Journal of Cancer Research & Therapy

Taylor MA et al., J Cancer Res Ther. 2021, 9(1):1-11
http://dx.doi.org/10.14312/2052-4994.2021-1

NobleResearch
An Open Access Publisher
www.nobleresearch.org

Review

Open Access

Advances in artificial neural networks as a disease prediction tool

Matthew A. Taylor[1], Charles L. Bennett[2], Martin W. Schoen[3], and Shamia Hoque[4,*]

[1]School of Medicine, University of South Carolina, Columbia, SC 29209, USA
[2]College of Pharmacy, University of South Carolina, Columbia, SC 29208, USA
[3]Department of Internal Medicine, Saint Louis University, Saint Louis, MO 63103, USA
[4]College of Engineering and Computing, University of South Carolina, Columbia, SC 29201, USA

Abstract

Throughout the last decade, utilization of machine learning has seen a sharp rise in fields such as computing, transportation, engineering, and medicine. Artificial neural networks (ANNs) have demonstrated increased application due to their versatility and ability to learn from large datasets. The emergence of electronic health records has propelled healthcare into an era of personalized medicine largely aided by computers. This review summarizes the current state of ANNs as a predictive tool in medicine and the downfalls of reliance on a self-adjusting computer network to make healthcare decisions. Medical ANN studies can be grouped into three categories - Diagnosis, Classification, and Prediction, with diagnostic studies currently dominating the field. However, recent trends show prediction studies may soon outnumber the remaining categories. ANN prediction studies dominate in fields such as cardiovascular disease, neurologic disease, and osteoporosis. Neural networks consistently show higher predictive accuracy than industry standards. But several pitfalls are preventing mainstream adoption. Clinicians often rely on situational pearls to make complex healthcare decisions, ANNs often do not account for intuitive variables during their analysis. Instead, ANNs rely on incomplete patient data and 'black box' computing to make decisions that are not completely transparent to the end-user. This has led to 'runaway' networks that may ultimately make inaccurate and harmful decisions. This review emphasizes the extensive potential of machine learning in medicine and the obstacles that must be overcome to utilize its full potential.

Keywords: neural network; artificial intelligence; machine learning; disease prediction; black box

1.0 – Introduction

Artificial intelligence (AI) is the capability of computers to learn from their environment and adapt over time to achieve a goal without human input. Also known as 'machine learning', computers have been using this concept for decades to advance products in a variety of fields. GPS navigation, virtual assistants, social media services, and search engines are all examples of services that rely heavily on machine learning to enhance their product [1–3]. A subset of machine learning, known as 'artificial neural networks' (ANN), has become one of the most popular machine learning modalities. ANNs are modeled after the neuronal networks that make up the human nervous system. ANNs gather information through a series of inputs, process those inputs according to their relative importance, and make a determination based on an assigned goal.

In medicine, ANNs are valued for their ability to quickly process vast amounts of information. Their medical use can be divided into three general categories: (1) Diagnosis (2) Classification (3) Prediction. Examples include medical image analysis in radiology, classifications of tumor biopsies, and models to predict the likelihood of coronary artery disease development [4]. ANN applications in medical diagnosis and classification have been well established for decades, but disease prediction remains in its early stages. This review focuses on the recent advances of ANNs in disease prediction, its pitfalls, and future pathways towards advancement.

1.1 – Artificial neural network design

An artificial neural network is a processing algorithm that is modeled after neurological pathways to emulate the human learning process, hence the name 'neural network'. Like the human brain, ANNs are powered by a series of external inputs. Each input represents a different attribute related to the desired output [5]. For example, when identifying a cat [output] one must first recognize

***Corresponding author:** Shamia Hoque, PhD, College of Engineering and Computing, University of South Carolina, 300 Main St., Columbia, SC 29201, USA. Tel.: 803-777-2288; Email: hoques@cec.sc.edu

Received 19 February 2021 *Revised* 11 May 2021 *Accepted* 19 May 2021 *Published* 26 May 2021

Citation: Taylor MA, Bennett CL, Schoen MW, Hoque S. Advances in artificial neural networks as a disease prediction tool. J Cancer Res Ther. 2021; 9(1):1-11. DOI: 10.14312/2052-4994.2021-1

attributes such as fur patterns, body shape, size, and tail [inputs]. Each input attribute is a node in the input layer. This layer is connected to the output via a hidden layer with adjustable number of nodes (Figure 1a). Weights or coefficients assigned to the nodes are adjusted depending on its estimated correlation with the desired output [6]. It is expected that fur patterns and body shape, which are the attributes that most easily identify cats, would be assigned the greatest weight. Size and tail are common to both cats and small dogs, hence they would be assigned 'smaller' weights comparatively [6].

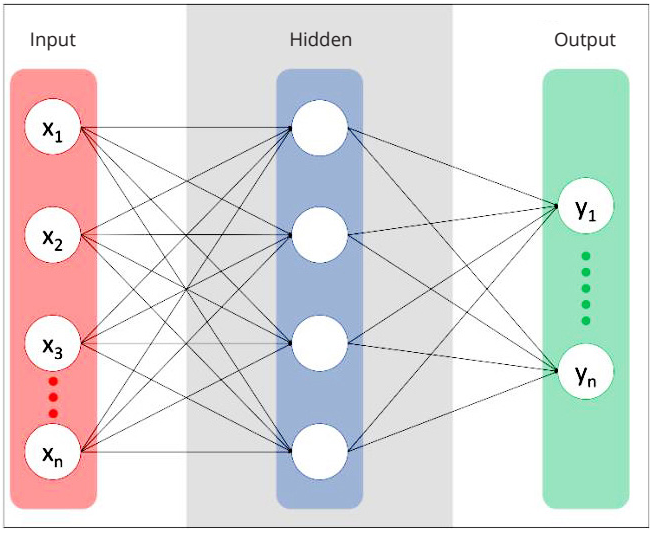

Figure 1a Architecture of a simple artificial neural network. Training data is entered into the neural network as input values [red circles] depicted by 'X_n'. Inputs are connected to nodes [blue circles] within the hidden layer [blue rectangle]. The hidden layer is often referred to as the 'black box' [gray rectangle]. The number of nodes is variable and determined by the user. The neural network assigns each node a bias, which determines how powerful it believes each particular node is at determining the desired output value, depicted by 'Y_n'. Each node connects to an available output [green circle], which is assigned by the user before the neural network is trained.

Similar to humans, ANNs must first learn which inputs likely correspond to an output. There are three different learning mechanisms that can be utilized: (1) supervised learning (2) unsupervised learning (3) reinforcement learning [7]. Initially, each input node is multiplied with an arbitrary weight. At each node of the hidden layer, these weighted inputs are summed, and a bias added. The combined input is passed through an activation function, propagated through the ANN and an output is produced with varying accuracy. The ANN then adjusts the input weights to compensate for the output errors recorded during the previous processing cycle. Time spent learning and degree of input weight adjustment are determined through the learning rate & momentum, respectively. Each additional training cycle and corresponding network adjustments will continue to improve the output accuracy of the ANN. These cycles are continued until the ANN is adequately trained and demonstrates sufficient accuracy to the preference of the user [5].

1.2 – Deep neural networks

Deep neural networks (DNN) are extensions of basic ANNs that are used to solve increasingly complex problems. Any ANN with more than three layers [input, hidden, and output] are considered DNNs (Figure 1b). Deep learning involves the development of algorithms that are more generalizable as opposed to task-specific [8]. The utility of deep learning for analyzing large amounts of data can also be its source of limitation - it requires data. Healthcare data is strictly protected and often constrained within a particular medical system [9].

