelines) for first cardiovascular event prediction over 10 years [46]. Their findings show that ML techniques outperformed the established algorithm, leading to a significantly more accurate risk prediction. Similarly, a team of researchers demonstrated that their hybrid ML approach to stroke prediction significantly reduced the false-negative rate in comparison to conventional approaches, while the overall error increased only slightly [59]. In addition to increasing prediction accuracy, authors also recognize the potential of ML-based approaches to help identify new potential risk factors and to generate a better understanding of the role of novel biomarkers [59, 60].

---

## 5.5 From Data-Driven Risk Prediction to Stroke Prevention

Accurate risk prediction allows clinicians and patients to act. Enabled by advances in AI technologies that can analyze vast volumes of health data in an efficient and accurate manner [4], precision medicine aims to provide treatment and prevention tailored to individuals' variability in genetics, environment, and lifestyle [1]. At present, doctors recommend lifestyle changes to their patients, advising them to change known, modifiable risk factors to prevent stroke. Yet, their advice often goes unheeded. We should eat healthy, refrain from smoking and eschew excessive alcohol consumption, exercise regularly, stay hydrated, and the list goes on and on. To adhere to all these health-promoting recommendations in a world full of competing

priorities, temptation, and imposed restrictions (e.g., financial constraints, poor access) may be too much to ask and simply not a realistic goal for many people. Earlier work has shown that there are incongruities between what people know they should do and their actual health behavior. So even though interventions (e.g., public health campaigns) may help to improve people's knowledge, these interventions may ultimately fail to induce, and more importantly, sustain behavior change—a phenomenon commonly referred to as the knowledge-behavior gap [61, 62].

Precision medicine is a promising approach to bridge this gap. It enables physicians and researchers to predict more accurately which prevention strategies will be most effective for which groups of people [1]. Understanding their natural predisposition to stroke may, in turn, motivate individuals to take on a more active role in their own health to reduce their individual stroke risk [14, 63]. In this context, the potential of mobile monitoring devices with real-time feedback systems has been highlighted as a tool for stroke prevention [10, 60, 64–67]. However, despite the promise these novel technologies hold for enabling personalized risk assessment and promoting stroke prevention, achieving stroke prevention via these means will largely depend on patients' acceptance and uptake of the technology. Tran et al. investigated chronic patients' perceptions of wearable biometric monitoring devices and AI systems that enable remote measurement and analysis of patient data in real-time [68]. In addition to capturing the perceived benefits and dangers of using these new technologies, the authors also assessed patients' readiness for using them. Their findings indicated that only half of the patients who participated in the study viewed digital tools and AI in healthcare as an opportunity, while 11% even considered them a danger, fearing that these will lead to the replacement of humans. In light of these findings, it is not surprising that 35% of patients indicated that they would refuse to integrate such devices into their care. More research is needed to better understand individuals' underlying motivations and fears that influence their attitudes toward the use of mobile monitoring devices and AI in healthcare. It is currently also unclear how well these new tools will be received by healthcare professionals. So, while AI-powered technologies are evolving rapidly, providing unprecedented opportunities for precision medicine in stroke prevention, the integration of these technologies into clinical practice raises several questions.

A project that will shed light on some of these questions is PRECISE4Q, a project funded under the European Union's Horizon 2020 Research and Innovation Program [69–71]. PRECISE4Q aims to identify and quantify risk factors and individual risk factor patterns. To do so, it combines heterogeneous data from a variety of sources, including large retrospective longitudinal stroke registry data, biobank data, and insurance data. What distinguishes PRECISE4Q from many other efforts in the field is its hybrid modeling approach, which combines ML methods and theory-driven (mechanistic modeling) approaches to risk prediction. Within the course of the project, a Digital Stroke Patient Platform will be established to collect and integrate large-scale data sets. This platform will also feature novel hybrid model architectures, structured prediction models, complex deep learning and gradient boosting models, as well as Clinical Decision Support Systems (CDSS) for stroke risk assessment, treatment outcomes, rehabilitation programs, and a

socio-economic planning tool. A thorough validation of the models is planned with clinical data generated by prospective clinical studies and retrospective analyses of health registries, cohort studies, health insurance data, and electronic health records. The CDSS envisioned by PRECISE4Q will allow clinicians to simulate how an individual's stroke risk will evolve and change under different circumstances over time. In other words, clinicians will be able to simulate how different risk factors (e.g., smoking) will contribute to disease occurrence and how the individual will respond to different possible interventions (e.g., lifestyle intervention, medication). This will assist them in providing individuals with tailored recommendations based on their natural predisposition. For individuals, this means that they will learn not only their individual stroke risk but also what they can do to reduce this risk.

