Annual Methodological Archive Research Review

http://amresearchreview.com/index.php/Journal/about

Volume 3, Issue 11 (2025)

Online ISSN 3007-3197

Print ISSN 3007-3189

http://amresearchreview.com/index.php/Journal/about

Predictive AI Healthcare Analytics for Early Disease Detection Leveraging AI and Machine Learning Models to Identify High-Risk Patients and Improve Preventive Healthcare Strategies

Ahmad Jamal
King Graduate School- Monroe University

Fatima Tauseef
Washington University of Science and Technology

Fahad Naseer
Washington University of Science and Technology

Almas Akbar
Rawalpindi Medical University, Pakistan

Mehwish Jabeen
King Edward Medical University, Pakistan

Fawad Nasim
Faculty of Computer Science and Information Technology, The Superior University, Lahore, 54600, Pakistan

ABSTRACT

With an increase in complexities in the field of healthcare resulting from issues like aging populations, increasing healthcare treatment costs, a shortage of workers and the high number of patients suffering from chronic diseases, predictive artificial intelligence is increasingly becoming important in the future of healthcare. In this article, the author examines the role played by the use of AI predictive healthcare analytics and machine learning methods in helping to identify risks associated with patients and early detection of disease risks. This research is conducted in a secondary research methodology in the form of a literature review of the existing studies on the subject. The paper evaluates several AI machine learning algorithms including logistic regression, decision trees, random forest, support vector machines, gradient boosting, neural networks, deep learning, NLP (natural language processing) and ensemble machine learning models. These AI algorithms are suggested to assist healthcare professionals in understanding the risks of disease among their patients through EHR (electronic health record), lab tests, imaging, devices, lifestyle data and genomics. Advantages of AI include clinical decision making, personalized prevention, improved patient outcome, reduction in hospital admission rates and better allocation of resources. However, there is a question concerning the privacy of information, discrimination through algorithms, non-explainability, false-positive predictions, false-negative predictions, the clinical validity of the tool, and accountability. In conclusion, the author states that

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 3, Issue 11 (2025)
Online ISSN 3007-3197
Print ISSN 3007-3189
http://amresearchreview.com/index.php/Journal/about

predictive AI should be responsibly utilized and support rather than replace health care specialists.

Keywords; Predictive healthcare analytics; artificial intelligence; machine learning; early disease detection; high-risk patients; preventive healthcare; clinical decision support.

1. Introduction

Healthcare systems across the world are facing increasing pressure due to the combined effects of chronic disease, ageing populations, workforce shortages, and rising treatment costs. Conditions such as diabetes, cardiovascular disease, cancer, chronic kidney disease, and respiratory illnesses often develop gradually and place a long-term burden on hospitals, primary care services, and public health systems (Usman, Khan and Butler, 2021). European healthcare systems face growing post-pandemic pressure from workforce shortages, ageing populations, chronic disease, and rising costs. The OECD reported a shortage of 1.2 million doctors, nurses and midwives in the EU in 2022. Ageing is also increasing demand, as people aged 65+ rose from 16% in 2000 to over 21% in 2023, and may reach nearly 30% by 2050. Over 40% of EU citizens aged 65+ have at least two chronic conditions. OECD health spending rose from 7% of GDP in 2000 to almost 9% in 2019, and may reach 11.8% by 2040 without reform, or 10.6% with transformational action.

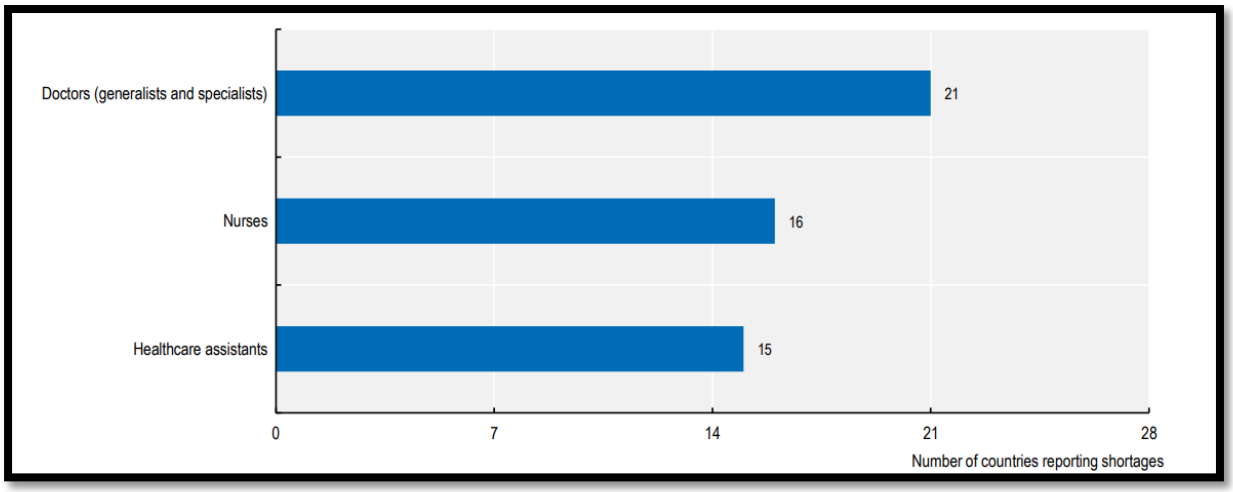


Figure 1 : Healthcare Workforce Shortages Reported Across European Countries
Source : (OECD, 2023)

Conventional health services have tended to be reactive that is, patients have only started treatment when they have symptoms or have taken action for treatment. Although this model is still required when it comes to acute care, it is inappropriate for diseases that are not easily detectable in the early stages (Shang et al., 2020). For instance, somebody with diabetes, cardiovascular disease, cancer, chronic kidney disease or sepsis may have quantifiable risk markers that are present, but not yet accompanied by significant clinical symptoms. Early detection of these risks allows for screening, lifestyle education, medication review, monitoring, and/or specialist referral. For this reason, the benefits of preventive and predictive healthcare are becoming more and more significant.

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 3, Issue 11 (2025)
Online ISSN 3007-3197
Print ISSN 3007-3189
http://amresearchreview.com/index.php/Journal/about

Predictive artificial intelligence is a possible path towards this change. By harnessing AI and machine learning algorithms, scientists are able to work with huge, complex sets of data collected in the healthcare sector like electronic health records, lab test results, medical images, genomic information, data obtained through wearable technology, and other lifestyle data (Nagendran et al., 2020). AI applications are capable of detecting some underlying patterns that could otherwise remain unnoticed using purely clinical reasoning. The fact that even the FDA maintains a list of medical devices featuring AI functionality proves that the use of AI tools is on the rise within the healthcare sector, but only after rigorous clinical verification.

However, it is important not to exaggerate the power of predictive AI. The emergence of technological advancements in AI does not always guarantee the improvement of health care services. Data quality, transparency, fairness, clinical validation, clinical privacy, and proper application within health care services are key components that must be considered in the successful application of AI. As mentioned in the guidelines issued by WHO on AI in health care (Reddy et al., 2020), ethical governance, transparency, equity, accountability, and human supervision of AI systems in health care settings are critical considerations. Thus, AI can be seen as a decision support tool for clinicians.

2. Aim and Objectives

This paper aims at exploring the opportunities of applying predictive AI healthcare analytics and machine learning in the area of disease detection at an early stage, high-risk patient recognition, and prevention. This paper provides a critical examination of the benefits and limitations of applying predictive AI in healthcare, as well as its viability and ethical issues.

- To examine the key healthcare data sources used in predictive AI analytics for early disease detection.
- To evaluate how machine learning models support the identification of high-risk patients across major disease areas.
- To critically assess the benefits, risks, and ethical challenges of using predictive AI to improve preventive healthcare strategies.



Figure 2 : Predictive AI Healthcare Analytics workflow for Early Diseases
Source: Author Created

3. Methodology: Secondary Data and Narrative Review Approach

The methodology adopted in this article is the secondary research by using the Research Onion model by Saunders. Overall, the study is at a philosophical level interpretivist because its main objective is to understand and interpret the discussion and application of predictive AI in the healthcare field as found in the literature. The research approach is inductive in nature, in that themes are being generated from the existing information in the academic, institutional and

Annual Methodological Archive Research Review

http://amresearchreview.com/index.php/Journal/about

Volume 3, Issue 11 (2025)

Online ISSN 3007-3197

Print ISSN 3007-3189

http://amresearchreview.com/index.php/Journal/about

regulatory field, rather than tested using primary numerical data. The method used is Qualitative as there are no patient records, hospital databases, clinical trials, interviews, surveys or statistical data used within the article (D' Costa, Kuhn and Fritz, 2020).

A narrative review methodology is used to evaluate published literature to date on predictive AI, machine learning models, early disease detection, identification of high-risk patients and planning for preventive healthcare. Peer-reviewed journal articles, systematic reviews, healthcare policy documents, AI ethics guidance, and regulatory documents are among the secondary sources. Focus on Electronic Health Records, Clinical Decision Support Systems, Medical Imaging, Wearables, and Risk Stratification and AI Governance. There are also a host of organisations that could come into play, such as the WHO, which has the capability to provide guidance on safety, ethics, accountability, and clinical applications with the use of AI in healthcare technologies. Additionally, any regulatory bodies that have a stake in the safety and clinical utility of AI-driven health care products are pertinent, such as the U.S. Food and Drug Administration.

When data is analysed to identify and link themes that emerge throughout the literature being reviewed, themetic analysis is used. This is appropriate as the purpose is not to statistically synthesize results, as in a meta-analysis, but to critically structure major trends found in the literature (Alowais et al., 2023). Themes involve around data sources in AI, machine learning models, disease applications, value of preventive healthcare, ethics/clinical limitations. As a result, the article will not be an empirical clinical study, but rather a secondary, qualitative, narrative review.