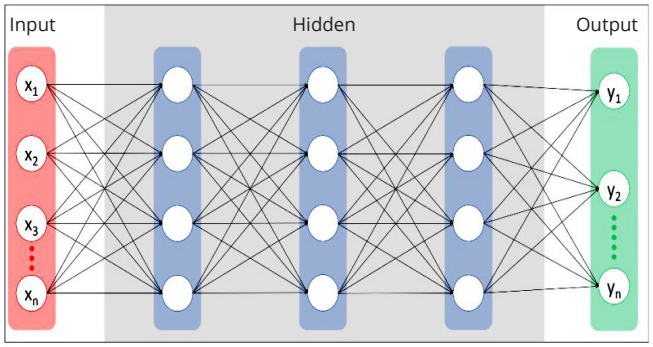

Figure 1b Architecture of a deep neural network. Each node within a hidden layer assigns a bias based on how powerful the neural network determines its data is at determining an output. The biased node then connects to the next hidden layer. This process repeats until the final hidden layer connects to an output, which is determined by the user. A neural network with > 1 hidden layers or > 3 total layers (input, hidden, output) is considered a deep neural network.

For example, a deep recurrent neural network was designed to predict severe post-operative complications including mortality, renal failure, and hemorrhage after cardiothoracic surgery. A database containing 42,007 was used and yielded positive predictive values between 0.84 - 0.9 [10]. Another study assessed the performance of different predictive modeling tools, including deep learning, in the urgent care setting. The models utilized a database of 58,976 unique patients to analyze medical histories, physiological time series, and demographics data to predict mortality, disease differentials, and disease markers. The deep learning recurrent neural network out-performed other standard machine learning models, random forest, gradient boost classifier, and feed-forward multilayer NN for mortality prediction and diagnostics [11-23].

2.0 – Disease prediction

A survey of the literature from the past twenty years confirms the increasing trend of the use of neural networks. Studies can be grouped into three categories - Diagnosis, Classification, and Prediction with the 'Diagnosis' category dominating (Figure 2). While there is little difference between the number of total studies focusing on 'Classification' or 'Prediction' over the last decade, investigations utilizing neural nets for 'Prediction' is trending upwards. Notably, total 'Prediction' studies in the most recent year are set to outnumber Diagnosis and Classification studies (Figure 3).

3.1 – Study selection

Study selection was performed through PubMed/ MEDLINE and Google Scholar searches for the terms 'neural network', 'disease', and 'prediction' for the periods of 1999-2019. Search totals and category comparison are visualized in (Figures 2 and 3). Articles were then selected based on (1) clinical relevance (2) study quality (3) English language, and

(4) article accessibility. Study quality was assessed by the presence of overfitting leading to improbable statistical conclusions, poorly described neural network design, and cohorts less than 50 patients. Non-peer reviewed studies, conference abstracts, and unverifiable titles were also excluded.

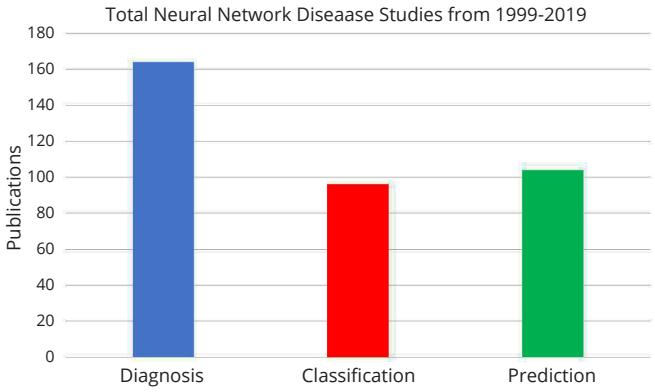

Figure 2 Total neural network prediction publications was determined by a Google Scholar analysis using Publish or Perish software [106]. The terms 'neural' + 'network' + 'disease' + 'diagnosis' or 'classification' or 'prediction' were input into the software. Titles from 1999 – 2019 were filtered by disqualifying non-peer reviewed publications, conference abstracts, and titles without a verifiable link.

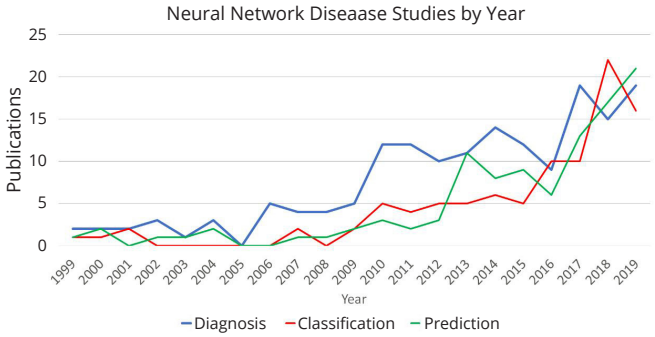

Figure 3 Total neural network prediction publications was determined by a Google Scholar meta-analysis using Publish or Perish software [106]. The titles were filtered by the previously described method in (Figure 2). Remaining titles were separated by year and quantified.

3.2 – Cardiovascular disease

Cardiovascular disease (CVD) is a leading cause of death across the world. Therefore, primary prevention remains a paramount public health initiative [24]. Risk assessments such as the ASCVD Risk Algorithm and the Reynolds Risk Score are used in every day practice to evaluate a patient's risk for CVD [25, 26]. These algorithms use static formulas based on limited risk factors to predict the likelihood of developing CVD. Recent studies have used neural networks to develop personalized CVD prediction models with accuracies that rival established clinical tools [27–29].

A study aiming to predict a first cardiovascular event over 10 years in non-diseased individuals analyzed 30 risk factors with four different machine learning algorithms and compared them to the American Heart Association (AHA) prediction model. They found that amongst all models, neural networks successfully predicted 4,998/7,404 cardiovascular events (sensitivity 67.5%) and 53,458/75,585 of non-cases. This amounts to a 7.6% improvement over the AHA prediction model [27]. Another study compared artificial neural networks versus the Diamond-Forrester (DF) and the Morise models (MM) to stratify the risk of inducible ischemia by cardiac stress testing. The study evaluated 486 patients undergoing radionucleotide and exercise stress testing and found that the ANN had a significant increase in predictive power compared to the DF and MM models. The ANN model demonstrated a sensitivity of 91%, specificity of 65%, PPV of 26%, and NPV of 98%, which ultimately reduced unnecessary stress imaging by 59% [30]. Finally, a study used a deep feed-forward neural network to analyze 5436 patients and predict the one-year mortality in patients diagnosed with an acute myocardial infarction. The model utilized a significant number of variables involving demographics, diagnostics, treatments, and lab values to ultimately arrive at a prediction accuracy of 85.12% and a peak AUC of 0.901 [31].

Medical experts commonly dispute the use of neural networks due to its unknown 'black box' prediction using one or more hidden layers. A study used data from the Korean Centers for Disease Control and Prevention to predict an individual's risk for developing CVD. They first identified each input by using CVD risk factor sensitivities as detected with a previously trained NN. The hidden layer was then bypassed by connecting correlated inputs to the hidden layer in a coupled connection, which avoided the hidden reorganization of the inputs within the layer. The model resulted in a more accurate CVD risk prediction than the compared Framingham Risk Score [32].

3.2.1 – Heart failure

Prolonged cardiovascular disease often leads to advanced heart failure, coronary artery bypass grafting, and prolonged stays in the intensive care unit [33–35]. Researchers in the year 2000 designed a ANN to predict the one-year mortality in 132 patients with heart failure. A network consisting of 62 inputs, 20 hidden nodes, and 3 output nodes successfully predicted 93.2% of one-year deaths, readmissions, and event-free survivals. One-year death rates were predicted with 95.2% sensitivity and 97.8% specificity with an AUC of 0.971. The authors note the results should be interpreted with caution due to the small sample size. Overfitting is likely to have occurred based on comparing the number of adjustable parameters and the sample size (Table 1) [36].