Another promising avenue for future research is the use of natural language processing to automatically extract information on lifestyle modification assessment and/or advice in clinical practice from electronic health records [72–74]. Such analyses can provide an objective evaluation of current clinical practice and improve our understanding of the timing of lifestyle modification and patient, clinic, and provider characteristics that are associated with or predictive of lifestyle modification documentation [73]. Understanding how and when clinicians assess lifestyle modification and provide advice to patients holds important implications for the development of prevention strategies. These insights can inform the improvement of care delivery and documentation in practice. Combining tools aimed at understanding current clinical practice with sophisticated risk prediction models, such as the ones described earlier, constitutes an opportunity to deepen our understanding of stroke prevention.

---

## 5.6 Technological, Methodological, and Ethical Challenges

Machine learning holds great promise for stroke prevention, yet it is also subject to some challenges and limitations. There are three common areas of challenges that clinicians and researchers should be mindful of as they seek to maximize the advantages of ML in stroke prevention, and in healthcare more generally: (1) challenges in data sourcing; (2) challenges in application development; (3) challenges in deployment in clinical practice [75]. Given that patients' health and well-being are at stake, it is of critical importance to investigate the technological and methodological challenges that arise at each stage and to consider their potential real-life consequences. It is also important to note that challenges occurring at one stage may have consequences for the subsequent stages. Challenges and limitations at the stage of data sourcing, for example, inevitably affect application development and deployment in clinical practice.

### 5.6.1 Data Sourcing

High-quality big data is key to accurate predictions. To develop ML systems that can be deployed in clinical practice, a continuous supply of large datasets is needed initially to train, validate, and improve algorithms [3, 76]. Yet, inadequate access to

well-established patient and population-based datasets constitutes a major challenge for many ML-based data scientists and developers [13]. These professionals lack access to data partly because effective data sharing is currently not sufficiently incentivized by the medical scientific community [3, 10, 13, 77]. International research collaborations can help to mitigate this challenge. In the long run, effective data sharing strategies also need to be in place to facilitate and incentivize data sharing across institutions.

Another challenge to data sourcing relates to data protection and privacy regulations. Personal data are often subject to protective regulations that may impede data sharing. The European General Data Protection Regulation (GDPR), for example, entails a comprehensive set of regulations for the collection, storage, and use of personal information that will affect AI implementation in healthcare in several ways [76, 78]. The GDPR requires that individuals give explicit and informed consent before any organization collects personal data. It also grants individuals the right to track what data organizations are collecting about them, and it empowers them to direct an organization to discard their data. While these regulations rightly aim to protect patient privacy, they of course also impose certain restrictions on researchers and clinicians who seek to utilize these data. At present, the long-term impact of the GDPR and similar regulations on the implementation of AI in healthcare remains to be seen.

Closely related to data sourcing, data harmonization across different sources can also be quite problematic for data scientists. Given that very few studies provide comprehensive datasets for large numbers of participants, collaborative efforts are currently underway in the scientific community to harmonize and synthesize heterogeneous data across studies [79]. However, data harmonization is a time-consuming task that demands significant technological and scientific investments [80, 81].

## 5.6.2 Application Development

As outlined in this chapter, there is substantial evidence to suggest that ML-based algorithms can provide robust and accurate models for cardiovascular and stroke risk assessment, and can often outperform traditional regression-based approaches. Yet, there are several potential challenges and pitfalls to be mindful of when it comes to developing apps based on these algorithms. One of the key challenges in application development is algorithmic bias, which leads to systemic and unfair discrimination against certain individuals or groups of individuals [82, 83]. Even if no discrimination is intended, we know that the way data is collected, selected, prepared, and used to train ML-based algorithms can introduce bias [82]. Datasets used to develop stroke risk prediction models may, for example, suffer from missing data, misclassification, and measurement error, which can lead researchers and clinicians to make inaccurate predictions for subgroups of patients [84]. In other words, bias can occur when data sources do not reflect the true epidemiology within a given demographic [75]. As an example, consider that cardiovascular disease is often underdiagnosed in women because their symptoms are described as atypical



[85, 86]. Using such data to train ML-based algorithms may further reinforce this trend.