Annual Methodological Archive Research Review

http://amresearchreview.com/index.php/Journal/about

Volume 3, Issue 11 (2025)

Online ISSN 3007-3197

Print ISSN 3007-3189

http://amresearchreview.com/index.php/Journal/about

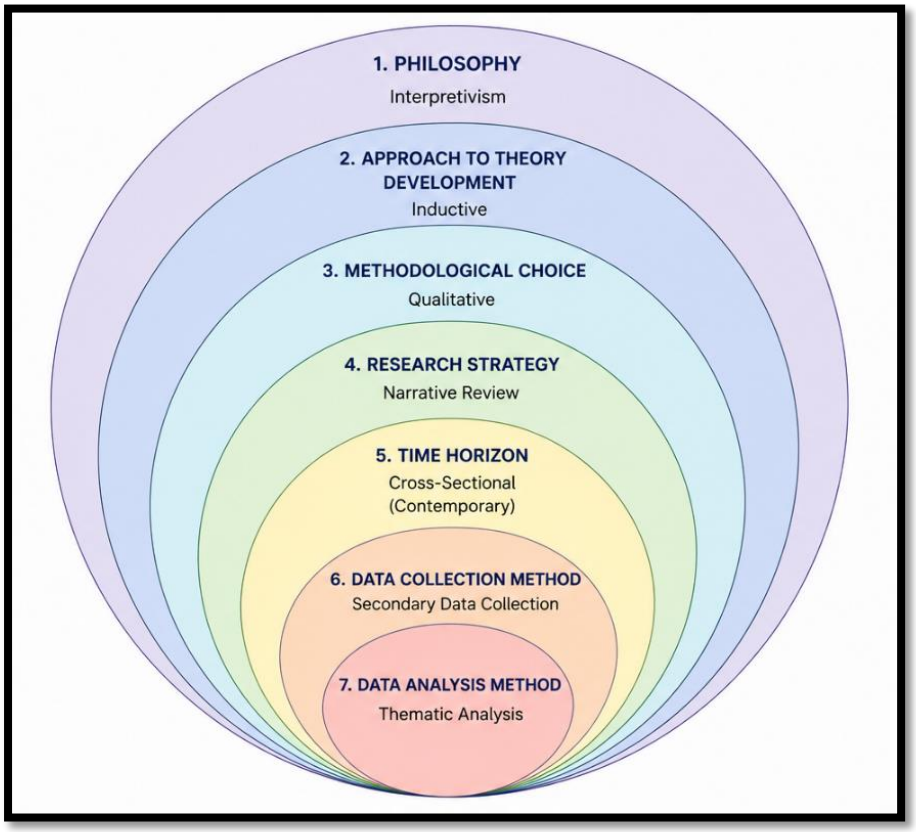


Figure 3: Research Onion Diagram

4. Conceptual Background of Predictive AI Healthcare Analytics

4.1 Predictive Healthcare Analytics

Predictive(Ali et al., 2025) healthcare analytics is the application of past and current health information to forecast the likelihood of future healthcare events. In the medical field, it could be predicting when a disease will occur, the likelihood of needing to be admitted to hospital, a patient's deterioration, the way in which they respond to treatment, or the likelihood of complications. Predictive analytics is more about predicting what will happen next, compared to descriptive analytics, which is about what has happened. It is therefore very relevant to the field of preventive medicine as it gives the clinicians the chance to intervene before the disease sets in and becomes severe (Razzak, Imran and Xu, 2019). For instance, trends in blood pressure, blood glucose, blood cholesterol, prior diagnoses, medications and lifestyle factors can be used to determine if the patient is more likely to develop diabetes, cardiovascular disease or chronic kidney disease.

4.2 Artificial Intelligence and Machine Learning

Artificial Intelligence is a large-scale area involving the creation of machines or systems that are able to do things that are typically done by humans, such as recognising patterns, making predictions, and helping with decisions. Machine learning is a part of AI where algorithms learn from data instead of having to be programmed with set rules (Nathani and Singh, 2021). For

Annual Methodological Archive Research Review

http://amresearchreview.com/index.php/Journal/about

Volume 3, Issue 11 (2025)

Online ISSN 3007-3197

Print ISSN 3007-3189

http://amresearchreview.com/index.php/Journal/about

example, in medicine, machine learning can be applied to huge datasets on patients and determine patterns that relate patient characteristics to certain diseases or conditions. Deep learning refers to an advanced form of machine learning that uses neural networks with multiple layers. In cases where image analysis is required, deep learning can be advantageous because it will help analyse the information contained within the image effectively.

4.3 Early Disease Detection

The early detection of disease plays a key role in preventive health care as many diseases are more manageable and less costly if discovered at an earlier stage. Early detection enables earlier treatment, helps track disease progression, helps make lifestyle changes and helps minimize avoidable complications (AbdulRaheem, 2023). It may also have beneficial effects on survival in certain diseases, like cardiovascular diseases and cancer, where a late diagnosis can have a dramatic effect. A health system's perspective of the early intervention may lead to fewer emergency admissions, longer term treatment costs, and strain on a hospital's resources.

4.4 High-Risk Patient Identification

High-risk patient identification can be supported by predictive AI which can classify patients into low risk, medium risk and high risk categories based on clinical and behavioural characteristics. It is now possible to concentrate screening, surveillance, and prevention programs only on those patients who stand to benefit the most from the interventions. However, these predictions cannot simply be accepted as true. These AI predictions should merely serve as an adjunct to professional judgment since they are fallible, leading possibly to misdiagnosis and mismanagement of cases.

Annual Methodological Archive Research Review

http://amresearchreview.com/index.php/Journal/about

Volume 3, Issue 11 (2025)

Online ISSN 3007-3197

Print ISSN 3007-3189

http://amresearchreview.com/index.php/Journal/about

5. Findings and Discussion: Thematic Analysis of Reviewed Literature

The literature review was conducted to identify key themes, and the results revealed five recurring themes: Sources of data for AI-driven disease prediction, machine learning models for disease early detection, specific disease applications, preventive healthcare strategies, and ethical & clinical implementation issues. The themes reveal that the application of predictive AI has considerable potential for early disease detection and recognition of patients who are at risk. On the other hand, the findings reveal that prediction using AI is not always accurate and positive. It relies on the quality of healthcare data, the appropriate use of machine learning models, clinical validation, fairness of the model, explainability(Nasim et al., 2025), and the responsible integration into healthcare workflows.

5.1 Theme 1: Data Sources Used in Predictive AI Healthcare Analytics

The predictive analytics AI system in healthcare relies significantly on the data available for use in training and testing the model. In the healthcare industry, the AI system does not make predictions on its own; rather, it identifies patterns using the patient-related data provided. Thus, the precision of disease prediction will rely on whether the information is adequate, comprehensive, clinically relevant, and managed well (van Smeden et al., 2021). The latest literature on AI in healthcare has indicated that the technology can utilize various types of data, including electronic health records, medical imaging, laboratory results, and patient-monitoring data, although it should be noted that low-quality data may lead to erroneous diagnoses and bias (Nagendran et al., 2020; Yogarajan et al., 2022).

Another crucial source of information for disease prediction using AI is EHRs. Such records contain demographic information about patients, their medical history, diagnoses, prescriptions, former treatment plans, admission to hospitals, and notes (Yang et al., 2022). The use of EHRs allows predicting algorithms to detect longitudinal associations, such as recurring abnormalities and the presence of multiple risk factors that may lead to disease development in the future (Lee and Kim, 2021). On the other hand, EHRs can have some gaps in the data, coding errors, and incomplete or erroneous notes.

Objective criteria can be found in the form of clinical outcomes and laboratory parameters for predicting various diseases such as diabetes mellitus, hypertension, and hypercholesterolemia. For example, glucose levels and hemoglobin A1C can be used to predict diabetes; serum cholesterol, blood pressure, and ECG can predict cardiovascular disorders; serum creatinine and estimated glomerular filtration rate (eGFR) can predict kidney disorders; and inflammatory biomarkers can predict infections and sepsis (Wang et al., 2022). The other important source of information for predicting medical conditions includes medical imaging, which involves x-ray imaging, CT scanning, MRI imaging, ultrasonography, mammography, pathology imaging, and retinal imaging(Aish et al., 2024).

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 3, Issue 11 (2025)
Online ISSN 3007-3197
Print ISSN 3007-3189
http://amresearchreview.com/index.php/Journal/about

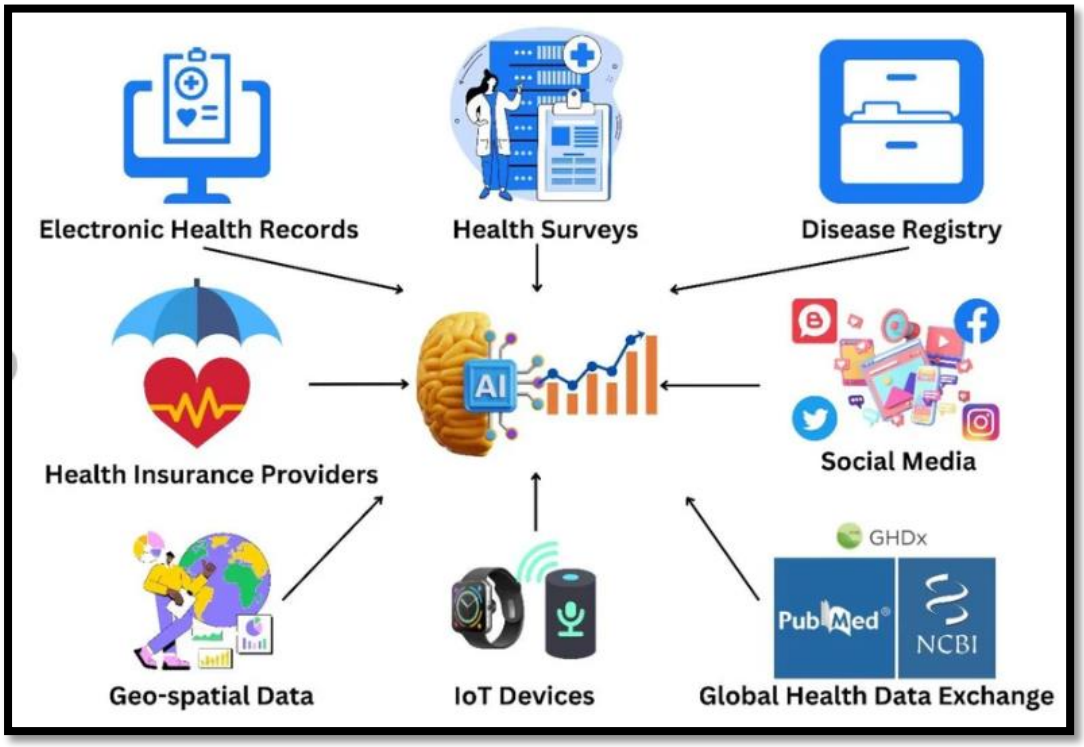


Figure 4 : Key Data Sources Used in Predictive AI Healthcare Analytics
Source: (Pranav Anjaria et al., 2023)

The role of wearable and remote data is becoming more prominent because it offers real-time information even outside of hospital environments. Wristbands and watches that monitor vital signs, glucose levels, blood pressure, and other vital signs may be helpful to detect any changes in patients' health. Moreover, genomic and lifestyle information will add value to prediction by identifying genetic risk factors for disease development, diet, smoking habits, consumption of alcohol, exercise, sleep, and family background (Firth et al., 2020). It should also be noted that social determinants of health such as financial status, housing conditions, education, job opportunities, access to care, and the environment are crucial due to the complex nature of health-related issues.

Crucially, any predictions made by artificial intelligence are only as good as the information fed into the system. This implies that any biases or inaccuracies within the training dataset will affect the results generated from a predictive model. WHO guidelines on the ethical deployment of artificial intelligence in healthcare indicate that the proper usage of artificial intelligence requires factors such as data accuracy, fairness, and transparency to be taken into consideration. The data gathering process, therefore, transcends mere technique.