Another study compared an ANN with logistic regression (LR) and Cox Proportional Hazard (CPH) models to predict the incidence of cardiovascular-related deaths in 2,635 patients with heart failure. The model used eight inputs in a feed-forward neural network and had slightly improved predictive abilities compared to the LR and CPH models, with receiver operating characteristic (ROC) curves of 0.72, 0.70, and 0.69, respectively [37]. A later study recognized that deep neural networks had little advantage over logistic regression models when predicting heart failure readmission and were susceptible to overfitting. Instead, a deep unified network (DUN) was designed to predict 30-day readmission rates in 11,510 patients with heart failure. DUNs are deep learning networks that bind each hidden layer together in a mesh-like pattern, allowing all hidden layers to learn directly from the input data. This design will theoretically reduce overfitting. Using AUC as the primary

Table 1 A summary table of the neural network prediction studies described in this review. In studies where multiple sensitivities/specificities were measured, the peak sensitivity value is recorded in the chart.

Topic	Primary author (Year)	*Sample size*	*Machine learning model* [inputs, hidden layer nodes, outputs]	*Adjustable parameters*	*Sensitivity/ Specificity AUC PPV/NPV*	*Accuracy*	*Study limitations*
Cardiovascular disease							
10 Year ASCVD risk	Weng et al. [27]	378,256	Artificial neural network [Not detailed]		67.5% / 70.7% 0.764 18.4% / 95.7%		Black box, overfitting
Inducible ischemia	Isma'eel et al. [30]	486	Artificial neural network [3,3,1]	Biases = 4 Weights = 12 AP = 16	91% / 65% -- 26% / 98%		Sample size, input variables and parameters
Myocardial infarction	Barrett et al. [31]	5,436	Feed forward deep neural network		-- 0.901 --	85.12%	Imbalanced dataset
ASCVD development	Kim et al. [32]	4,146	Artificial neural network [16,4,1]	Biases = 5 Weights = 68 AP = 73	81.5% / * -- 67.6% / 85.1%	81.1%	Study focused on overcoming black box limitations
Heart failure one-year prognosis	Atienza et al. [36]	132	Artificial neural network [62,20,3]	Biases = 23 Weights = 1300 AP = 1323	95.2% / 97.8% 0.971 --	93.2%	Input weights not well defined
Heart failure mortality	Myers et al. [37]	2,635	Artificial neural network [8,7,2]	Biases = 9 Weights = 70 AP = 79	79% / 63% -- --	72%	Data bias, incomplete input variable dataset
Heart failure 30-day hospital readmission	Golas et al. [38]	11,510	Deep neural network [N/A]		-- 0.705 --	76.4%	Overfitting, dataset, generalizability to other health care systems
Stroke							
Ischemic stroke outcome	Heo et al. [44]	2604	Deep neural network [38,15,15,15,2]	Biases = 47 Weights = 1050 AP = 1097	-- 0.888 --		Single-center study, embolectomy patients excluded
Post -tPA outcome	Nielsen et al. [45]	222	Deep convolutional neural network [N/A]		-- 0.88 --		Data bias, dataset
Post-thrombectomy outcome	Asadi et al. [46]	107	Artificial neural network [Not detailed]		-- 0.6 --		Black box, lack of interpretability of models
Ischemic to hemorrhagic stroke transformation	Yu et al. [53]	165	Artificial neural network [Not detailed]		-- 0.693 --	80.7%	Single-center study, small dataset, random sampling
ACA aneurysm rupture risk	Liu et al. [54]	594	Artificial neural network [17,34,34,2]	Biases = 70 Weights = 1802 AP = 1872	95.0% / 92.6% 0.953 --	94.8%	Single-center study, imbalanced dataset, no long-term follow-up
Parkinson disease							
Genetic risk of parkinson disease	Kumudini et al. [59]	306	Artificial neural network [21,*,2]		-- 0.86 --		
CT imaging to predict Parkinson disease progression	Tang et al. [61]	69	Artificial neural network [98,5,2]	Biases = 5 Weights = 900 AP = 905		75%	
Alzheimer disease							
MRI image analysis predicts AD	Lin et al. [70]	818	Deep convolutional neural network [4,32,32,32,32,64,64,1024,2]		84% / 74.8% 0.878 --	79.9%	No long-term follow-up
Clinical score prediction using neuroimaging	Bhagwat et al. [71]	1,606	Series of neural networks in stages				Subjective data as inputs, computing constraints

Osteoporosis							
Bone mineral density assessment	Yoo et al. [83]	1,674	Artificial neural network [Not detailed]		76.6% / 74.4% 0.807 62.9% / 84.8%	75.2%	Dataset, data bias, input variables
Hip fracture risk assessment	Liu et al. [84]	725	Artificial neural network [74,140,1] (F) [74,37,1] (F) [67,140,1] (M) [67,34,1] (M)		-- 0.91 (F)*,0.99 (M) --	85% (F)* 93% (M)	
Adverse drug reactions							
Detecting potential ADRs	Wang et al. [87]	746	Deep neural network [17,1024,512,128,1325]		-- 0.844 --		Limited dataset, model only predicted existing ADRs
Drug molecular analysis to predict ADRs	Dey et al. [89]	[N/A]	Deep convolutional neural network [N/A]		-- 0.957 74.2% / 95.0%	94.5%	No patient data used
Drug-drug interactions	Rohani et al. [90]	2,062	Artificial neural network [7,5,4,2]	Biases = 11 Weights = 63 AP = 74	-- 0.954 – 0.994 --		No patient data used

Abbreviation: PPV = positive predictive value; NPV = negative predictive value; AUC = area under curve; ASCVD = atherosclerotic cardiovascular disease; ACA = anterior communicating artery; *F = females, M = males.

comparison metric, the DUN model had a higher predictive power than logistic regression, gradient boosting, and maxout networks. Ultimately the DUN model successfully predicted 76.4% of 30-day readmissions [38].

Cardiology is a rapidly expanding field that is increasingly seeing favorable results from the use of neural networks. Model details and outcomes for cardiovascular disease have been summarized in (Table 1). Most studies described in the previous section utilized prediction accuracy as a success metric. The success of neural networks in CVD is because CVD has a well-defined series of risk factors that accurately predict a person's risk of disease, including age, cholesterol, blood pressure, weight, and physical activity. CVD is also characterized by defined criteria that dictate the presence and severity of disease. For example, systolic heart failure is defined by the measurement of ejection fraction and prognosis can be predicted by New York Heart Association classifications [39]. Myocardial infarctions are readily detected by an electrocardiogram (EKG), elevated serum troponin, and follow the same risk factors as CVD. Reliance on qualitative interpretations data is less imperative, as easily obtained quantitative patient data can accurately predict a person's risk.

3.3 – Stroke

A stroke is the process of reduced or absent perfusion of blood flow to a region of the brain, causing severe neurological deficits and often death. Generally, about 80% of strokes are ischemic and 20% are hemorrhagic. Treatment of ischemic strokes involve careful assessment of the disease etiology, medical history, imaging with CT or MRI, and potential treatment with thrombolytics or embolectomy. Hemorrhagic strokes are contraindications to thrombolytics, and often require neurosurgical intervention to relieve intracranial pressure and induce hemostasis [40, 41].

Patient outcomes from an ischemic stroke inversely correlate with the amount of infarcted brain tissue [42]. Predicting patient mortality is paramount in directing the treatment strategy. Several studies have used imaging data to train neural networks to predict ischemic stroke outcomes. A retrospective cohort study compared a deep neural network, random forest, and logistic regression models to Acute Stroke Registry and Analysis of Lausanne (ASTRAL) scores at predicting 3 month ischemic stroke outcomes [43]. The machine learning models analyzed 38 inputs and found that the deep neural network significantly outperformed the ASTRAL score, while the random forest and logistic regression models showed little difference. However, when the machine learning models were limited to the same six inputs as the ASTRAL score, the deep neural network performed similarly to the ASTRAL score. The authors conclude that the primary advantage of machine learning models is their ability to analyze large amounts of diverse information to make an outcome prediction [44].