It has also been shown that ML methods perform poorly on imbalanced datasets, as they will be biased towards the majority group [59, 87, 88]. In other words, insufficient training samples and imbalanced class distribution will limit predictive performance in cases of rare occurrences [89]. In the case of stroke risk prediction, this may, for instance, pose limitations when we aim to develop predictive models for younger populations since the vast majority of available records likely describe older age groups [89]. Even though several balancing techniques have been developed, it is still a challenge to detect and address this bias in ML models [88].

But what does persistent algorithmic bias mean in practice? Algorithmic bias can cause enormous harm and contribute to increasing existing health inequalities in the real world [83]. A prominent example is the case of racial bias in commercial algorithms used in the U.S. healthcare system. In their 2019 study, Obermeyer et al. [90] found evidence indicating that a widely used algorithm was significantly biased against black patients. Due to this racial bias, a significantly lower number of black patients were identified for extra care. The authors demonstrated that bias occurred because the algorithm predicted healthcare costs rather than illness, not accounting for the fact that unequal access to care means that healthcare spending is lower for black patients than for white patients. The study carried out by Obermeyer et al. [90] serves as a striking example of how ML-based algorithms can reinforce existing inequalities and cause harm. It also raises the question: how many biased algorithms are still out there operating day in, day out? Importantly, this kind of bias is by no means limited to the US or to US race demographics. Similar problems can just as well be embedded in European algorithms, hiding similar (or different) kinds of social disparity.

### 5.6.3 Deployment in Clinical Practice

Finally, the practical implementation of AI technologies in healthcare is not without its own challenges [76, 91]. Trust plays a fundamental role in the implementation process. To obtain acceptance, AI-powered tools must first gain healthcare providers' and patients' trust [92]. As a first important step to gaining trust, tools should comply with existing data protection requirements and be transparent as to how outcomes and recommendations are derived [75]. However, at present, many ML models are considered black boxes that do not explain how their predictions are derived in a way that humans can grasp [93]. Unlike well-established regression-based methods where a clear relationship can be observed between the input variables and the output variable, the internal workings of ML algorithms are not easy to interpret for most clinicians [10]. As a result, clinicians may be wary of ML-based algorithms and reluctant to adopt them in practice [13]. This may also have to do with the fact that clinicians owe their patients explanations as to how certain recommendations were derived. Patients may, in turn, be more likely to follow recommendations regarding stroke prevention if they receive a clear explanation of why certain

prevention measures (e.g., exercise regime, medication) are preferable over others in their particular situation. Even though concepts like AI explainability, interpretability, and transparency have gained traction in the scientific community, there is a need for strengthening cooperation among medical practitioners and data scientists to tackle these issues in a collaborative manner [13].

There is also uncertainty regarding who can be held liable for adverse events that result from the use of ML-based algorithms. This uncertainty may, in turn, hamper trust and impede the adoption of these technologies in practice [75]. This point is also linked to clinical validation and efficacy. To foster trust in ML-based algorithms, data scientists and researchers have to show that their algorithms yield accurate predictions and that they can be integrated into clinical practice securely and efficiently for the benefit of patients [10]. In the case of stroke risk prediction and prevention, this means that novel ML-based approaches will have to compete against established models to win over clinicians' and patients' trust. Clinicians and patients, in turn, will have to exercise good judgment about what and whom to trust.

---

## 5.7 Conclusion

Novel ML-driven approaches to stroke risk prediction allow researchers to overcome some of the challenges frequently associated with traditional risk prediction models. Capitalizing on the advantages of ML, physicians, and researchers will also be able to predict more accurately which type of interventions will be most effective for which groups of people. This will, in turn, help them to provide patients with tailored recommendations based on their natural predisposition, empowering them to reduce their individual risk of suffering a stroke. Yet, while ML methods offer unprecedented opportunities for precision medicine in stroke prevention, several technological and methodological challenges remain. As outlined in this chapter, challenges can be grouped into three broad categories: (1) challenges in data sourcing, (2) challenges in application development, (3) challenges in deployment in clinical practice.