5.2 Theme 2: Machine Learning Models for Early Disease Detection

Machine learning algorithms (Mehdi et al., 2024) are at the heart of the predictive analysis carried out using artificial intelligence in health care because they enable analysis of the information gathered by hospitals and predict the probability of disease development even before symptoms occur (Badawy, 2023). There is a variety of models depending on the type of data analyzed, and their efficiency depends not only on their accuracy but also on other criteria.

Annual Methodological Archive Research Review

http://amresearchreview.com/index.php/Journal/about

Volume 3, Issue 11 (2025)

Online ISSN 3007-3197

Print ISSN 3007-3189

http://amresearchreview.com/index.php/Journal/about

5.2.1 *Logistic Regression*

Logistic regression model is among those statistical models that are widely employed in forecasting health risks, owing to its ease of use and clarity. Logistic regression model can be applied in situations where it is necessary to predict the probability of occurrence of some events, such as contracting diabetes, cardiovascular disease, or returning to hospitals on account of health reasons (Van Grootven et al., 2021). The primary advantage of using logistic regression is that physicians can comprehend the relationships between various factors and diseases. However, logistic regression can have limitations in complicated scenarios.

5.2.2 *Decision Trees and Random Forests*

Decision Trees and Random Forest models are mostly employed in order to categorize patients into risk classes by using structured health care features such as age, blood pressure, cholesterol, glucose, medical records, and lifestyle features. According to Esmaily et al., random forest is more accurate than using a single decision tree model in forecasting the risks of developing heart diseases, hence making it effective for use in predicting risk factors. The ease of interpretation is one of the merits associated with using decision trees; however, random forest becomes less interpretable due to its composition of many decision trees (Esmaily et al., 2018).

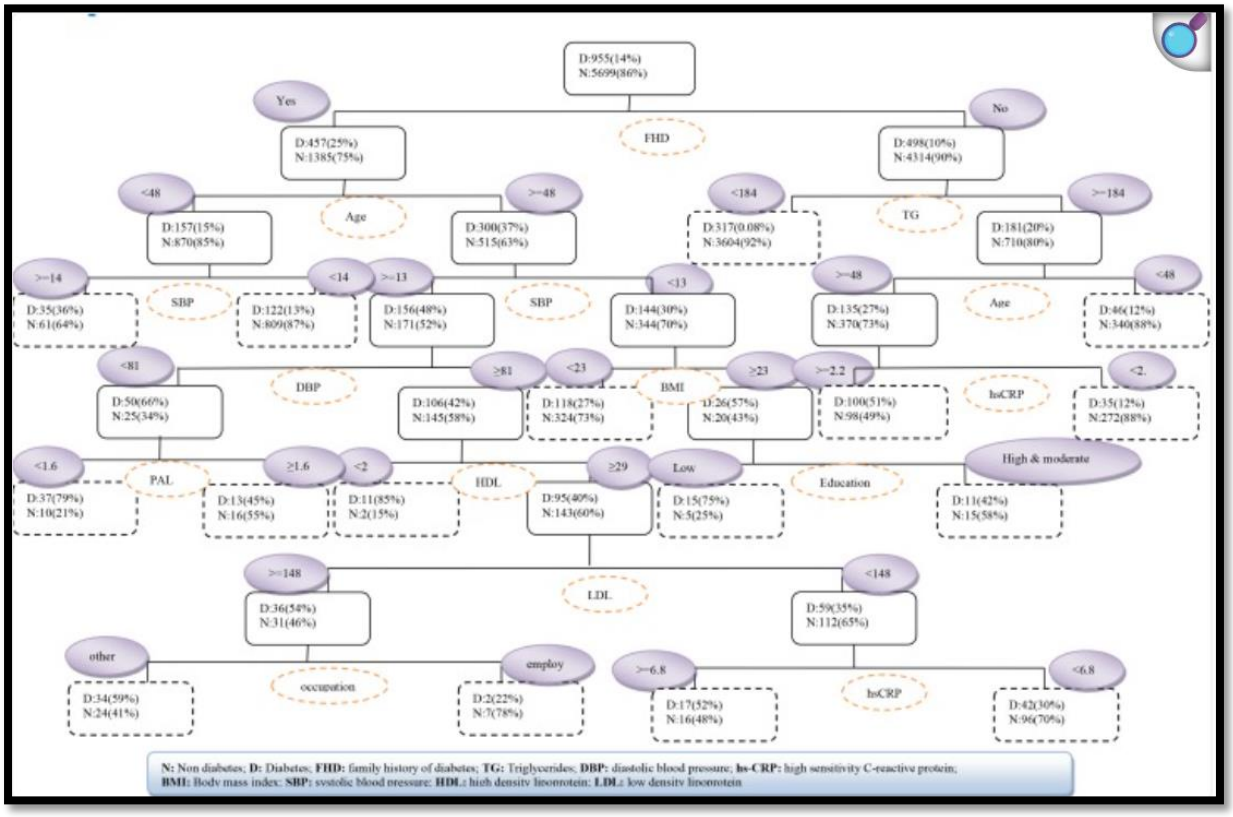


Figure 5 : Example of a Decision Tree Model for Diabetes Risk Classification
Source: (Esmaily et al., 2018)

5.2.3 *Support Vector Machines*

The support vector machine is widely applied in binary classifications, like classifying patients into at-risk and not-at-risk categories for certain diseases. It is applicable even to relatively

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 3, Issue 11 (2025)
Online ISSN 3007-3197
Print ISSN 3007-3189
http://amresearchreview.com/index.php/Journal/about

small data sets and complex spaces (Hermann and Lampinen, 2020). Nevertheless, their application in medicine can be hindered by low interpretability compared to simple classifiers, especially if complicated kernels are applied.

5.2.4 Gradient Boosting Models

Gradient Boosting algorithms such as XGBoost, LightGBM, and CatBoost are strong machine learning techniques applicable to structured datasets of healthcare. Gradient boosting models are sequentially learned, with each subsequent model being built on the mistakes made by previous models. Gradient boosting predictive analysis is generally highly accurate in predicting risks of disease, hospital admissions, and patient deterioration. Nevertheless, such models require optimization and proper explanations since doctors must comprehend why a particular individual is at risk.

5.2.5 Neural Networks and Deep Learning

Deep learning and neural networks work well in cases where complex pattern recognition is involved (Ahmed et al., 2023). In health care, deep learning is used extensively in medical imaging, pathological slides, retinal scanning, radiology, and big data (Akinnuwesi et al., 2023). Deep learning techniques can spot patterns that would be hard for people to discern. On the other hand, they are described as "black boxes" due to the difficulty in explaining the process through which they come up with their outputs.

5.2.6 Natural Language Processing

NLP makes it possible for artificial intelligence (AI) to interpret unstructured clinical data, such as physicians' notes, discharge data, radiology reports, and patient documents. This is crucial since most healthcare data is not documented in structured numeric databases. It is capable of extracting information on symptoms, diagnosis, drug therapy, and other clinical findings. Nevertheless, misinterpretation of language could result in errors.

5.2.7 Ensemble Models

Ensemble models use more than one algorithm for better results and greater reliability. For instance, the model could utilize the outputs from logistic regression, random forest, and gradient boosting to enhance the prediction of the risk (Kavzoglu and Teke, 2022). Although ensemble models enhance accuracy, they can also decrease transparency if medical practitioners fail to comprehend the process used to create the prediction.

Table 1 : Machine Learning Models and Their Healthcare Applications

Machine Learning Model	Main Healthcare Use
Logistic Regression	Interpretable disease risk scoring
Decision Tree	Simple clinical decision pathways
Random Forest	High-risk patient classification
Support Vector Machine	Disease/no disease classification
Gradient Boosting	Strong prediction using structured clinical data
Neural Network	Complex disease pattern recognition
Deep Learning	Medical imaging and cancer detection
Natural Language Processing	Clinical notes and patient record analysis
Ensemble Models	Improved prediction through model combination

The performance of models is generally evaluated through the measures of accuracy, precision, recall or sensitivity, specificity, F1 score, ROC-AUC curve, false positives, and false negatives. Nonetheless, in medical settings, mere accuracy of a model is insufficient. The model can have

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 3, Issue 11 (2025)
Online ISSN 3007-3197
Print ISSN 3007-3189
http://amresearchreview.com/index.php/Journal/about

high accuracy but still be considered unsafe if it does not detect high-risk patients. The failure to detect high-risk patients may lead to delays in treatments when a false negative occurs or induce undue stress and medical testing when a false positive happens.

5.3 Theme 3: Disease Areas Where Predictive AI Is Applied

AI predictions have useful applications for many disease categories due to the nature of some diseases that send early risk alerts before developing into full-blown disorders. The use of AI through structured and unstructured healthcare data is expected to help medical practitioners detect individuals at risk and provide them with appropriate treatment before a disease advances (Goh et al., 2021). It is vital to acknowledge that different diseases have varying levels of evidence, making the application of AI technology selective and prudent.

5.3.1 Cardiovascular Disease

One of the best cases of predictive AI is cardiovascular disease, in that many risk factors including blood pressure, cholesterol levels, whether a patient smokes, obesity, diabetes, ECG findings, lifestyle risk factors, and family history are usually tracked by health-care facilities. Predictive algorithms can use these metrics to identify the likelihood of having a heart attack, stroke, heart failure, or other forms of heart problems (Patel and Sengupta, 2020). While predictive models can possibly identify complex correlations among risk factors better than conventional scores, whether this translates into clinical utility is determined by the action taken based on prediction such as drug assessment, lifestyle advice, diagnostic workup, or referral to specialist care. Figure 6 depicts an AI-supported ECG algorithm designed to predict coronary revascularization in the general population.