Another study used MRI imaging data to compare various convolutional neural network configurations ability to predict the outcomes of 222 ischemic stroke patients who were treated with tissue plasminogen activator (tPA) versus those who were not treated with tPA. They found the deep CNN had significantly greater predictive capability compared to the other CNNs, with an AUC of 0.88 in the deep CNN and AUCs of 0.85 and 0.72 in the $CNN_{shallow}$ and CNN_{Tmax}, respectively. The investigation concluded that additional layers added to the network improved its ability to predict outcomes [45]. The outcomes of the previously described investigations are further enforced by a study investigating outcomes in patients recovering from post-thrombectomy and a study combining both MRI imaging data and the thrombolysis in cerebral infarction scale. Both show improved predictive power compared to the study controls [46, 47].

Hemorrhagic stroke is a devastating condition resulting in a one-year mortality greater than 50% [48]. Etiologies include hypertension, cerebral amyloid angiopathy, arteriovenous malformations, and hemorrhagic transformation following acute ischemic strokes [49-52]. A study evaluating several machine learning algorithms' ability to predict hemorrhagic transformation severity from ischemic strokes found that the Kernel spectral regression model had a predictive accuracy of 83.7% [53]. Another ANN model was developed to predict anterior communicating artery aneurysm rupture risk. The authors used 13 inputs including computed tomography (CT) angiography data, demographics, and lifestyle risks and found their model successfully had a prediction accuracy of 94.8% in 594 patients [54].

Present day stroke outcome predictions involve a qualitative analysis of many variables including patient demographics, injury severity, and prior comorbidities. The previously described machine learning models mirror that process but enhance the prediction by adding a quantitative dimension. Asadi et al. showed that their machine learning model independently prioritized the NIH Stroke Scale as a leading prognostic factor, which directly corresponds to current medical practices [46]. Heo et al. and Nielsen et al. demonstrated the primary advantage to their machine learning models versus current medical standards is their ability to analyze a higher input volume and complexity [44, 45]. When those inputs were adjusted to match the current standard, there were no differences in predictive capabilities.

3.4 – Parkinson disease

Parkinson disease (PD) is a neurodegenerative disorder that affects around 1 in 500 people over the age of 40 [55]. It involves loss of dopaminergic neurons in the substantia nigra, leading to physical manifestations such as tremor, bradykinesia, rigidity, and postural instability [56]. Advanced age is the most prominent risk factor, but others include traumatic brain injuries, pesticides, certain medications, and genetic predisposition [57, 58]. As average lifespans continue to increase, the prevalence of PD and other neuromuscular disorders are expected to rise. Improved predictive mechanisms can promote early onset therapies and neuroprotective interventions.

Generally, PD arises from a multifactorial etiology. But certain genetic variants have been shown to increase risk of PD. Using 21 single nucleotide polymorphisms (SNPs) related to PD as inputs, a comparative analysis of four different predictive models was performed to assess which model had the highest predictive value for future PD development. Additive, multifactor dimensionality reduction (MDR), recursive partitioning (RP), and artificial neural network (ANN) were all compared. The results showed that the ANN had the highest diagnostic utility with an AUC = 0.86 by utilizing a sigmoid function from hidden to output layer for the categorical output variable. Comparatively, the additive, MDR, and RP models had AUCs of 0.76, 0.69, and 0.82 respectively [59].

The overall prognosis of PD is variable, with advanced age and dementia associated with increased mortality. Studies show varying results, but median survival after initial diagnosis is approximately 7-14 years [60]. Disease progression is typically assessed through its physical symptoms. However, a recent study utilized a trained ANN to analyze dopamine transporter single-photon emission computed tomography images as a potential mechanism to monitor PD disease progression. Analysis of images obtained from regions commonly affected by PD [caudate, putamen, and ventral striatum], they found that the ANN model could predict 4 year disease progression with 75% accuracy [61].

Studies predicting PD highlight the capabilities of neural networks when disease risk factors and variables contributing to prognosis are not well defined. Kumudini et al. analyzed numerous gene sequences and their individual mutations to predict the onset of Parkinson disease [59]. Considering the complexity of an individual's genetic code, analysis of that code and its relationships to other genes in the human genome require complex computing systems like neural networks. Tang et al. found that their neural network independently concluded that brain imaging and the Unified Parkinson's Disease Rating Scale (UPDRS) complemented each other when predicting PD disease progression [61]. Although this conclusion is intuitive to clinicians, this is an example of computers replicating this human intuition.

3.5 – Alzheimer disease

Alzheimer disease (AD) is another disease process that utilizes brain imaging for neural network training. AD is a neurodegenerative disorder that is the most common cause of dementia in the elderly [62]. AD arises through the extracellular deposition of amyloid beta plaques and intracellular deposition of neurofibrillary tangles, causing widespread central nervous system degeneration and subsequent decline in mental function [63]. The etiology of disease onset is multifactorial, but studies have shown that genetics play a major role in age of onset and rate of progression [64–66]. Primary treatment is supportive, but many clinical trials have been conducted aimed at prediction, early detection, and prolonging the duration between diagnosis and complete debilitation [67].

Identifying AD is primarily clinical, but structural MRI images detailing white matter lesions and hippocampal atrophy can aid in the diagnosis [68, 69]. A convolutional neural network (CNN), was created to evaluate hippocampal changes via MRI imaging in patients with mild cognitive impairment to predict future AD conversion with an accuracy of 79.9% [70]. Another study used an ANN to predict Mini Mental State Examination (MMSE) and Alzheimer Disease Assessment Scale (ADAS) scores in AD patients based on hippocampal degeneration and cortical thickness seen in MRI imaging. The authors found their ANN could predict scores with r-values ranging between 0.60 – 0.68 for the MMSE and 0.52 – 0.55 for the ADAS [71] (Table 1).

These studies demonstrate advanced prediction models to differentiate between someone with mild cognitive

impairment from advancing age and someone who is in the early stages of Alzheimer disease. A diagnosis of AD is typically made when symptoms have become too severe to be classified as mild cognitive impairment. But at that stage, patients may no longer be in the ideal clinical therapeutic window to maximally delay severe symptom onset [72]. Studies performed by Lin et al. and Bhagwat et al. used common clinical tools such as the MMSE and MRI imaging to make accurate mathematical predictions of whether a person with mild cognitive impairment will develop into AD [70, 71]. These results could help guide clinicians towards preventative therapies to slow this eventual progression.

3.6 – Osteoporosis

As the average lifespan continues to grow, diseases of aging continue to increase in prevalence [73]. Bone continuously remodels throughout a person's life to maintain its high density and reduce risk of fracture. Osteoclasts resorb old or damaged bone and osteoblasts replace those areas with new, stronger bone [74]. All persons over the age of 30 demonstrate a gradual decline in bone density due to a variety of factors, with postmenopausal women experiencing the most drastic change [75]. This is largely due to decreased estrogen production, which reduces the new bone formation by osteoblasts [76]. Other factors that increase risk for osteoporosis include age, decreased physical activity, genetics, and medications [77–80]. Using these risk factors to predict the likelihood of disease development can allow early preventative measures and reduce fracture-related mortality in elderly patients.