Having identified some of the opportunities and challenges of machine learning in stroke risk prediction and prevention, it is time to ask ourselves what impact these dynamics will have on individuals and the delivery of care, more generally. Even though it will certainly take some time before ML-based tools can (at least partially) replace established approaches for stroke risk assessment and prevention, we should already prepare for the questions that will arise as these applications are broadly adopted in practice: how will they impact the doctor-patient relationship? How will they affect public trust in the healthcare system? As great strides are made in precision medicine for stroke, how can we ensure everyone will benefit from these gains—what about low- to middle-income countries where stroke incidence rates exceed those observed in high-income countries? What about individuals who refuse to have their data collected and analyzed? These and several other questions raise important ethical concerns that require further investigation. Only by committing to ethical conduct, methodological rigor, and patient safety will we harness the full potential of data-driven predictive modeling in stroke.

**Funding** This work was supported by funding from the European Union's Horizon 2020 research and innovation program under grant agreement No. 777107 (PRECISE4Q).

**Acknowledgments** The author would like to thank Vince I Madai and Stephanie Bishop for critically reviewing this chapter.

---

## References

1. Mesko B. The role of artificial intelligence in precision medicine. *Exp Rev Precis Med Drug Develop.* 2017;2(5):239–41. <https://doi.org/10.1080/23808993.2017.1380516>.
2. Huang BE, Mulyasasmita W, Rajagopal G. The path from big data to precision medicine. *Exp Rev Precis Med Drug Develop.* 2016;1(2):129–43.
3. Jiang F, Jiang Y, Zhi H, Dong Y, Li H, Ma S, et al. Artificial intelligence in healthcare: past, present and future. *Stroke Vasc Neurol.* 2017;2(4):230–43.
4. Patel UK, Anwar A, Saleem S, Malik P, Rasul B, Patel K, et al. Artificial intelligence as an emerging technology in the current care of neurological disorders. *J Neurol.* 2019:1–20.
5. Jha S, Topol EJ. Adapting to artificial intelligence: radiologists and pathologists as information specialists. *JAMA.* 2016;316(22):2353–4.
6. Pesapane F, Codari M, Sardanelli F. Artificial intelligence in medical imaging: threat or opportunity? Radiologists again at the forefront of innovation in medicine. *Eur Radiol Exp.* 2018;2(1):35.
7. Attia ZI, Noseworthy PA, Lopez-Jimenez F, Asirvatham SJ, Deshmukh AJ, Gersh BJ, et al. An artificial intelligence-enabled ECG algorithm for the identification of patients with atrial fibrillation during sinus rhythm: a retrospective analysis of outcome prediction. *Lancet.* 2019;394(10201):861–7.
8. Tran BX, Vu GT, Ha GH, Vuong Q-H, Ho M-T, Vuong T-T, et al. Global evolution of research in artificial intelligence in health and medicine: a bibliometric study. *J Clin Med.* 2019;8(3):360.
9. Ienca M, Ferretti A, Hurst S, Puhan M, Lovis C, Vayena E. Considerations for ethics review of big data health research: a scoping review. *PLoS One.* 2018;13(10):e0204937.
10. Saber H, Somai M, Rajah GB, Scalzo F, Liebeskind DS. Predictive analytics and machine learning in stroke and neurovascular medicine. *Neurol Res.* 2019;41(8):681–90.
11. Sakai K, Yamada K. Machine learning studies on major brain diseases: 5-year trends of 2014–2018. *Jpn J Radiol.* 2019;37(1):34–72.
12. Feng R, Badgeley M, Mocco J, Oermann EK. Deep learning guided stroke management: a review of clinical applications. *J Neurointervent Surg.* 2018;10(4):358–62.
13. Jamthikar A, Gupta D, Khanna NN, Araki T, Saba L, Nicolaides A, et al. A special report on changing trends in preventive stroke/cardiovascular risk assessment via B-mode ultrasonography. *Curr Atheroscler Rep.* 2019;21(7):25.
14. Feigin VL, Norrving B, George MG, Foltz JL, Roth GA, Mensah GA. Prevention of stroke: a strategic global imperative. *Nat Rev Neurol.* 2016;12(9):501.
15. Feigin VL, Abajobir AA, Abate KH, Abd-Allah F, Abdulle AM, Abera SF, et al. Global, regional, and national burden of neurological disorders during 1990–2015: a systematic analysis for the Global Burden of Disease Study 2015. *The Lancet Neurol.* 2017;16(11):877–97.
16. Feigin VL. Anthology of stroke epidemiology in the 20th and 21st centuries: assessing the past, the present, and envisioning the future. *Int J Stroke.* 2019;14(3):223–37.
17. Feigin VL, Krishnamurthi RV, Parmar P, Norrving B, Mensah GA, Bennett DA, et al. Update on the global burden of ischemic and hemorrhagic stroke in 1990–2013: the GBD 2013 study. *Neuroepidemiology.* 2015;45(3):161–76.
18. Feigin VL, Lawes CM, Bennett DA, Barker-Collo SL, Parag V. Worldwide stroke incidence and early case fatality reported in 56 population-based studies: a systematic review. *Lancet Neurol.* 2009;8(4):355–69.