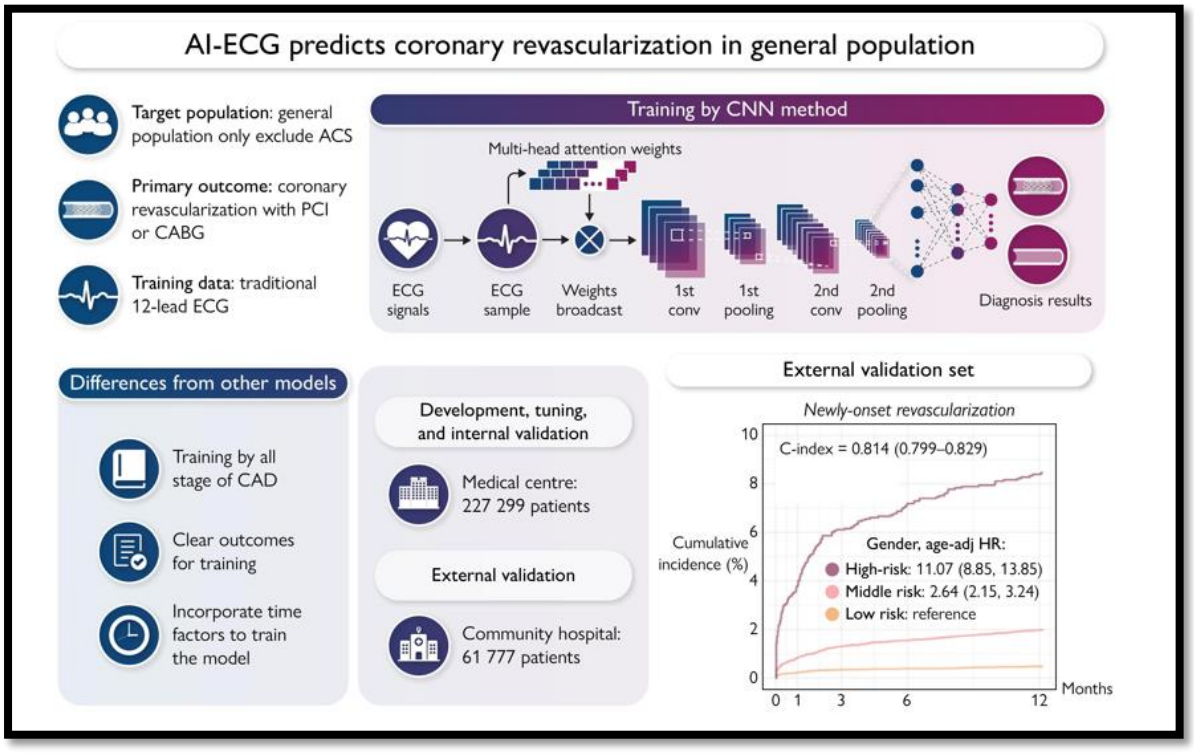


Figure 6 : AI-ECG Model for Predicting Coronary Revascularization
Source : (Chang et al., 2023)

Annual Methodological Archive Research Review

http://amresearchreview.com/index.php/Journal/about

Volume 3, Issue 11 (2025)

Online ISSN 3007-3197

Print ISSN 3007-3189

http://amresearchreview.com/index.php/Journal/about

5.3.2 Diabetes

In addition, AI can help in identifying patients at high risk of developing type 2 diabetes early on. Using machine learning algorithms, it is possible to predict the development of diabetes in the future by taking into consideration factors such as blood sugar levels, glycated hemoglobin level, body mass index, age, dietary habits, exercise routines, family history, and prior clinical history (Abbas et al., 2019). This is important because diabetes can be postponed or even prevented if appropriate lifestyle changes, weight loss, and other precautions are taken earlier on.

5.3.3 Cancer

The emergence of deep learning technologies has been critical to cancer diagnosis since they have the ability to interpret complicated imaging and pathological data. Artificial intelligence technologies can be useful in the interpretation of mammograms, computer tomography (CT) scans, radiological imaging, pathology slides, retinal scans, and genomics (McKinney et al., 2020; Jiang et al., 2023). The use of these technologies could enable early identification of possible cancers through the detection of any abnormality that needs further clinical assessment (Swanson et al., 2023). Nevertheless, the prediction of cancer is very risky due to the risk of missing some cancers and the risk of false alarms that could trigger stress and unnecessary invasive tests.

Annual Methodological Archive Research Review

http://amresearchreview.com/index.php/Journal/about

Volume 3, Issue 11 (2025)

Online ISSN 3007-3197

Print ISSN 3007-3189

http://amresearchreview.com/index.php/Journal/about

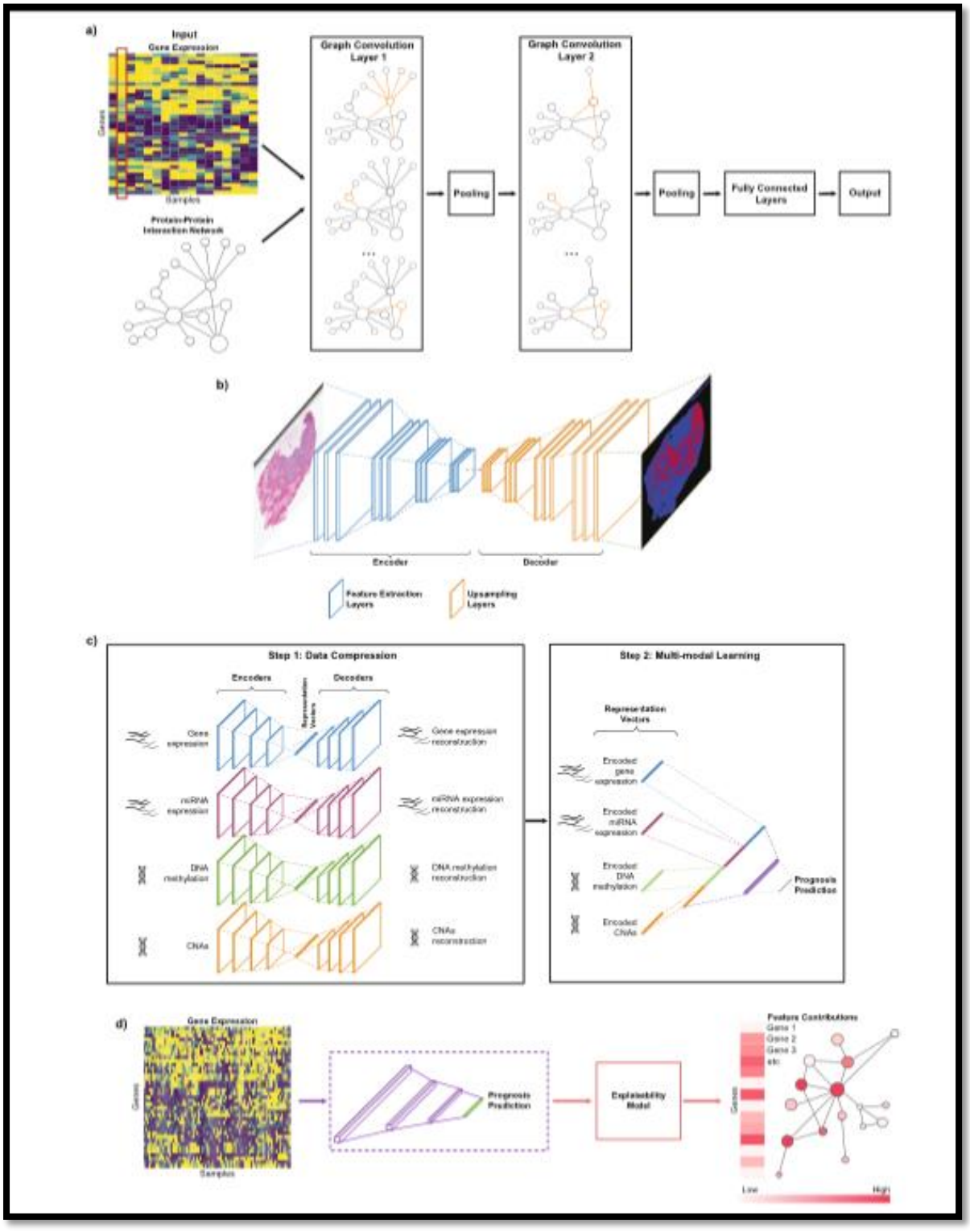


Figure 7 : Deep learning in cancer diagnosis, prognosis and treatment selection
Source : (Tran et al., 2021)

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 3, Issue 11 (2025)
Online ISSN 3007-3197
Print ISSN 3007-3189
http://amresearchreview.com/index.php/Journal/about

5.3.4 Sepsis and Acute Deterioration

Models of prediction can be helpful in detecting early warnings for sepsis or acute patient deterioration through analysis of vital signs, laboratory values, signs of infection, nurse documentation, and charting (Islam et al., 2023).

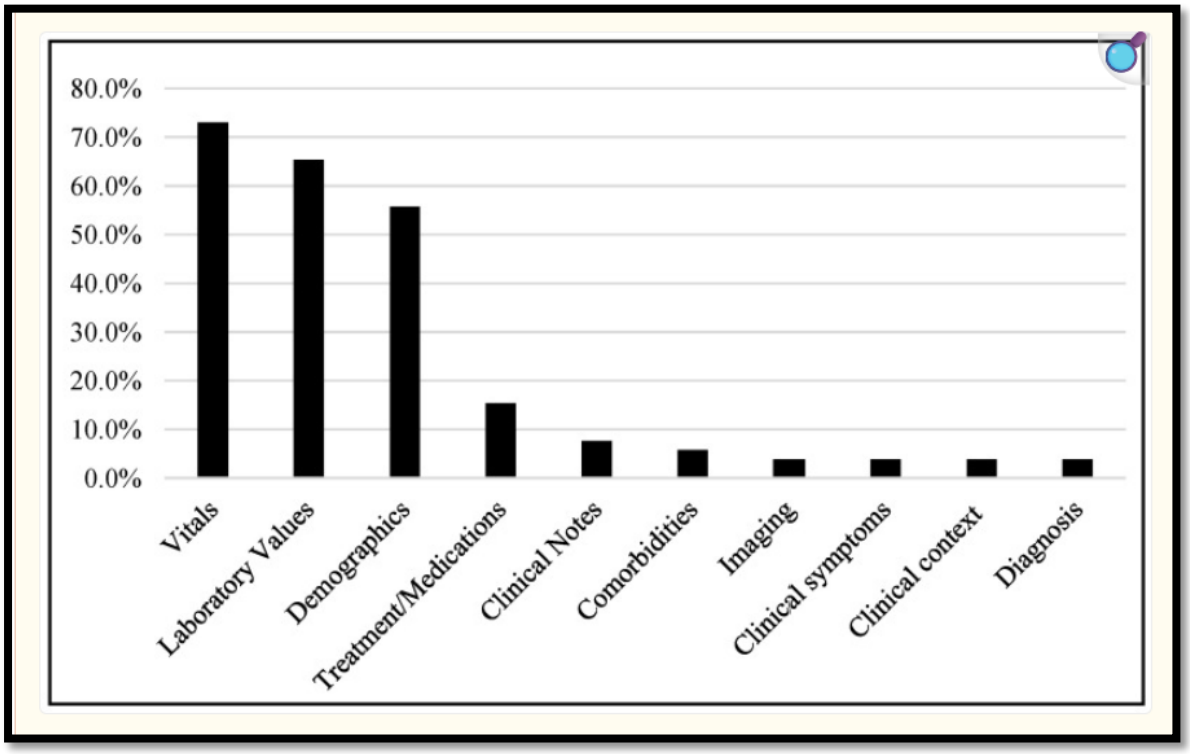


Figure 8 : Data Types Used in Machine Learning-Based Sepsis Prediction Studies
Source: Adapted from Islam et al. (2023)

This is significant because sepsis is a condition that develops fast, and early intervention reduces deaths (Schlapbach et al., 2020). While AI alerts can enable early detection, poor design can lead to alert fatigue due to false alerts being generated.