Past studies have found that ANNs outperformed standard statistical models when predicting osteoporosis development [81, 82]. A study analyzed data from the Korea National Health and Nutrition Examination Surveys using various machine learning algorithms to predict osteoporosis risk in postmenopausal women. The study found that support vector machines had the highest predictive value with an AUC of 0.827 compared to the ANN with an AUC of 0.807 [83]. Considering that hip fractures are one of the most significant predictors of mortality in the elderly, an ANN used survey data from 725 respondents to stratify their risk of hip fracture. They found that a simple 3-layer ANN successfully predicted hip fractures in 85% of cases in females and 93% of cases in males. Their model also identified the 10 greatest risk factors for hip fracture by ranking the importance of each input through leave 1 out and connection weight methods. Neither method proved to be better than the other [84].

In osteoporosis, the primary goals of treatment are to slow bone degeneration and prevent bone fractures. Yoo et al. concluded that their neural network excelled at incorporating epidemiological risk factors with bone density measurements to stratify a person's risk of osteoporosis. However, their model was unable to account for concurrent medication use that could also play a large role in bone density [83]. Liu et al. analyzed 74 and 67 risk factors for bone fracture in women and men respectively and used various computational techniques to stratify the top hip fracture risk factors independent of human input. The neural network determined total bone mineral density, declining cognition, and self-assessment of a person's health were leading risk factors leading to hip fracture [84]. These findings directly correlate to current clinical practice, and further reinforce the capability of neural networks to stratify risk factors in disease prediction.

3.7 – Adverse drug reactions

It has been estimated that over 2.2 million people per year in the United States were hospitalized due to adverse drug reactions (ADRs), with approximately 130,000 of those reactions becoming fatal [85]. Prediction and detection of ADRs remains difficult, often relying on time-consuming reporting processes, insufficient data, and professional repercussions [86]. In the last several years, many have proposed utilizing machine learning algorithms to analyze elaborate troves of computerized healthcare data to more successfully predict and prevent future ADRs [87, 88].

Researchers from the National Cheng Kung University in Taiwan designed a deep neural network to predict adverse drug reactions (ADRs). The neural network model utilized a drug's molecular properties, biological effects, and reports of previous ADRs. The results showed the network had a precision of 0.523 and AUC of 0.844 for predicting ADRs. The study was limited due to the ANN model requiring a pre-defined set of ADRs for the ANN to predict. The authors limited their outputs to 1325 possible reactions. However, the true number of possible ADRs may be limitless [87]. Another study analyzed chemical structures commonly associated with ADRs with a deep learning framework. They compared their results to ten other molecular fingerprint models and found the neural networks had the highest predictive utility. Perioral dermatitis was the most successfully predicted ADR with a sensitivity of 74.2%, specificity of 95.0%, and AUC of 0.957 [89]. These results have been further expanded by using a neural network to predict drug-drug interactions. The authors analyzed drug structures, targets, side effects, indications, and molecular pathways. When comparing numerous drug combinations, their model had an AUC between 0.954 – 0.994 for predicting adverse drug interactions [90].

Studies using neural networks to predict adverse drug reactions are scarce. However, these findings are not unexpected. Adverse drug reactions are discovered through the drug development process, involving *in vitro*, *in vivo*, animal, and Phase I-IV Clinical Trials. The process relies on medication trials and careful monitoring of the effects, with the findings of the preceding stages alluding to the findings at each subsequent stage. Similarly, neural networks rely on past data for training. But often that data is incomplete or unknown (e.g., genetics, environment, medical data). Neural networks can predict a defined set of outcomes. But when those outcomes are not defined, neural networks lose their utility. Neither study described utilizes patient-level data or assesses if such data has an influence on the probability of an adverse reaction.

4.0 – Discussion

Neural networks and machine learning are beginning to make their effects known in the healthcare industry. Healthcare systems have already implemented machine learning algorithms to predict a patient's risk for transfer to the intensive care unit (ICU) and EKG strips are often

pre-screened by computer systems to quickly identify pathologic patterns [91, 92]. A recent study identified 29 FDA-approved artificial intelligence-based medical devices and algorithms. Interestingly, only one algorithm involved prediction while the rest focused on diagnosis and classification analysis. These results correlate with the findings in (Figures 2 & 3) indicating that disease detection with machine learning has seen the greatest success in healthcare thus far. But despite its promising potential, machine learning has not taken the massive foothold in everyday medicine, as some had predicted [93].

The representative investigations in this literature review highlight the success, pitfalls, and future potential of neural networks. The aforenoted studies showed both simple and deep neural networks had improved disease prediction when compared to other predictive computer models. Some studies showed the ability of neural networks to independently identify the input variables that had the greatest impact on outcome prediction. But notably, the weights and biases that connect each layer of the network were not adequately described in any of the studies. Another trained network architecture could lead to similar results, but study repetition and subsequent application of the successfully trained neural network would be difficult to implement without knowing the magnitudes each node has on the entirety of the system. In (Table 1), AUC is the most described success metric across all studies. This makes sense as AUC is a better metric than accuracy when the total amount of each measured outcome is drastically skewed. For example, in studies predicting Stroke mortality (Section 3.3) AUC was utilized in all studies while accuracy was only used in two studies. Overall stroke mortality is estimated to be 7.5% [94]. If the neural network picked the outcomes, survival vs. death, by random chance, it would still have a favorable accuracy due to the overall data bias towards one outcome. AUC accounts for data bias by not relying on classifier thresholds, thus discriminating against true positives and false positives in a dataset vulnerable to a high false positive rate.

The size of the neural net architecture varies widely within and across diseases. Investigators emphasize three challenges consistently, (1) lack of context recognition (2) data availability (3) black box computing. The average cohort size for the previously described studies is 1,897 patients when adjusted for outliers. This is an expected number for conventional medical studies, but neural networks thrive in data environments with an abundance of datapoints. Privacy laws and fragmented medical records hinder the ability of neural networks to reach their full predictive potential. Additionally, reproducibility and reliability of neural networks is hindered by the 'black box' phenomenon. Inability to stratify the influence of adjustable parameters, learning rates, and function forms in each network raises a significant barrier to troubleshooting volatile neural networks.

4.1 – Context recognition

Machine learning algorithms excel in environments where pattern recognition holds true. Chronic HIV infection almost always originates from high risk behaviors, presents with eventual immunologic compromise, and warrants itself to highly accurate screening tests [95, 96]. This has made machine learning successful in HIV detection and prevention [97]. But most diseases have varying presentations that rely heavily on context and the personal expertise of clinicians. Cabitza et al. notes that presently, machine learning algorithms are unable to account for the qualitative aspects of medicine and only rely on the interpretation of their quantitative inputs [98]. This point is reinforced when neural networks are used to predict adverse drug reactions. Ignoring context can lead to misleading conclusions. A machine learning model designed to predict mortality in patients with pneumonia versus those with pneumonia and asthma concluded that patients with pneumonia and asthma had a lower risk for mortality. Statistically this result is correct, but this result remains to be untrue from an intuitive sense. The machine learning model was unable to account for asthmatic patients with pneumonia being directly admitted to intensive care units, which resulted in higher levels of care and lower incidence of complications [99].

4.2 – Data availability

Computing has made the analysis of large datasets a manageable reality. Electronic health records have digitized troves of health data that may hold secrets to unlocking many mysteries in medicine. But with most data, maintaining privacy is a priority. The Health Insurance Portability and Accountability Act (HIPAA) has reasonably protected the dissemination of identifying healthcare data. This has made accessibility to healthcare data difficult to attain, leading to fragmented datasets that hinder a neural network's predictive capabilities [100]. For example, a patient with heart failure, diabetes, and a prior stroke is likely receiving care from three separate specialist physicians. If those physicians do not belong to the same hospital system their records are likely not shared with one another, leading to a fragmented medical record for that patient. Several attempts have been made to design networks that compensate for this data discrepancy, with inconsistent results that would not withstand rigorous quality assurance measures required for widespread implementation in healthcare [101, 102]. Results from isolated studies are promising. But without seamless confluence of medical records nation-wide, neural networks may never have the complete data necessary to accurately predict individual patient outcomes.