19. Luengo-Fernandez R, Violato M, Candio P, Leal J. Economic burden of stroke across Europe: a population-based cost analysis. *Eur Stroke J.* 2020;5(1):17–25.
20. Roth GA, Forouzanfar MH, Moran AE, Barber R, Nguyen G, Feigin VL, et al. Demographic and epidemiologic drivers of global cardiovascular mortality. *N Engl J Med.* 2015;372(14):1333–41.
21. Di Carlo A. Human and economic burden of stroke. Oxford University Press; 2009.
22. Graven C, Sansonetti D, Moloczij N, Cadilhac D, Joubert L. Stroke survivor and carer perspectives of the concept of recovery: a qualitative study. *Disabil Rehabil.* 2013;35(7):578–85.
23. Forsberg-Wärleby G, Möller A, Blomstrand C. Psychological Well-being of spouses of stroke patients during the first year after stroke. *Clin Rehabil.* 2004;18(4):430–7.
24. Hill V. Live well after stroke: methods of a community-based, occupational therapist-led, life management intervention. *Ann Phys Rehabil Med.* 2018;61:e514.
25. Redfern J, Gordon C, Cadilhac D. Longer-term support for survivors of stroke and their carers. *Stroke Nurs.* 2019;2:323–45.
26. Wray F, Clarke D. Longer-term needs of stroke survivors with communication difficulties living in the community: a systematic review and thematic synthesis of qualitative studies. *BMJ Open.* 2017;7(10):e017944.
27. Pindus DM, Mullis R, Lim L, Wellwood I, Rundell AV, Aziz NAA, et al. Stroke survivors' and informal caregivers' experiences of primary care and community healthcare services—a systematic review and meta-ethnography. *PLoS One.* 2018;13(2):e0192533.
28. Rajscic S, Gothe H, Borba H, Sroczyński G, Vujicic J, Toell T, et al. Economic burden of stroke: a systematic review on post-stroke care. *Eur J Health Econ.* 2019;20(1):107–34.
29. Mozaffarian D, Benjamin E, Go A, Arnett D, Blaha M, Cushman M, et al. Heart disease and stroke statistics-2016 update: a report from the American Heart Association. *Circulation.* 2016;133(4):e38.
30. Saka Ö, McGuire A, Wolfe C. Cost of stroke in the United Kingdom. *Age Ageing.* 2009;38(1):27–32.
31. Benjamin EJ, Muntner P, Bittencourt MS. Heart disease and stroke statistics-2019 update: a report from the American Heart Association. *Circulation.* 2019;139(10):e56–e528.
32. Mendis S, Davis S, Norrving B. Organizational update: the world health organization global status report on noncommunicable diseases 2014; one more landmark step in the combat against stroke and vascular disease. *Stroke.* 2015;46(5):e121–e2.
33. Mendis S, Armstrong T, Bettcher D, Branca F, Lauer J, Mace C, et al. Global status report on noncommunicable diseases 2014. World Health Organization; 2014.
34. Aarli J, Tarun D, Janca A, Muscetta A. Neurological disorders: public health challenges. World Health Organization; 2006.
35. Meschia JF, Bushnell C, Boden-Albala B, Braun LT, Bravata DM, Chaturvedi S, et al. Guidelines for the primary prevention of stroke: a statement for healthcare professionals from the American Heart Association/American Stroke Association. *Stroke.* 2014;45(12):3754–832.
36. Goldstein LB, Bushnell CD, Adams RJ, Appel LJ, Braun LT, Chaturvedi S, et al. Guidelines for the primary prevention of stroke: a guideline for healthcare professionals from the American Heart Association/American Stroke Association. *Stroke.* 2011;42(2):517–84.
37. World Health Organization. Prevention of cardiovascular disease. World Health Organization; 2007.
38. Boehme AK, Esenwa C, Elkind MS. Stroke risk factors, genetics, and prevention. *Circ Res.* 2017;120(3):472–95.
39. O'donnell MJ, Xavier D, Liu L, Zhang H, Chin SL, Rao-Melacini P, et al. Risk factors for ischaemic and intracerebral haemorrhagic stroke in 22 countries (the INTERSTROKE study): a case-control study. *Lancet.* 2010;376(9735):112–23.
40. Rose G. Sick individuals and sick populations. *Int J Epidemiol.* 1985;14(1):32–8.
41. Feigin VL, Krishnamurthi R, Bhattacherjee R, Parmar P, Theadom A, Hussein T, et al. New strategy to reduce the global burden of stroke. *Stroke.* 2015;46(6):1740–7.
42. Parmar P, Krishnamurthi R, Ikram MA, Hofman A, Mirza SS, Varakin Y, et al. The Stroke Riskometer(TM) App: Validation of a data collection tool and stroke risk predictor. *Int J Stroke.* 2015;10(2):231–44.