Table 2 : Evidence from Machine Learning Models for Sepsis Prediction

Study	Dataset	Model	Main finding
Gholamzadeh et al. (2023)	MIMIC-III	DT, RF, XGB	AUROC 0.918
Duan et al. (2023)	Shanghai Hospital	Double fusion deep learning	AUC 0.92, 24 hours before onset
Goh et al. (2021)	Singapore hospital	SERA algorithm	AUC 0.94, 12 hours before onset

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 3, Issue 11 (2025)
Online ISSN 3007-3197
Print ISSN 3007-3189
http://amresearchreview.com/index.php/Journal/about

Shashikumar et al. (2021)	Emory, UCSD, MIMIC-III	Deep AISE	AUC 0.90, 4 hours before onset
Kim et al. (2020)	Tertiary academic hospital	SVM, GB, RF	AUROC 0.93
Ibrahim et al. (2020)	MIMIC-III	RF, GB, SVM	AUROC 0.96

5.3.5 Chronic Kidney Disease

AI can facilitate the early diagnosis of chronic kidney disease through the analysis of creatinine values, glomerular filtration rate estimation, blood pressure, presence of diabetes mellitus, urinalysis findings, medication intake, and other biomarkers (Debal and Sitote, 2022). Kidney disease is often asymptomatic and silent, and predictive models may assist in recognizing patients requiring surveillance, diet and lifestyle counseling, and medication management.

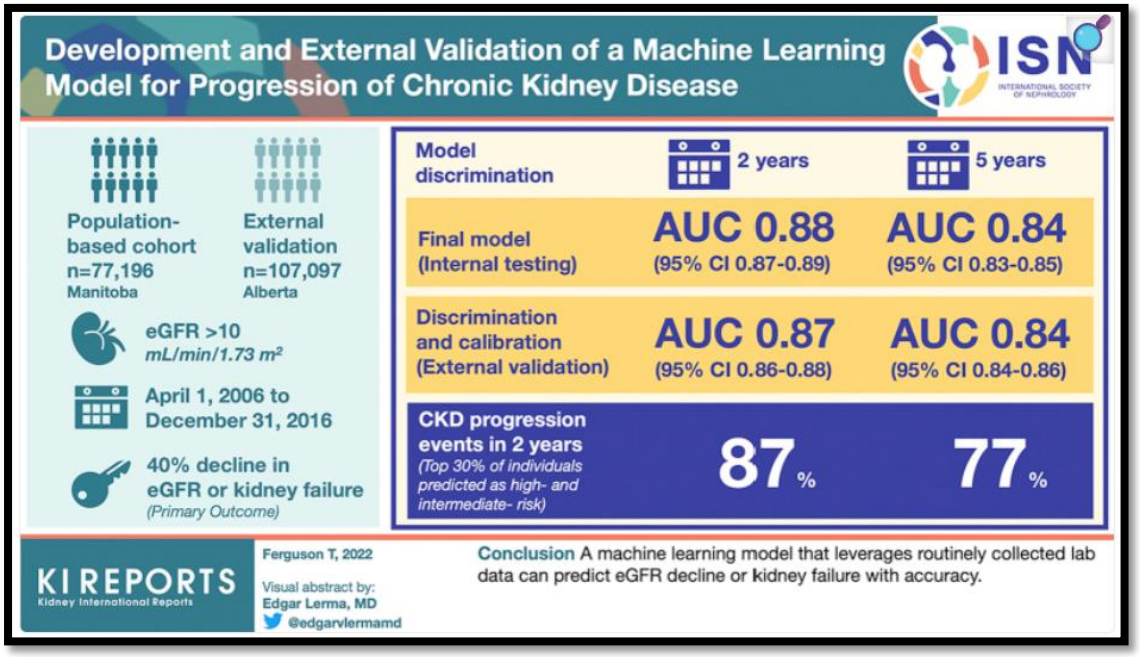


Figure 9 : Machine Learning Model for Predicting Chronic Kidney Disease Progression
Source: (Ferguson et al., 2022)

The machine learning algorithm presented in figure 9 was designed by Ferguson et al. (2022) for prediction of CKD progression on the basis of laboratory parameters commonly measured during the course of treatment. The results showed that this model had the ability to predict a 40% drop in eGFR with an AUC value of 0.88 and renal failure with an AUC value of 0.87.

5.3.6 Mental Health Risk

Moreover, the use of AI might also help detect early symptoms of depression, anxiety, suicide, or relapse through electronic medical records, digital behavioral data, and patient reports (Nickson et al., 2023). But this field needs to be ethically protected since mental health information is very personal and the results can have implications for stigmatization and privacy.

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 3, Issue 11 (2025)
Online ISSN 3007-3197
Print ISSN 3007-3189
http://amresearchreview.com/index.php/Journal/about

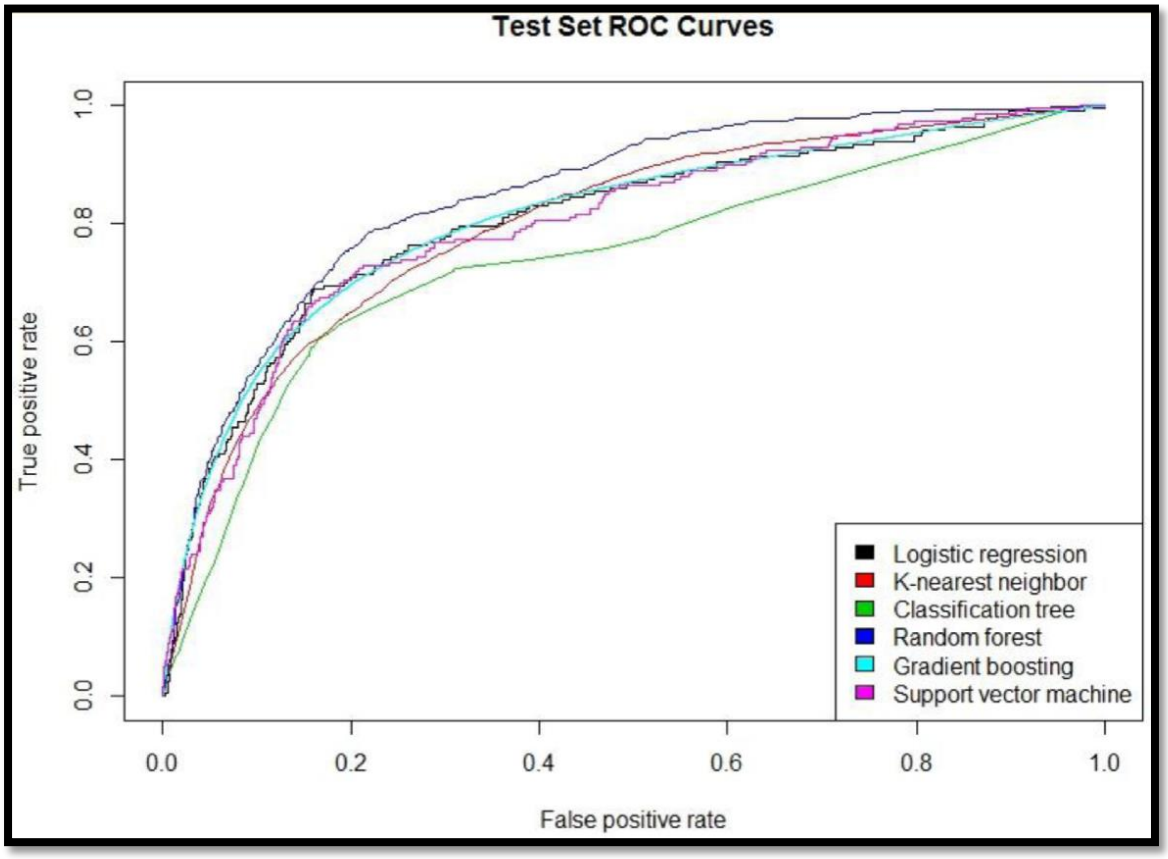


Figure 10 : ROC Curves of Machine Learning Models Predicting Future Suicidal Ideation
Source: (Van Mens et al., 2020)

In conclusion, the above examples demonstrate that predictive AI can be applied in different areas and not restricted to any particular disease. Predictive AI has the highest evidentiary standards for imaging-based predictions, clinically structured prediction, and biomarker based measurable conditions. Mental health prediction should be exercised with caution.

5.4 Theme 4 : Preventive Healthcare Strategies Supported by Predictive AI

Predictive AI only acquires value for medical applications once risk prediction can be converted into actionable clinical decision-making. A machine learning algorithm may classify a patient as being at high risk; however, that piece of information will be clinically insignificant unless it enables screening, treatment, tracking, or further diagnostic measures. Hence, the importance of predictive AI in preventive medicine lies not only in risk prediction but in better healthcare decisions through early identification and personalization.

5.4.1 Risk Stratification

Classification of risk based on prediction by the artificial intelligence is among the key advantages. The classification is performed using various patient's medical data including those obtained through laboratory studies, imaging scans, personal data including those collected from wearables used by the patient. As a result, there are groups of low-risk, medium-risk, and high-risk patients. The doctor thus knows which patients need urgent treatment and which should be monitored regularly. For example, patients with increased glucose levels, obese individuals, family history of type 2 diabetes, and patients with inactive lifestyles are at

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 3, Issue 11 (2025)
Online ISSN 3007-3197
Print ISSN 3007-3189
http://amresearchreview.com/index.php/Journal/about

higher risks for getting type 2 diabetes. Similarly, patients with hypertension, high level of cholesterol, smoking, and ECG abnormalities are high-priority patients for heart disease problems.

5.4.2 Targeted Screening

The use of artificial intelligence in predicting the future will assist in enhancing targeted screenings. Rather than applying the same screening techniques to all patients, it would be possible to conduct diagnostics on those who would benefit most from them. High-risk patients can be selected for blood work, scans, genetic testing, cancer screening, heart health evaluation, and so on.

5.4.3 Personalized Prevention

Predictions by AI can help achieve personalized prevention by customizing the prevention measures based on the risk profile of an individual. Prevention strategies include giving out information on lifestyle, review of drugs, immunizations, improvement of eating habits, stopping smoking, losing weight, and exercising. It is crucial because preventive medicine will be more effective when customized based on the risk factors associated with the patient.

5.4.4 Remote Monitoring

Use of wearable technology and remote monitoring solutions can enhance preventive healthcare by offering data continuously from the outside environment, which can be achieved through use of gadgets such as smart watches, glucose measuring gadgets, heart rate meters, and even blood pressure apparatus. These technologies can offer insights into a patient's health condition when combined with predictive AI systems.