4.3 – Black box computing

Another important obstacle is the presence of the 'black box' computing that takes place within the hidden layer of an ANN. In fields like medicine where decisions have life-altering consequences, the experts making those decisions value having intricate knowledge of each step required to make those decisions. Neural networks process large amounts of data with a process unknown to both the end-user and developer. Neural networks are meant to adapt over time to improve accuracy, but this can lead to unpredictable behavior patterns that can become dangerous [103]. A recent example is Watson Health by IBM suggesting erroneous cancer treatments. Critics say Watson was unable to adapt to the complexity of individualized oncologic therapy; partly due to incomplete data and lack of context recognition, leading to further

distrust of machine learning in medicine [104]. However, London 2019, argues that many decisions in medicine must be made with uncertainty. Often the value of a beneficial result triumphs knowing the exact mechanistic details behind the result [105].

5.0 – Conclusion

In medicine, machine learning plays the role of a double-edged sword. One perspective views machine learning as a future indispensable tool that improves diagnostic accuracy and clinical efficiency. The other perspective emphasizes caution, knowing that intelligent computer systems lack insight into context and can make technically accurate, but situationally incorrect decisions that may cause harm to patients. Rather than viewing machine learning in medicine as a threat to job security and patient safety, it should be viewed as a powerful tool in its primal stages. Machine learning has the unique capability of analyzing the functional relationships of vast complex medical data. But obstacles such as contextual recognition, fragmented patient data, and black box computing are hindering neural networks from reaching their full potential. Continued research into this tool is imperative to its success, where it may globally improve patient outcomes and completely reshape the field of medicine.

Funding

This work was supported by the American Cancer Society [IRG-17-179-04 to C.B.].

Conflicts of interest

The authors report no conflicts of interest.

References

[1] Ellis K, Godbole S, Marshall S, Lanckriet G, Staudenmayer J, et al. Identifying active travel behaviors in challenging environments using GPS, accelerometers, and machine learning algorithms. Front Public Heal. 2014; 2:36.

[2] Lee K, Caverlee J, Webb S. Uncovering social spammers. Proceeding 33rd Int. ACM SIGIR Conf Res Dev Inf Retr – SIGIR'10. 2010; 435–442.

[3] Chau M, Chen H. A machine learning approach to web page filtering using content and structure analysis. Decis Support Syst. 2008; 44(2):482–494.

[4] Patel JL, Goyal RK. Applications of artificial neural networks in medical science. Curr Clin Pharmacol. 2007; 2(3):217–226.

[5] Jain AK, Jianchang Mao, Mohiuddin KM. Artificial neural networks: a tutorial. Computer (Long Beach Calif). 1996; 29(3):31–44.

[6] Wang SC. Chapter 5: Artificial neural network. Interdiscip Comput Java Program. 2003; 81–100.

[7] Towell GG, Shavlik JW. Knowledge-based artificial neural networks. Artif Intell. 1994; 70(1-2):119–165.

[8] Abatemarco D, Perera S, Bao SH, Desai S, Assuncao B, et al. Training augmented intelligent capabilities for pharmacovigilance: Applying deep-learning approaches to individual case safety report processing. Pharmaceut Med. 2018; 32(6):391–401.

[9] Jiang F, Jiang Y, Zhi H, Dong Y, Li H, et al. Artificial intelligence in healthcare: past, present and future. Stroke Vasc. Neurol. 2017; 2(4):230–243.

[10] Meyer A, Zverinski D, Pfahringer B, Kempfert J, Kuehne T, et al. Machine learning for real-time prediction of complications in critical care: a retrospective study. Lancet Respir Med. 2018; 6(12):905–914.

[11] Tang F, Xiao C, Wang F, Zhou J. Predictive modeling in urgent care: a comparative study of machine learning approaches. JAMIA Open. 2018; 1(1):87–98.

[12] Bascil MS, Oztekin H. A Study on hepatitis disease diagnosis using probabilistic neural network. J Med Syst. 2012; 36(3):1603–1606.

[13] Hirschauer TJ, Adeli H, Buford JA. Computer-aided diagnosis of parkinson's disease using enhanced probabilistic neural network. J Med Syst. 2015; 39(11):179.

[14] Mirzaei G, Adeli A, Adeli H. Imaging and machine learning techniques for diagnosis of Alzheimer's disease. Rev Neurosci. 2016; 27(8):857–870.

[15] Świetlik D, Białowąs J. Application of artificial neural networks to identify alzheimer's disease using cerebral perfusion SPECT data. Int J Environ Res Public Health. 2019; 16(7):1303.

[16] Liu X, Faes L, Kale AU, Wagner SK, Fu DJ, et al. A comparison of deep learning performance against health-care professionals in detecting diseases from medical imaging: a systematic review and meta-analysis. Lancet Digit Heal. 2019; 1(6):e271–e297.

[17] Courtiol P, Maussion C, Moarii M, Pronier E, Pilcer S, et al. Deep learning-based classification of mesothelioma improves prediction of patient outcome. Nat Med. 2019; 25(10):1519–1525.

[18] Acharya UR, Oh SL, Hagiwara Y, Tan JH, Adam M, et al. A deep convolutional neural network model to classify heartbeats. Comput Biol Med. 2017; 89:389–396.

[19] Komeda Y, Handa H, Watanabe T, Nomura T, Kitahashi M, et al. Computer-Aided Diagnosis Based on Convolutional Neural Network System for Colorectal Polyp Classification: Preliminary Experience. Oncology. 2017; 93(Supp 1):30–34.

[20] de Dumast P, Mirabel C, Cevidanes L, Ruellas A, Yatabe M, et al. A web-based system for neural network based classification in temporomandibular joint osteoarthritis. Comput Med Imaging Graph. 2018; 67:45–54.

[21] Ross TR, Ng D, Brown JS, Pardee R, Hornbrook MC, et al. The HMO research network virtual data warehouse: a public data model to support collaboration. EGEMS (Washington, DC) 2014; 2(1):1049.

[22] Shillan D, Sterne JAC, Champneys A, Gibbison B. Use of machine learning to analyse routinely collected intensive care unit data: a systematic review. Crit Care. 2019; 23(1):284.

[23] Raghupathi W, Raghupathi V. Big data analytics in healthcare: promise and potential. Health Inf Sci Syst. 2014; 2:3.

[24] Roth GA, Johnson C, Abajobir A, Abd-Allah F, Abera SF, et al. Global, regional, and national burden of cardiovascular diseases for 10 causes, 1990 to 2015. J Am Coll Cardiol. 2017; 70(1):1–25.

[25] Goff DC, Lloyd-Jones DM, Bennett G, Coady S, D'Agostino RB, et al. 2013 ACC/AHA guideline on the assessment of cardiovascular risk: A report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines. J Am Coll Cardiol. 2014; 63(25 Pt B):2935–2959.

[26] Ridker PM, Buring JE, Rifai N. Development and validation of improved algorithms for the assessment of global cardiovascular risk in women: the Reynolds Risk Score. JAMA. 2007; 297(6):611–619.

[27] 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:e0174944.

[28] Narain R, Saxena S, Goyal AK. Cardiovascular risk prediction: a comparative study of Framingham and quantum neural network based approach. Patient Prefer Adherence. 2016; 10:1259–1270.

[29] Vallée A, Cinaud A, Blachier V, Lelong H, Safar ME, et al. Coronary heart disease diagnosis by artificial neural networks including aortic pulse wave velocity index and clinical parameters. J. Hypertens. 2019; 37(8):1682–1688.

[30] Isma'eel HA, Sakr GE, Serhan M, Lamaa N, Hakim A, et al. Artificial neural network-based model enhances risk stratification and reduces non-invasive cardiac stress imaging compared to Diamond–Forrester and Morise risk assessment models: A prospective study. J Nucl Cardiol. 2018; 25(5):1601–1609.