43. Feigin VL, Brainin M, Norrving B, Gorelick PB, Dichgans M, Wang W, et al. What is the best mix of population-wide and high-risk targeted strategies of primary stroke and cardiovascular disease prevention? *J Am Heart Assoc.* 2020;9(3):e014494.
44. Feigin VL, Norrving B, Mensah GA. Primary prevention of cardiovascular disease through population-wide motivational strategies: insights from using smartphones in stroke prevention. *BMJ Glob Health.* 2017;2(2):e000306.
45. Diener A, Celemin-Heinrich S, Wegscheider K, Kolpatzik K, Tomaschko K, Altiner A, et al. In-vivo-validation of a cardiovascular risk prediction tool: the Arriba-pro study. *BMC Fam Pract.* 2013;14:7. <https://doi.org/10.1186/1471-2296-14-13>.
46. Weng SF, Reps J, Kai J, Garibaldi JM, Qureshi N. Can machine-learning improve cardiovascular risk prediction using routine clinical data? *PLoS One.* 2017;12(4):e0174944.
47. Jamthikar A, Gupta D, Khanna NN, Saba L, Araki T, Viskovic K, et al. A low-cost machine learning-based cardiovascular/stroke risk assessment system: integration of conventional factors with image phenotypes. *Cardiovas Diagn Ther.* 2019;9(5):420.
48. D'agostino RB, Vasan RS, Pencina MJ, Wolf PA, Cobain M, Massaro JM, et al. General cardiovascular risk profile for use in primary care. *Circulation.* 2008;117(6):743–53.
49. Nobel L, Mayo NE, Hanley J, Nadeau L, Daskalopoulou SS. MyRisk\_Stroke calculator: a personalized stroke risk assessment tool for the general population. *J Clin Neurol.* 2014;10(1):1–9.
50. Goldstein BA, Navar AM, Carter RE. Moving beyond regression techniques in cardiovascular risk prediction: applying machine learning to address analytic challenges. *Eur Heart J.* 2017;38(23):1805–14.
51. Kakadiaris IA, Vrigkas M, Yen AA, Kuznetsova T, Budoff M, Naghavi M. Machine learning outperforms ACC/AHA CVD risk calculator in MESA. *J Am Heart Assoc.* 2018;7(22):e009476.
52. Garg N, Muduli SK, Kapoor A, Tewari S, Kumar S, Khanna R, et al. Comparison of different cardiovascular risk score calculators for cardiovascular risk prediction and guideline recommended statin uses. *Indian Heart J.* 2017;69(4):458–63.
53. Vieira S, Pinaya WHL, Mechelli A. Introduction to machine learning. In: *Machine learning.* Elsevier; 2020. p. 1–20.
54. Rajkomar A, Dean J, Kohane I. Machine learning in medicine. *N Engl J Med.* 2019;380(14):1347–58.
55. Beam AL, Kohane IS. Big data and machine learning in healthcare. *JAMA.* 2018;319(13):1317–8.
56. Olesen AE, Grønlund D, Gram M, Skorpen F, Drewes AM, Klepstad P. Prediction of opioid dose in cancer pain patients using genetic profiling: not yet an option with support vector machine learning. *BMC Res Notes.* 2018;11(1):78.
57. Ngiam KY, Khor W. Big data and machine learning algorithms for health-care delivery. *Lancet Oncol.* 2019;20(5):e262–e73.
58. Ambale-Venkatesh B, Yang X, Wu CO, Liu K, Hundley WG, McClelland R, et al. Cardiovascular event prediction by machine learning: the multi-ethnic study of atherosclerosis. *Circ Res.* 2017;121(9):1092–101.
59. Liu T, Fan W, Wu C. A hybrid machine learning approach to cerebral stroke prediction based on imbalanced medical dataset. *Artif Intell Med.* 2019;101:101723.
60. Li X, Liu H, Du X, Zhang P, Hu G, Xie G, et al, editors. Integrated machine learning approaches for predicting ischemic stroke and thromboembolism in atrial fibrillation. In: *AMIA annual symposium proceedings.* American Medical Informatics Association; 2016.
61. Petosa R. Using behavioral contracts to promote health behavior change: application in a college level health course. *Health Educ.* 1984;15(2):22–7.
62. Lira M, Kunstmann S, Caballero E, Guarda E, Villarroel L, Molina J. Cardiovascular prevention and attitude of people towards behavior changes: state of the art. *Revista Medica de Chile.* 2006;134(2):223–30.
63. Garrido P, Aldaz A, Vera R, Calleja M, de Alava E, Martín M, et al. Proposal for the creation of a national strategy for precision medicine in cancer: a position statement of SEOM, SEAP, and SEFH. *Clin Transl Oncol.* 2018;20(4):443–7.