5.4.5 Resource Allocation

The predictive power of data analytics could also be used by hospitals, clinics, and public health organizations to optimize their allocation of limited resources. Predicting who will end up in an emergency department visit or suffer from complications due to an illness would enable better scheduling of patient visits and other preventive measures (Khunti et al., 2022).

It is imperative to note that there is no value in the prediction done through predictive AI technology without any subsequent action from it. A risk score will mean nothing when used to predict patient outcomes without taking action. The key question to ask is whether clinicians put this prediction into practice in decision-making processes.

6. Benefits, Risks, and Ethical Implementation Challenges

Predictive analysis in healthcare through artificial intelligence brings about many advantages in the modern world especially with the increasing importance of early diagnosis and prevention in today's health care services. The first advantage that is of great importance is that of early diagnosis. Through artificial intelligence, diseases are diagnosed much earlier through the analysis of various health record data such as laboratory findings, images, and patient devices' outputs.

Predictive AI may also contribute to decision-making processes within the medical field through provision of relevant risk predictions for evidence-based practice. For instance, an AI model can help identify patients who might need additional testing, more monitoring, or referral to a specialist. However, it is essential not to replace doctors with predictive AI since the function of AI in the process is solely supportive (Baker et al., 2020).

Another benefit of predictive analysis is the individualized approach to therapy. With the help of predictive analysis, prevention will be personalized depending on personal characteristics.

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 3, Issue 11 (2025)
Online ISSN 3007-3197
Print ISSN 3007-3189
http://amresearchreview.com/index.php/Journal/about

Those individuals who are more likely to get diabetes, heart diseases, cancer, or kidney diseases will be examined and offered advice on how to stay healthy, medications, and remote monitoring.

Similarly, predictive AI may prove helpful in improving patient outcomes, lowering costs incurred by health care organizations, and lowering hospital admission rates due to identification of possible threats before the occurrence of any complications. As such, this application of predictive AI comes at a time when the health industry is under considerable strain as a result of increasing demands, staffing issues, and aging populations. According to the OECD, European health systems remain under considerable strain as a consequence of staffing shortages and system stress. Predictive AI will also allow remote communities to benefit from the technology by facilitating quicker processing of complex information as well as assisting clinicians in the absence of specialists.

7. Challenges, Risks, and Ethical Considerations

Predictive AI healthcare analytics, although promising, comes with many significant technical, ethical, and clinical challenges. It is essential to understand that a successful predictive AI healthcare algorithm does not necessarily guarantee safety or fairness when applied in a clinical setting.

7.1 Data Privacy and Security

Personal data in the healthcare sector is highly sensitive since it involves details related to one's diagnosis, treatment options, genetics, psychological well-being, lifestyle, and health records. Since artificial intelligence systems rely on vast amounts of data for training their models, they can expose such information to data abuse, illegal access, cyber threats, and breaches in the confidentiality of patient's information. Hence, predictive AI should adhere to data regulations and proper data management techniques.

7.2 Algorithmic Bias

AI algorithm bias is among the major threats in predictive healthcare. If data sets used to train algorithms do not contain comprehensive information or represent different demographic groups properly, algorithms will be inefficient in identifying diseases for particular ethnic groups, certain age cohorts, gender, people with disabilities, and poor people. Consequently, AI technology will increase existing disparities between social groups in access to healthcare rather than diminish these disparities.

7.3 Lack of Explainability

In most advanced systems, including deep learning-based models, there tends to be an aspect known as "black boxes." In this case, there could be an indication of a high-risk patient by assigning a risk score to a patient, but the explanation behind why that particular score was generated is not clearly specified.

7.4 Clinical Accountability

It is also necessary to understand who is responsible for errors in the AI prediction. For example, if the AI tool fails to detect a patient at high risk or misclassifies a patient who was initially considered low risk as high risk, then it is not always easy to determine whose fault it is (Asress Adimi Gikay et al., 2023).

7.5 False Positives and False Negatives

The occurrence of false positives might result in unnecessary tests being carried out, patient distress, higher expenses, and increased pressures on the health care system. However, the

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 3, Issue 11 (2025)
Online ISSN 3007-3197
Print ISSN 3007-3189
http://amresearchreview.com/index.php/Journal/about

negative effect that could arise from false negatives is worse since it results in delayed diagnosis and treatment.

7.6 Clinical Workflow Integration

AI solutions need to be integrated with clinical processes on the ground. If they cause more paperwork or alarm fatigue for healthcare professionals, they are likely to be ignored. Thus, an AI solution is required to be clinical in nature and not technologically-driven.

7.7 Validation and Regulation

The models of artificial intelligence should be tested in real-life clinical scenarios before being used in practice to ensure their reliability. As one can see from the FDA's list of AI-powered medical devices, numerous AI solutions are now being implemented within the healthcare sector, which means that the necessity for further safety control still remains relevant. Thus, predictive artificial intelligence should be clinically safe, ethical, understandable, and governed.

8. Future Directions

It is expected that in the future, predictive AI analytics in healthcare will progress beyond mere risk stratification to incorporate explainability, privacy, and clinical integration. Explainable AI will gain increased importance due to the necessity for both clinicians and patients to comprehend how and why a certain prediction was generated by the model.

Another area that might gain traction in the future is federated learning since it will be possible for AI systems to learn using data from various hospitals without actually disclosing any patient data. This will ensure privacy-preserving analytics while making AI models more efficient (Bi, Liu and Kato, 2022). Multi-modal AI is another key development that should be highlighted. Rather than working on a single data source, future systems will make use of electronic health data, images, labs, genomics, wearables, and clinical narratives.

There is a likelihood that there will be an expansion of real-time patient monitoring using wearable devices, remote monitoring, and telemedicine systems. This could enable early detection of deterioration in patients with chronic illnesses. More clinical validation and regulation would be required. Several studies carried out on FDA-approved medical devices incorporating AI have demonstrated how quickly AI applications have penetrated into medicine, but there is a need to comprehend how such applications are actually used in practice. Developments in the future regarding predictive AI technology should not only be about how technically adept they become. They should consider other issues like safety, ethical concerns, transparency, patient assurance, and usefulness in actual clinical practice.

9. Conclusion

The application of AI in predicting healthcare outcomes comes with many advantages when it comes to detecting various diseases early on. With the help of data analysis from various sources, including clinical, behavioral, imaging, genomic, and wearables data, it is possible to identify people with potential risks of developing certain diseases such as diabetes, heart disease, cancer, kidney disorders, sepsis, or deteriorating mental wellbeing. It means that healthcare professionals will be able to change their approach from reactionary to preventive methods.

The most important advantage of employing predictive artificial intelligence technologies is the ability to help with risk stratification, screening, prevention, telemedicine, and optimal resource allocation. If the use of AI technology is done properly, then there is a good chance

Annual Methodological Archive Research Review

http://amresearchreview.com/index.php/Journal/about

Volume 3, Issue 11 (2025)

Online ISSN 3007-3197

Print ISSN 3007-3189

http://amresearchreview.com/index.php/Journal/about

that it may help in detecting risks and avoiding any unnecessary complications, thus resulting in optimized results. The most important thing about artificial intelligence that must be noted is that the use of AI should not replace medical professionals.

Nevertheless, there are a number of problems that remain unresolved. Among such problems, one can identify data privacy, algorithmic bias, non-transparency of the decision-making process, errors of type I and II, clinical responsibility, and integration into workflow as factors affecting the safety of the application of predictive AI. Excellent results in terms of model performance in an experiment do not guarantee implementation of the algorithm into practice. Thus, one needs to move in the direction of responsible use. It is crucial for achieving the expected result in the context of prediction.

Annual Methodological Archive Research Review

http://amresearchreview.com/index.php/Journal/about

Volume 3, Issue 11 (2025)

Online ISSN 3007-3197

Print ISSN 3007-3189

http://amresearchreview.com/index.php/Journal/about

10. References

AbdulRaheem, Y. (2023). Unveiling the significance and challenges of integrating prevention levels in healthcare practice. Journal of Primary Care & Community Health, [online] 14(1), pp.1–6. doi:https://doi.org/10.1177/21501319231186500.

Ahmed, S.F., Bin, S., Hassan, M., Mahtabin Rodela Rozbu, Taoseef Ishtiak, Rafa, N., M. Mofijur, Ali and Gandomi, A.H. (2023). Deep learning modelling techniques: current progress, applications, advantages, and challenges. Artificial Intelligence Review, 56(1). doi:https://doi.org/10.1007/s10462-023-10466-8.

Alowais, S.A., Alghamdi, S.S., Alsuhebany, N., Alqahtani, T., Alshaya, A., Almohareb, S.N., Aldairem, A., Alrashed, M., Saleh, K.B., Badreldin, H.A., Yami, A., Harbi, S.A. and Albekairy, A.M. (2023). Revolutionizing healthcare: the Role of Artificial Intelligence in Clinical Practice. BMC Medical Education, [online] 23(1), pp.1–15. doi:https://doi.org/10.1186/s12909-023-04698-z.

Asress Adimi Gikay, Pin Lean Lau, Cigdem Sengul, Miron, A. and Malin, B. (2023). High-Risk Artificial Intelligence Systems under the European Union's Artificial Intelligence Act: Systemic Flaws and Practical Challenges. Social Science Research Network. [online] doi:https://doi.org/10.2139/ssrn.4621605.

Badawy, M. (2023). Healthcare predictive analytics using machine learning and deep learning techniques: a survey. ProQuest, [online] 10(40), p.40. doi:https://doi.org/10.1186/s43067-023-00108-y.

Baker, A., Perov, Y., Middleton, K., Baxter, J., Mullarkey, D., Sangar, D., Butt, M., DoRosario, A. and Johri, S. (2020). A Comparison of Artificial Intelligence and Human Doctors for the Purpose of Triage and Diagnosis. Frontiers in Artificial Intelligence, 3(543405). doi:https://doi.org/10.3389/frai.2020.543405.

Bi, H., Liu, J. and Kato, N. (2022). Deep Learning-Based Privacy Preservation and Data Analytics for IoT Enabled Healthcare. IEEE Transactions on Industrial Informatics, [online] 18(7), pp.4798–4807. doi:https://doi.org/10.1109/TII.2021.3117285.

D'Costa, S.N., Kuhn, I.L. and Fritz, Z. (2020). A systematic review of patient access to medical records in the acute setting: practicalities, perspectives and ethical consequences. BMC Medical Ethics, 21(18). doi:https://doi.org/10.1186/s12910-020-0459-6.

Firth, J., Solmi, M., Wootton, R.E., Vancampfort, D., Schuch, F.B., Hoare, E., Gilbody, S., Torous, J., Teasdale, S.B., Jackson, S.E., Smith, L., Eaton, M., Jacka, F.N., Veronese, N., Marx, W., Ashdown-Franks, G., Siskind, D., Sarris, J., Rosenbaum, S. and Carvalho, A.F. (2020). A meta-review of 'lifestyle psychiatry': The role of exercise, smoking, diet and sleep in the prevention and treatment of mental disorders. World Psychiatry, 19(3), pp.360–380. doi:https://doi.org/10.1002/wps.20773.