[31] Barrett LA, Payrovnaziri SN, Bian J, He Z. Building computational models to predict one-year mortality in ICU patients with acute myocardial infarction and post myocardial infarction syndrome. AMIA Jt Summits Transl Sci proceedings. AMIA Jt Summits Transl Sci. 2019; 2019:407–416.

[32] Kim JK, Kang S. Neural network-based coronary heart disease risk prediction using feature correlation analysis. J Healthc Eng. 2017; 2017:2780501.

[33] Low Wang CC, Hess CN, Hiatt WR, et al. Clinical update: Cardiovascular disease in diabetes mellitus: atherosclerotic cardiovascular disease and heart failure in type 2 diabetes mellitus - mechanisms, management, and clinical considerations. Circulation. 2016; 133(24):2459–502.

[34] Gaudino M, Taggart D, Suma H, Puskas JD, Crea F, et al. The choice of conduits in coronary artery bypass surgery. J Am Coll Cardiol. 2015; 66(15):1729–1737.

[35] Loughran J, Puthawala T, Sutton BS, Brown LE, Pronovost PJ, et al. The cardiovascular intensive care unit-an evolving model for health care delivery. J Intensive Care Med. 2017; 32(2):116–123.

[36] Atienza F, Martinez-Alzamora N, De Velasco JA, Dreiseitl S, Ohno-Machado L. Risk stratification in heart failure using artificial neural networks. Proc AMIA Symp. 2000; 32–36.

[37] Myers J, De Souza CR, Borghi-Silva A, Guazzi M, Chase P, et al. A neural network approach to predicting outcomes in heart failure using cardiopulmonary exercise testing. Int J Cardiol. 2014; 171(2):265–269.

[38] Golas SB, Shibahara T, Agboola S, Otaki H, Sato J, et al. A machine learning model to predict the risk of 30-day readmissions in patients with heart failure: A retrospective analysis of electronic medical records data. BMC Med Inform Decis Mak. 2018; 18(1):1–17.

[39] Raphael C, Briscoe C, Davies J, Manisty C, Sutton R, et al. Limitations of the New York Heart Association functional classification system and self-reported walking distances in chronic heart failure. Heart. 2007; 93(4):476–482.

[40] Powers WJ, Rabinstein AA, Ackerson T, Adeoye PM, Bambakidis NC, et al. 2018 Guidelines for the early management of patients with acute ischemic stroke: A guideline for healthcare professionals from the American Heart Association/American Stroke Association. Stroke. 2018; 49(3):e46–e110.

[41] Hemphill JC, Greenberg SM, Anderson CS, Becker K, Bendok BR, et al. Guidelines for the management of spontaneous intracerebral hemorrhage: A guideline for healthcare professionals from the American Heart Association/American Stroke Association. Stroke. 2015; 46(7):2032–2060.

[42] El-Koussy M, Schroth G, Brekenfeld C, Arnold M. Imaging of acute ischemic stroke. Eur Neurol. 2014; 72(5-6):309–316.

[43] Ntaios G, Faouzi M, Ferrari J, Lang W, Vemmos K, et al. An integer-based score to predict functional outcome in acute ischemic stroke: the ASTRAL score. Neurology. 2012; 78(24):1916–1922.

[44] Heo JN, Yoon JG, Park H, Kim YD, Nam HS, et al. Machine learning-based model for prediction of outcomes in acute stroke. Stroke. 2019; 50(5):1263–1265.

[45] Nielsen A, Hansen MB, Tietze A, Mouridsen K. Prediction of tissue outcome and assessment of treatment effect in acute ischemic stroke using deep learning. Stroke. 2018; 49(6):1394–1401.

[46] Asadi H, Dowling R, Yan B, Mitchell P. Machine learning for outcome prediction of acute ischemic stroke post intra-arterial therapy. PLoS One. 2014; 9(2):14–19.

[47] Pinto A, Mckinley R, Alves V, Wiest R, Silva CA, et al. Stroke lesion outcome prediction based on MRI imaging combined with clinical information. Front Neurol. 2018; 9:1–10.

[48] Flaherty ML, Haverbusch M, Sekar P, Kissela B, Kleindorfer D, et al. Long-term mortality after intracerebral hemorrhage. Neurology. 2006; 66(8):1182–1186.

[49] Wang Q, Gao C, Liu H, Li W, Zhao Y, et al. Hypertension modifies the short-term effects of temperature on morbidity of hemorrhagic stroke. Sci Total Environ. 2017; 598:198–203.

[50] Block F, Dafotakis M. Cerebral amyloid angiopathy in stroke medicine. Dtsch Arztebl Int. 2017; 114(3):37–42.

[51] Gross BA, Jankowitz BT, Friedlander RM. Cerebral Intraparenchymal Hemorrhage: A Review. JAMA. 2019; 321(13):1295–1303.

[52] Khatri R, McKinney AM, Swenson B, Janardhan V. Blood-brain barrier, reperfusion injury, and hemorrhagic transformation in acute ischemic stroke. Neurology. 2012; 79(13 Suppl 1):S52–S57.

[53] Yu Y, Guo D, Lou M, Liebeskind D, Scalzo F. Prediction of hemorrhagic transformation severity in acute stroke from source perfusion MRI. IEEE Trans Biomed Eng. 2018; 65(9):2058–2065.

[54] Liu J, Chen Y, Lan L, Lin B, Chen W, et al. Prediction of rupture risk in anterior communicating artery aneurysms with a feed-forward artificial neural network. Eur Radiol. 2018; 28(8):3268–3275.

[55] Tysnes OB, Storstein A. Epidemiology of Parkinson's disease. J Neural Transm. 2017; 124(8):901–905.

[56] Lang AE, Lozano AM. Parkinson's Disease. N Engl J Med. 1998; 339(15):1044–1053.

[57] Bellou V, Belbasis L, Tzoulaki I, Evangelou E, Ioannidis JPA. Environmental risk factors and Parkinson's disease: An umbrella review of meta-analyses. Parkinsonism Relat Disord. 2016; 23:1–9.

[58] Delamarre A, Meissner WG. Epidemiology, environmental risk factors and genetics of Parkinson's disease. Presse Med. 2017; 46(2 Pt 1):175–181.

[59] Kumudini N, Naushad SM, Alex Stanley B, Niveditha M, Sharmila G, et al. Comparative analysis of four disease prediction models of Parkinson's disease. Mol Cell Biochem. 2016; 411(1-2):127–134.

[60] Macleod AD, Taylor KSM, Counsell CE. Mortality in Parkinson's disease: a systematic review and meta-analysis. Mov Disord. 2014; 29(13):1615–1622.

[61] Tang J, Yang B, Adams MP, Shenkov NN, Klyuzhin IS, et al. Artificial neural network–based prediction of outcome in Parkinson's disease patients using DaTscan SPECT imaging features. Mol Imaging Biol. 2019; 21(6):1165–1173.

[62] Ballard C, Gauthier S, Corbett A, Brayne C, Aarsland D, et al. Alzheimer's disease. Lancet. 2011; 377 (9770):1019–1031.

[63] Parihar MS, Hemnani T. Alzheimer's disease pathogenesis and therapeutic interventions. J Clin Neurosci. 2004; 11(5):456–467.

[64] Karch CM, Cruchaga C, Goate AM. Alzheimer's disease genetics: from the bench to the clinic. Neuron. 2014; 83(1):11–26.

[65] Zhu JB, Tan CC, Tan L, Yu JT. State of Play in Alzheimer's disease genetics. J Alzheimers Dis. 2017; 58:631–659.

[66] Cacace R, Sleegers K, Van Broeckhoven C. Molecular genetics of early-onset Alzheimer's disease revisited. Alzheimers Dement. 2016; 12(6):733–748.