64. Kökciyan N, Chapman M, Balatsoukas P, Sassoon I, Essers K, Ashworth M, et al. A collaborative decision support tool for managing chronic conditions. *Stud Health Technol Inform.* 2019;264:644–8.
65. Kario K. Perfect 24-h management of hypertension: clinical relevance and perspectives. *J Hum Hypertens.* 2017;31(4):231–43.
66. Li KHC, White FA, Tipoe T, Liu T, Wong MC, Jesuthasan A, et al. The current state of mobile phone apps for monitoring heart rate, heart rate variability, and atrial fibrillation: narrative review. *JMIR Mhealth Uhealth.* 2019;7(2):e11606.
67. Lowres N, Neubeck L, Salkeld G, Krass I, McLachlan AJ, Redfern J, et al. Feasibility and cost-effectiveness of stroke prevention through community screening for atrial fibrillation using iPhone ECG in pharmacies. *Thromb Haemost.* 2014;111(06):1167–76.
68. Tran V-T, Riveros C, Ravaud P. Patients' views of wearable devices and AI in healthcare: findings from the ComPaRe e-cohort. *NPJ Digit Med.* 2019;2(1):1–8.
69. PRECISE4Q Consortium. PRECISE4Q: predictive modelling in stroke. 2020. [www.precise4q.eu](http://www.precise4q.eu). Accessed 25 Mar 2020.
70. Frey D. Schlaganfallbehandlung: Künstliche Intelligenz als Game-Changer. *kma-Das Gesundheitswirtschaftsmagazin.* 2018;23(11):32–4.
71. CORDIS EU Reserach Results. Personalised medicine by predictive modeling in stroke for better quality of life. 2020. <https://cordis.europa.eu/project/id/777107>. Accessed 23 Mar 2020.
72. Shoenbill K, Song Y, Craven M, Johnson H, Smith M, Mendonca EA. Identifying patterns and predictors of lifestyle modification in electronic health record documentation using statistical and machine learning methods. *Prev Med.* 2020;136:106061.
73. Shoenbill K, Song Y, Gress L, Johnson H, Smith M, Mendonca EA. Natural language processing of lifestyle modification documentation. *Health Informatics J.* 2020;26(1):388–405.
74. Liu F, Weng C, Yu H. Natural language processing, electronic health records, and clinical research. In: *Clinical research informatics.* London: Springer; 2012. p. 293–310.
75. Vayena E, Blasimme A, Cohen IG. Machine learning in medicine: addressing ethical challenges. *PLoS Med.* 2018;15(11):e1002689.
76. He J, Baxter SL, Xu J, Xu J, Zhou X, Zhang K. The practical implementation of artificial intelligence technologies in medicine. *Nat Med.* 2019;25(1):30–6.
77. Blasimme A, Fadda M, Schneider M, Vayena E. Data sharing for precision medicine: policy lessons and future directions. *Health Aff.* 2018;37(5):702–9.
78. McCall B. What does the GDPR mean for the medical community? *Lancet.* 2018;391(10127):1249.