Aish, M. A., Ahmad, J., Nasim, F., & Iqbal, M. J. (2024). Brain Tumor Segmentation and Classification Using ResNet50 and U-Net with TCGA-LGG and TCIA MRI Scans. Journal of Computing & Biomedical Informatics, 8(01).

Goh, K.H., Wang, L., Yeow, A.Y.K., Poh, H., Li, K., Yeow, J.J.L. and Tan, G.Y.H. (2021). Artificial intelligence in sepsis early prediction and diagnosis using unstructured data in healthcare. Nature Communications, [online] 12(1), p.711. doi:https://doi.org/10.1038/s41467-021-20910-4.

Online ISSN 3007-3197

Print ISSN 3007-3189

http://amresearchreview.com/index.php/Journal/about

Hermann, K. and Lampinen, A. (2020). What shapes feature representations? Exploring datasets, architectures, and training. Advances in Neural Information Processing Systems, [online] 33, pp.9995–10006. Available at: https://proceedings.neurips.cc/paper/2020/hash/71e9c6620d381d60196ebe694840aaaa-Abstract.html [Accessed 2 May 2026].

Kavzoglu, T. and Teke, A. (2022). Predictive Performances of Ensemble Machine Learning Algorithms in Landslide Susceptibility Mapping Using Random Forest, Extreme Gradient Boosting (XGBoost) and Natural Gradient Boosting (NGBoost). Arabian Journal for Science and Engineering. doi:https://doi.org/10.1007/s13369-022-06560-8.

Ali, J., Rehman, F. U., Badar, M. A., & Nasim, F. (2025). Predictive analytics in supply chain management. Journal of Applied Linguistics and TESOL (JALT), 8(1), 225-244.

Khunti, K., Aroda, V.R., Aschner, P., Chan, J.C.N., Del Prato, S., Hambling, C.E., Harris, S., Lamptey, R., McKee, M., Tandon, N., Valabhji, J. and Seidu, S. (2022). The impact of the COVID-19 pandemic on diabetes services: planning for a global recovery. The Lancet Diabetes & Endocrinology, 10(12), pp.890–900. doi:https://doi.org/10.1016/s2213-8587(22)00278-9.

Mohammadagha, M. (2025). Hyperparameter Optimization Strategies for Tree-Based Machine Learning Models Prediction: A Comparative Study of AdaBoost, Decision Trees, and Random Forest. SSRN Electronic Journal. doi:https://doi.org/10.2139/ssrn.5226457.

Nagendran, M., Chen, Y., Lovejoy, C.A., Gordon, A.C., Komorowski, M., Harvey, H., Topol, E.J., Ioannidis, J.P.A., Collins, G.S. and Maruthappu, M. (2020). Artificial intelligence versus clinicians: systematic review of design, reporting standards, and claims of deep learning studies. BMJ, 368, p.m689. doi:https://doi.org/10.1136/bmj.m689.

Nathani, N. and Singh, A. (2021). Foundations of Machine Learning. Introduction to AI Techniques for Renewable Energy Systems, pp.43–64. doi:https://doi.org/10.1201/9781003104445-4.

Patel, B. and Sengupta, P. (2020). Machine learning for predicting cardiac events: what does the future hold? Expert Review of Cardiovascular Therapy, 18(2), pp.77–84. doi:https://doi.org/10.1080/14779072.2020.1732208.

Razzak, M.I., Imran, M. and Xu, G. (2019). Big data analytics for preventive medicine. Neural Computing and Applications, 32(9), pp.4417–4451. doi:https://doi.org/10.1007/s00521-019-04095-y.

Reddy, S., Allan, S., Coghlan, S. and Cooper, P. (2020). A governance model for the application of AI in health care. Journal of the American Medical Informatics Association, 27(3), pp.491–497.

Schlapbach, L.J., Kissoon, N., Alhawsawi, A., Aljuaid, M.H., Daniels, R., Gorordo-Delsol, L.A., Machado, F., Malik, I., Nsutebu, E.F., Finfer, S. and Reinhart, K. (2020). World Sepsis Day: a global agenda to target a leading cause of morbidity and mortality. American Journal of Physiology-Lung Cellular and Molecular Physiology, 319(3), pp.L518–L522. doi:https://doi.org/10.1152/ajplung.00369.2020.

Shang, Y., Pan, C., Yang, X., Zhong, M., Shang, X., Wu, Z., Yu, Z., Zhang, W., Zhong, Q., Zheng, X., Sang, L., Jiang, L., Zhang, J., Xiong, W., Liu, J. and Chen, D. (2020). Management of critically ill patients with COVID-19 in ICU: statement from front-line

Annual Methodological Archive Research Review

http://amresearchreview.com/index.php/Journal/about

Volume 3, Issue 11 (2025)

Online ISSN 3007-3197

Print ISSN 3007-3189

http://amresearchreview.com/index.php/Journal/about

intensive care experts in Wuhan, China. Annals of Intensive Care, 10(1). doi:https://doi.org/10.1186/s13613-020-00689-1.

Mehdi, M., Nasim, F., & Munir, M. Q. (2024). Comparative Risk Analysis and Price Prediction of Corporate Shares Using Deep Learning Models like LSTM and Machine Learning Models. Journal of Computing & Biomedical Informatics, 7(02).

Swanson, K., Wu, E., Zhang, A., Alizadeh, A.A. and Zou, J. (2023). From patterns to patients: Advances in clinical machine learning for cancer diagnosis, prognosis, and treatment. Cell, [online] 186(8), pp.S0092-8674(23)000946. doi:https://doi.org/10.1016/j.cell.2023.01.035.

Usman, M.S., Khan, M.S. and Butler, J. (2021). The Interplay Between Diabetes, Cardiovascular Disease, and Kidney Disease. ADA Clinical Compendia, 2021(1), pp.13–18. doi:https://doi.org/10.2337/db20211-13.

Van Grootven, B., Jepma, P., Rijpkema, C., Verweij, L., Leeflang, M., Daams, J., Deschodt, M., Milisen, K., Flamaing, J. and Buurman, B. (2021). Prediction models for hospital readmissions in patients with heart disease: a systematic review and meta-analysis. BMJ Open, 11(8), p.e047576. doi:https://doi.org/10.1136/bmjopen-2020-047576.

Nasim, M. F., Anwar, M., Alorfi, A. S., Ibrahim, H. A., Ahmed, A., Jaffar, A., ... & Zeeshan, H. M. (2025). Cognitively inspired sound-based automobile problem detection: A step toward explainable AI (XAI). International Journal of Advanced and Applied Sciences, 12(8), 1-15.

van Smeden, M., Reitsma, J.B., Riley, R.D., Collins, G.S. and Moons, K.G. (2021). Clinical prediction models: diagnosis versus prognosis. Journal of Clinical Epidemiology, 132, pp.142–145. doi:https://doi.org/10.1016/j.jclinepi.2021.01.009.

Wang, J., Wang, Y., Li, Y., Hu, Y., Jin, L., Wang, W., Gao, Z., Tang, X., Yan, L., Wan, Q., Luo, Z., Qin, G., Chen, L., Gu, W., Zhaohui Lyv and Mu, Y. (2022). High Normal Urinary Albumin–Creatinine Ratio Is Associated With Hypertension, Type 2 Diabetes Mellitus, HTN With T2DM, Dyslipidemia, and Cardiovascular Diseases in the Chinese Population: A Report From the REACTION Study. Frontiers in Endocrinology, [online] 13. doi:https://doi.org/10.3389/fendo.2022.864562.

Abbas, H.T., Alic, L., Erraguntla , M. and Abdul-Ghani, M. (2019). *Predicting long-term type 2 diabetes with support vector machine using oral glucose tolerance test*. [online] Nih.gov. Available at: https://pmc.ncbi.nlm.nih.gov/articles/PMC6905529/ [Accessed 7 May 2026].

Akinnuwesi, B.A., Olayanju, K.A., Aribisala, B.S., Stephen Gbenga Fashoto, Mbunge, E., Okpeku, M. and Owate, P. (2023). Application of support vector machine algorithm for early differential diagnosis of prostate cancer. *Data Science and Management*, 6(1), pp.1–12. doi:https://doi.org/10.1016/j.dsm.2022.10.001.

Chang, C.-H., Lin, C.-S., Lee, C.-H., Lin, C., Lee, C.-C., Liu, W.-T., Lee, Y.-T. and Tsai, D.-J. (2025). Real-world application of deep learning for ECG-based prediction of coronary artery disease and revascularization needs. *European Heart Journal - Digital Health*, [online] 6(6), pp.1124–1133. doi:https://doi.org/10.1093/ehjdh/ztaf096.

Debal, D.A. and Sitote, T.M. (2022). Chronic kidney disease prediction using machine learning techniques. *Journal of Big Data*, 9(1). doi:https://doi.org/10.1186/s40537-022-00657-5.

Online ISSN 3007-3197

Print ISSN 3007-3189

http://amresearchreview.com/index.php/Journal/about

Duan, Y., Huo, J., Chen, M., Hou, F., Yan, G., Li, S. and Wang, H. (2023). Early prediction of sepsis using double fusion of deep features and handcrafted features. *Applied Intelligence*, 53(14), pp.17903–17919. doi:https://doi.org/10.1007/s10489-022-04425-z.

Esmaily, H., Tayefi, M., Doosti, H., Majid Ghayour-Mobarhan, Nezami, H. and Amirabadizadeh, A. (2018). A Comparison between Decision Tree and Random Forest in Determining the Risk Factors Associated with Type 2 Diabetes. *Journal of Research in Health Sciences*, [online] 18(2), p.412. Available at: https://pmc.ncbi.nlm.nih.gov/articles/PMC7204421/?utm_source=chatgpt.com [Accessed 16 May 2026].

Ferguson, T., Ravani, P., Sood, M.M., Clarke, A., Komenda, P., Rigatto, C. and Tangri, N. (2022). Development and External Validation of a Machine Learning Model for Progression of CKD. *Kidney International Reports*, [online] 7(8), pp.1772–1781. doi:https://doi.org/10.1016/j.ekir.2022.05.004.

Gholamzadeh, M., Abtahi, H. and Safdari, R. (2023). Comparison of different machine learning algorithms to classify patients suspected of having sepsis infection in the intensive care unit. *Informatics in Medicine Unlocked*, [online] 38, p.101236. doi:https://doi.org/10.1016/j.imu.2023.101236.

Goh, K.H., Wang, L., Yeow, A.Y.K., Poh, H., Li, K., Yeow, J.J.L. and Tan, G.Y.H. (2021). Artificial intelligence in sepsis early prediction and diagnosis using unstructured data in healthcare. *Nature Communications*, [online] 12(1), p.711. doi:https://doi.org/10.1038/s41467-021-20910-4.