[67] The need for early detection and treatment in Alzheimer's disease. EBioMedicine. 2016; 9:1–2.

[68] Richard E, Gouw AA, Scheltens P, van Gool WA. Vascular care in patients with Alzheimer disease with cerebrovascular lesions slows progression of white matter lesions on MRI: the evaluation of vascular care in Alzheimer's disease (EVA) study. Stroke. 2010; 41(3):554–556.

[69] Wahlund LO, Almkvist O, Blennow K, Engedahl K, Johansson A, et al. Evidence-based evaluation of magnetic resonance imaging as a diagnostic tool in dementia workup. Top Magn Reson Imaging. 2005; 16(6):427–437.

[70] Lin W, Tong T, Gao Q, Guo D, Du X, et al. Convolutional neural networks-based MRI image analysis for the Alzheimer's disease prediction from mild cognitive impairment. Front Neurosci. 2018; 12:1–13.

[71] Bhagwat N, Pipitone J, Voineskos AN, Chakravarty MM. An artificial neural network model for clinical score prediction in Alzheimer disease using structural neuroimaging measures. J Psychiatry Neurosci. 2019; 44(4):246–250.

[72] Rasmussen J, Langerman H. Alzheimer's disease - Why we need early diagnosis. Degener Neurol Neuromuscul Dis. 2019; 9:123–130.

[73] Childs BG, Durik M, Baker DJ, van Deursen JM. Cellular senescence in aging and age-related disease: from mechanisms to therapy. Nat Med. 2015; 21(12):1424–1435.

[74] Johnston CC, Slemenda CW. Pathogenesis of osteoporosis. Bone. 1995; 17(2 Suppl):19S-22S.

[75] Runolfsdottir HL, Sigurdsson G, Franzson L, Indridason OS. Gender comparison of factors associated with age-related differences in bone mineral density. Arch Osteoporos. 2015; 10:214.

[76] Khosla S, Oursler MJ, Monroe DG. Estrogen and the skeleton. Trends Endocrinol Metab. 2012; 23(11):576–581.

[77] Ensrud KE, Crandall CJ. Osteoporosis. Ann Intern Med. 2017; 167(3):ITC17–ITC32.

[78] Cai X, Yi X, Zhang Y, Zhang D, Zhi L, et al. Genetic susceptibility of postmenopausal osteoporosis on sulfide quinone reductase-like gene. Osteoporos Int. 2018; 29(9):2041–2047.

[79] Bijelic R, Milicevic S, Balaban J. Risk Factors for Osteoporosis in Postmenopausal Women. Med Arch. 2017; 71(1):25–28.

[80] Coughlan T, Dockery F. Osteoporosis and fracture risk in older people. Clin Med. 2014; 14(2):187–191.

[81] Queraltó JM, Torres J, Guinot M. Neural networks for the biochemical prediction of bone mass loss. Clin Chem Lab Med. 1999; 37(8):831–838.

[82] Chiu JS, Li YC, Yu FC, Wang YF. Applying an artificial neural network to predict osteoporosis in the elderly. Stud Health Technol Inform. 2006; 124:609–614.

[83] Yoo TK, Kim SK, Kim DW, Choi JY, Lee WH, et al. Osteoporosis risk prediction for bone mineral density assessment of postmenopausal women using machine learning. Yonsei Med J. 2013; 54(6):1321–1330.

[84] Liu Q, Cui X, Chou YC, Abbodd MF, Lin J, et al. Ensemble artificial neural networks applied to predict the key risk factors of hip bone fracture for elders. Biomed Signal Process Control. 2015; 21:146–156.

[85] Lazarou J, Pomeranz BH, Corey PN. Incidence of Adverse Drug Reactions in Hospitalized Patients. JAMA 1998; 279(15):1200.

[86] Bennett CL, Schooley B, Taylor MA, Witherspoon BJ, Godwin A, et al. Caveat medicus: Clinician experiences in publishing reports of serious oncology-associated adverse drug reactions. PLoS One 2019; 14(7):1–14.

[87] Wang CS, Lin PJ, Cheng CL, Tai SH, Yang YHK, et al. Detecting potential adverse drug reactions using a deep neural network model. J Med Internet Res. 2019; 21(2):e11016.

[88] Zheng Y, Peng H, Zhang X, Zhao Z, Yin J, et al. Predicting adverse drug reactions of combined medication from heterogeneous pharmacologic databases. BMC Bioinformatics. 2018; 19(Suppl 19):517.

[89] Dey S, Luo H, Fokoue A, Hu J, Zhang P. Predicting adverse drug reactions through interpretable deep learning framework. BMC Bioinformatics 2018; 19(Suppl 21):476.

[90] Rohani N, Eslahchi C. Drug-drug interaction predicting by neural network using integrated similarity. Sci Rep. 2019; 9(1):13645.

[91] Escobar GJ, Turk BJ, Ragins A, Ha J, Hoberman B, et al. Piloting electronic medical record-based early detection of inpatient deterioration in community hospitals. J Hosp Med. 2016; 11(Suppl 1):S18–S24.

[92] Parvaneh S, Rubin J, Babaeizadeh S, Xu-Wilson M. Cardiac arrhythmia detection using deep learning: A review. J Electrocardiol. 2019; 57S:S70–S74.

[93] Ramesh AN, Kambhampati C, Monson JRT, Drew PJ. Artificial intelligence in medicine. Ann R Coll Surg Engl. 2004; 86(5):334–338.

[94] D'Alton M, Coughlan T, Cogan N, Greene S, McCabe DJH, et al. Patterns of mortality in modern stroke care. Ir Med J. 2018; 111(5):750.

[95] Deeks SG, Overbaugh J, Phillips A, Buchbinder S. HIV infection. Nat Rev Dis Prim. 2015; 1:15035.

[96] Hamada Y, Lujan J, Schenkel K, Ford N, Getahun H. Sensitivity and specificity of WHO's recommended four-symptom screening rule for tuberculosis in people living with HIV: a systematic review and meta-analysis. Lancet HIV. 2018; 5(9):e515–e523.

[97] Young SD, Yu W, Wang W. Toward automating HIV identification: Machine learning for rapid identification of HIV-related social media data. J Acquir Immune Defic Syndr. 2017; 74(Suppl 2):S128–S131.

[98] Cabitza F, Rasoini R, Gensini GF. Unintended Consequences of Machine Learning in Medicine. JAMA. 2017; 318(6):517–518.

[99] Caruana R, Lou Y, Gehrke J, Koch P, Sturm M, et al. Intelligible models for healthcare: Predicting pneumonia risk and hospital 30-day readmission. Proc 21th ACM SIGKDD Int Conf Knowl Discov Data Mining. 2015; 1721–1730.

[100] Chen M, Hao Y, Hwang K, Wang L, Wang L. Disease prediction by machine learning over big data from healthcare communities. IEEE Access. 2017; 5:8869–8879.

[101] Mazurowski MA, Habas PA, Zurada JM, Baker JA, Tourassi GD. Training neural network classifiers for medical decision making: The effects of imbalanced datasets on classification performance. Neural Networks. 2008; 21(2-3):427–436.

[102] Khaing HW. Data mining based fragmentation and prediction of medical data. 2011 3rd Int Conf Comput Res Dev. 2011; 480–485.

[103] The Lancet Respiratory Medicine. Opening the black box of machine learning. Lancet Respir Med. 2018; 6(11):801.

[104] Pan H, Tao J, Qian M, Zhou W, Qian Y, et al. Concordance assessment of Watson for oncology in breast cancer chemotherapy: first China experience. Transl Cancer Res. 2019; 8(2):389–401.

[105] London AJ. Artificial intelligence and black-box medical decisions: Accuracy versus Explainability. Hastings Cent Rep. 2019; 49(1):15–21.

[106] Harzing AW. Publish or Perish. 2007.