79. Fortier I, Doiron D, Burton P, Raina P. Invited commentary: consolidating data harmonization—how to obtain quality and applicability? *Am J Epidemiol.* 2011;174(3):261–4.
80. Fortier I, Raina P, Van den Heuvel ER, Griffith LE, Craig C, Saliba M, et al. Maelstrom research guidelines for rigorous retrospective data harmonization. *Int J Epidemiol.* 2017;46(1):103–5.
81. PRECISE4Q Consortium. How to tackle the challenges of Data Integration. In: PRECISE4Q: predictive modelling in stroke. 2020. <https://precise4q.eu/how-to-tackle-the-challenges-of-data-integration>. Accessed 24 May 2021.
82. Aysolmaz B, Iren D, Dau N, editors. Preventing algorithmic bias in the development of algorithmic decision-making systems: a Delphi study. In: *Proceedings of the 53rd Hawaii International Conference on System Sciences.* 2020.
83. Wong P-H. Democratizing algorithmic fairness. *Philos Technol.* 2019:1–20.
84. Luxtona DD. Ethical implications of conversational agents in global public health. *Bull World Health Organ.* 2020;98:285–7.
85. Baron AA, Baron SB. High levels of HDL cholesterol do not predict protection from cardiovascular disease in women. *Prev Cardiol.* 2007;10(3):125–7.
86. Lau ES, Sarma A. Utility of imaging in risk stratification of chest pain in women. *Curr Treat Options Cardiovasc Med.* 2017;19(9):72.
87. Wu Y, Fang Y. Stroke prediction with machine learning methods among older Chinese. *Int J Environ Res Public Health.* 2020;17(6):1828.

88. Krawczyk B. Learning from imbalanced data: open challenges and future directions. *Prog Artif Intell.* 2016;5(4):221–32.
89. Hung C-Y, Chen W-C, Lai P-T, Lin C-H, Lee C-C, editors. Comparing deep neural network and other machine learning algorithms for stroke prediction in a large-scale population-based electronic medical claims database. In: 2017 39th annual international conference of the IEEE engineering in medicine and biology society (EMBC). IEEE; 2017.
90. Obermeyer Z, Powers B, Vogeli C, Mullainathan S. Dissecting racial bias in an algorithm used to manage the health of populations. *Science.* 2019;366(6464):447–53.
91. Higgins D, Madai VI. From bit to bedside: a practical framework for artificial intelligence product development in healthcare. *Adv Intell Syst.* 2020;2(10) <https://doi.org/10.1002/aisy.202000052>.
92. Siau K, Wang W. Building trust in artificial intelligence, machine learning, and robotics. *Cutt Bus Technol J.* 2018;31(2):47–53.
93. Rudin C. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nat Mach Intell.* 2019;1(5):206–15.

**Open Access** This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