Ibrahim , Z.M., Wu , H., Hamoud, A., Stappen, L. and Dobson, R.J.B. (2020). *On classifying sepsis heterogeneity in the ICU: insight using machine learning*. [online] Available at: https://academic.oup.com/jamia/article/27/3/437/5709027?guestAccessKey=.

Islam, K.R., Prithula, J., Kumar, J., Tan, T.L., Reaz, M.B.I., Sumon, Md.S.I. and Chowdhury, M.E.H. (2023). Machine Learning-Based Early Prediction of Sepsis Using Electronic Health Records: A Systematic Review. *Journal of Clinical Medicine*, [online] 12(17), p.5658. doi:https://doi.org/10.3390/jcm12175658.

Jiang, X., Hu, Z., Wang, S. and Zhang, Y. (2023a). Deep Learning for Medical Image-Based Cancer Diagnosis. *Cancers*, 15(14), pp.3608–3608. doi:https://doi.org/10.3390/cancers15143608.

Jiang, X., Hu, Z., Wang, S. and Zhang, Y. (2023b). Deep Learning for Medical Image-Based Cancer Diagnosis. *Cancers*, [online] 15(14), p.3608. doi:https://doi.org/10.3390/cancers15143608.

Kim, J., Chang , H., Kim, D. x, Jang , D.-H., Park, I. and Kim, K. (2020). *Machine learning for prediction of septic shock at initial triage in emergency department*. [online] Available at: https://www.sciencedirect.com/science/article/abs/pii/S0883944119301534.

Lee, S. and Kim, H.-S. (2021). Prospect of Artificial Intelligence Based on Electronic Medical Records. *Journal of Lipid and Atherosclerosis*, 10(3), p.282. doi:https://doi.org/10.12997/jla.2021.10.3.282.

McKinney, S.M., Sieniek, M., Godbole, V., Godwin, J., Antropova, N., Ashrafian, H., Back, T., Chesus, M., Corrado, G.S., Darzi, A., Etemadi, M., Garcia-Vicente, F., Gilbert, F.J.,

Annual Methodological Archive Research Review

http://amresearchreview.com/index.php/Journal/about

Volume 3, Issue 11 (2025)

Online ISSN 3007-3197

Print ISSN 3007-3189

http://amresearchreview.com/index.php/Journal/about

Halling-Brown, M., Hassabis, D., Jansen, S., Karthikesalingam, A., Kelly, C.J., King, D. and Ledsam, J.R. (2020). International evaluation of an AI system for breast cancer screening. *Nature*, [online] 577(7788), pp.89–94. doi:https://doi.org/10.1038/s41586-019-1799-6.

Nasim, F., Masood, S., Jaffar, A., Ahmad, U., & Rashid, M. (2023). Intelligent sound-based early fault detection system for vehicles. *Computer Systems Science & Engineering*, *46*(3).

Nagendran, M., Chen, Y., Lovejoy, C.A., Gordon, A.C., Komorowski, M., Harvey, H., Topol, E.J., Ioannidis, J.P.A., Collins, G.S. and Maruthappu, M. (2020). Artificial intelligence versus clinicians: systematic review of design, reporting standards, and claims of deep learning studies. *BMJ*, 368, p.m689. doi:https://doi.org/10.1136/bmj.m689.

Nickson, D., Meyer, C., Walasek, L. and Toro, C. (2023). Prediction and diagnosis of depression using machine learning with electronic health records data: a systematic review. *BMC Medical Informatics & Decision Making*, [online] 23(1), pp.1–38. doi:https://doi.org/10.1186/s12911-023-02341-x.

Pranav Anjaria, Varun Asediya, Bhavsar, P., Pathak, A., Desai, D.N. and Patil, V. (2023). Artificial Intelligence in Public Health: Revolutionizing Epidemiological Surveillance for Pandemic Preparedness and Equitable Vaccine Access. *Vaccines*, 11(7), pp.1154–1154. doi:https://doi.org/10.3390/vaccines11071154.

Shashikumar, S.P., Josef, C.S., Sharma, A. and Nemati, S. (2021). DeepAISE – An interpretable and recurrent neural survival model for early prediction of sepsis. *Artificial Intelligence in Medicine*, [online] 113, p.102036. doi:https://doi.org/10.1016/j.artmed.2021.102036.

Tran, K.A., Kondrashova, O., Bradley, A., Williams, E.D., Pearson, J.V. and Waddell, N. (2021). Deep learning in cancer diagnosis, prognosis and treatment selection. *Genome Medicine*, [online] 13(1). doi:https://doi.org/10.1186/s13073-021-00968-x.

Van Mens, K., de Schepper, C., Wijnen, B., Koldijk, S.J., Schnack, H., de Looff, P., Lokkerbol, J., Wetherall, K., Cleare, S., C O'Connor, R. and de Beurs, D. (2020). Predicting future suicidal behaviour in young adults, with different machine learning techniques: A population-based longitudinal study. *Journal of Affective Disorders*, 271, pp.169–177. doi:https://doi.org/10.1016/j.jad.2020.03.081.

Yang, X., Mu, D., Peng, H., Li, H., Wang, Y., Wang, P., Wang, Y. and Han, S. (2022). Research and Application of Artificial Intelligence Based on Electronic Health Records of Patients With Cancer: Systematic Review. *JMIR Medical Informatics*, 10(4), p.e33799. doi:https://doi.org/10.2196/33799.

Yogarajan, V., Dobbie, G., Leitch, S., Keegan, T.T., Bensemann, J., Witbrock, M., Asrani, V. and Reith, D. (2022). Data and model bias in artificial intelligence for healthcare applications in New Zealand. *Frontiers in Computer Science*, 4. doi:https://doi.org/10.3389/fcomp.2022.1070493.

Annual Methodological Archive Research Review

http://amresearchreview.com/index.php/Journal/about

Volume 3, Issue 11 (2025)

Online ISSN 3007-3197

Print ISSN 3007-3189

http://amresearchreview.com/index.php/Journal/about

11. Appendix

11.1 Table 3 : Thematic Coding Matrix of Reviewed Literature

Article / Source	Theme 1: Data Sources Used in Predictive AI	Theme 2: Machine Learning Models for Early Disease Detection	Theme 3: Disease Areas Where Predictive AI Is Applied	Theme 4: Preventive Healthcare Strategies Supported by Predictive AI
Nagendran et al. (2020)	Medical imaging and clinical AI data	Deep learning models	Diagnostic AI in healthcare	Supports earlier diagnosis through AI-assisted clinical review
van Smeden et al. (2021)	Clinical prediction data	Clinical prediction models	Diagnosis and prognosis	Supports risk prediction for earlier clinical intervention
Yogarajan et al. (2022)	Healthcare datasets and bias-related data	Data-driven AI models	General healthcare AI	Highlights the need for fair and reliable preventive prediction
Yang et al. (2022)	Electronic health records	AI models using EHR data	Cancer-related AI prediction	Supports EHR-based monitoring and early identification
Wang et al. (2022)	Laboratory biomarkers	Risk prediction using clinical indicators	Diabetes, hypertension, cardiovascular and kidney disease risks	Supports biomarker-based monitoring and prevention
Firth et al. (2020)	Lifestyle data	Not model-specific	Mental health and lifestyle-related risks	Supports lifestyle-based prevention and early risk reduction
Badawy (2023)	Healthcare predictive datasets	Machine learning and deep learning models	General disease prediction	Supports predictive analytics for early detection
Van Grootven et al. (2021)	Clinical readmission data	Prediction models including	Heart disease	Supports readmission risk prediction and

Annual Methodological Archive Research Review

http://amresearchreview.com/index.php/Journal/about

Volume 3, Issue 11 (2025)

Online ISSN 3007-3197

Print ISSN 3007-3189

http://amresearchreview.com/index.php/Journal/about

		logistic regression		follow-up planning
Esmaily et al. (2018)	Clinical risk variables	Decision tree and random forest	Diabetes risk classification	Supports high-risk patient identification
Abbas et al. (2019)	Oral glucose tolerance and clinical diabetes data	Support vector machine	Type 2 diabetes	Supports early diabetes risk prediction
Ahmed et al. (2023)	Large-scale clinical and imaging datasets	Neural networks and deep learning	Complex disease pattern recognition	Supports advanced AI prediction for early detection
Kavzoglu and Teke (2022)	Structured predictive datasets	Ensemble models, random forest and XGBoost	General prediction modelling	Supports stronger prediction through model combination
Patel and Sengupta (2020)	Cardiovascular risk factors	Machine learning models	Cardiovascular disease	Supports cardiac risk prediction and earlier intervention
Chang et al. (2025)	ECG data	Deep learning / AI-ECG model	Coronary revascularisation risk	Supports early cardiovascular risk detection
McKinney et al. (2020)	Mammography images	AI screening system	Breast cancer	Supports earlier cancer detection and screening accuracy
Jiang et al. (2023)	Medical imaging data	Deep learning	Cancer diagnosis	Supports diagnostic assistance and lesion detection
Tran et al. (2021)	Genomic, pathology and imaging data	Deep learning	Cancer diagnosis, prognosis and treatment selection	Supports personalised cancer prediction and clinical planning
Islam et al. (2023)	EHRs, laboratory data and vital signs	ML sepsis prediction models	Sepsis	Supports early warning prediction before clinical deterioration

Annual Methodological Archive Research Review

http://amresearchreview.com/index.php/Journal/about

Volume 3, Issue 11 (2025)

Online ISSN 3007-3197

Print ISSN 3007-3189

http://amresearchreview.com/index.php/Journal/about

Goh et al. (2021)	Structured clinical data and clinical notes	SERA algorithm	Sepsis	Supports early sepsis detection and timely response
Shashikumar et al. (2021)	Clinical data	DeepAISE model	Sepsis and acute deterioration	Supports early deterioration prediction
Ferguson et al. (2022)	Routine laboratory data	Machine learning prediction model	Chronic kidney disease	Supports CKD progression prediction and early referral
Debal and Sitote (2022)	CKD clinical variables	Logistic regression and other ML models	Chronic kidney disease	Supports early CKD prediction and monitoring
Nickson et al. (2023)	Electronic health records	ML models for depression prediction	Depression and mental health risk	Supports early mental health risk identification
Van Mens et al. (2020)	Mental health follow-up data	Logistic regression, random forest, gradient boosting, SVM and KNN	Suicidal ideation	Supports future mental health risk prediction
Baker et al. (2020)	Clinical triage and diagnostic data	AI decision-support tools	Triage and diagnosis	Supports AI-assisted clinical decision-making
Khunti et al. (2022)	Diabetes service data	Not model-specific	Diabetes care and service disruption	Supports resource planning and preventive service recovery
Bi, Liu and Kato (2022)	IoT healthcare data	Privacy-preserving deep learning	Remote healthcare monitoring	Supports remote monitoring and privacy-preserving prevention