World Journal of Advanced Research and Reviews
eISSN: 2581-9615 CODEN (USA): WJARAI
Cross Ref DOI: 10.30574/wjarr
Journal homepage: https://wjarr.com/

(Research Article)

Check for updates

AI in healthcare: Predictive modeling, explainability and clinical impact

Sylvester Tafirenyika *

University of Zimbabwe.

World Journal of Advanced Research and Reviews, 2023, 19(03), 1700-1718

Publication history: Received on 18 August 2023; revised on 27 September 2023; accepted on 29 September 2023

Article DOI: https://doi.org/10.30574/wjarr.2023.19.3.1986

Abstract

Artificial Intelligence (AI) is revolutionizing our generation's health care model in the context of enhancing precision, effectiveness, and speed of clinical decision-making. This essay presents an overview of how AI technology has been used in the health care industry with specific reference to three most important pillars: predictive modeling, explainable AI (XAI), and their ultimate clinical impact. Predictive modeling methods, driven by machine learning algorithms and big health data, enable disease diagnosis at an earlier stage, risk stratification, and individualized treatment protocols. In the absence of transparency in the majority of AI models, transparency, trust, and accountability problems emerged, particularly in clinical high-risk applications. To counter these issues, the paper delineates the growing role of explainable AI (XAI) as a means for establishing confidence among clinicians, facilitating regulatory compliance, and maintaining ethical standards. The research integrates the latest breakthroughs, challenges, and real-world applications and explains how XAI frameworks can fill the algorithmic prediction-to-human interpretability gap. Other than this, the article also explains the clinical role of AI solutions in maximizing diagnostic accuracy, reducing healthcare disparities, and maximizing resource utilization in various healthcare facilities. As great as boundless potential exists in AI, according to the report, there is a cluster of issues associated with data quality, bias mitigation, model explainability, and clinical validation that need to be solved to support solid and credible implementation. Ethically based AI over the long term based on clinical transparency, fairness, and effectiveness within the clinical environment will be the foundation of transformative patient outcomes.

Keywords: Artificial Intelligence; Predictive Modeling; Explainable AI; Clinical Decision Support; Healthcare Analytics; Machine Learning; Trust In AI; Healthcare Outcomes; Deep Learning; Ethical AI

1. Introduction

The field of medicine is driving a revolution of transformation, and the mastermind behind this is the deployment of high-order digital technology. Riding the steering wheel of this revolution is Artificial Intelligence (AI), revolutionizing the acquisition, processing, and use of medical data. Adoption of AI-driven technologies like machine learning (ML) and deep learning (DL) has opened new vistas of medical diagnosis, treatment planning, resource management, and patient tracking. Their similarities lie in that they can both be used to provide work off and learn from unimaginably large collections of exceedingly heterogenous and high-complexity data—e.g., radiologic images, EHRs, genomics, and wearables. AI has thus come to demonstrate impressive potential to drive care efficiency, if not quality, in healthcare.

Predictive modeling is the most over-hyped clinical application of AI, a data-driven exercise that utilizes historical and current data to predict future events or trends. Predictive models will be capable of predicting the chances of developing established disease, re-admission to hospital, and signs of deterioration in advance; even individualized courses of treatment. These interventions render the health workers strong enough to respond reactively and, in the process, improve clinical performance as well as the efficiency with which healthcare resources are being utilized. For example,

* Corresponding author: Sylvester Tafirenyika

AI systems are trained to predict sepsis in ICU patients several hours before clinical features and time to deliver life-saving intervention.

As AI technologies are gaining more traction, some of the hottest topics among them right now are the model explainability and transparency issue. Many of the advanced AI systems, as well as deep learning systems like CNNs and RNNs, are "black boxes." They may produce high-accuracy prediction or classification, but their internal decision-making process may be unclear and hard for human beings to comprehend. In as open and high-stakes a profession as medicine, this unexplability is a higher-order revolutionary clinical, legal, and ethical risk and not an engineering limitation.

Physicians are trained to think on the basis of categorical logical and fact-based reasoning. Clinicians will not accept, or react to, uninterpretable or unintelligible AI system output—especially when outputs contradict clinicians' intuition and knowledge as clinician-experts. This gap will assist in postponing the deployment of AI in real clinical workloads following demonstration of model functionality. The U.S. Food and Drug Administration (FDA) and the European Medicines Agency (EMA) also increasingly require that AI systems be interpretable and explainable by some mechanism clinicians and patients would understand how the decisions are being built. Explainable AI (XAI) has thus been R&D intensive in making the AI system transparent, trust-worthy, and accountable.

Explanation of AI is not only an underpinning foundation of clinician trust but also an ethical accountability platform and a legal defensibility foundation. Explainability can be used to discover and minimize the risk of bias in training data, a matter especially salient in consideration of the fact that health systems govern diverse populations of heterogenous patients. Until and unless the AI models are made interpretable, there can be high chances that the models would be reflecting just the prevailing healthcare inequalities or otherwise they would be having gaps in them which would never be eliminated because they would ultimately turn to be unintelligible. For instance, one model trained from one population subgroup will be recognizably bad when used on another population subgroup from the same population and therefore produce diagnostic errors or treatable prescribed. Explainability platforms are capable of allowing one to audit the AI models, and one is able to see which are the most important features to the model and how exactly they are being used in combination when producing the output.

The marriage of explainable AI and predictive modeling is the start of a new generation of clinically oriented AI models. With rational, comprehensible predictive models, the clinician is no longer a tool but a decision-support tool on which they can depend and that allows them to deliver high-value, high-quality, personalized care. This marriage has already started paying dividends in radiology, oncology, cardiology, and critical care. For instance, explainable AI models were applied to identify suspicious mammogram lesions not just to outline the region of interest but also to identify in pixel intensity and spatial coordinates, thereby simplifying the decision-making process of the radiologists.

Success with the use of AI in medicine, however, is still a matter of interdisciplinarity among computer scientists, clinicians, ethicists, policymakers, and patients, and not technological progress. Building operationally consistent, ethical, and clinically relevant AI systems takes deep insight into clinical and computational processes. Clinical validation of AI models is the greatest success. Nicely behaved models in the well-controlled lab environment shatter in pieces when used on noisy, missing, or time-varying clinical data.

The purpose of this paper is to map the landscape of AI in medicine with explainability and predictability, and combined clinical effect. It attempts to provide a systematic review of the way in which AI technology is being employed to make predictability possible in medicine, how the path is being paved to make explainability possible, and actual impact that these technologies are having towards patient outcomes, clinician decision-making, and the health system. This is thus a book trying to fill the gap between innovation and AI and application into clinical practice on a daily basis.

We then proceed in the next sections to re-visiting some of the basics and some of the developments in medical predictive modeling. We then proceed to the new area of explainable AI, explaining methods like SHAP (SHapley Additive exPlanations), LIME (Local Interpretable Model-agnostic Explanations), and model-specific interpretability methods. We are looking to identify healthy examples of intersection of explainability and predictability, case studies and fields of application where AI has shown to create value in health outcomes. The essay merely mentions some of the problems that are confronting data such as data privacy and algorithmic bias to regulation horizons and integration issues to provide directions for future use and study.

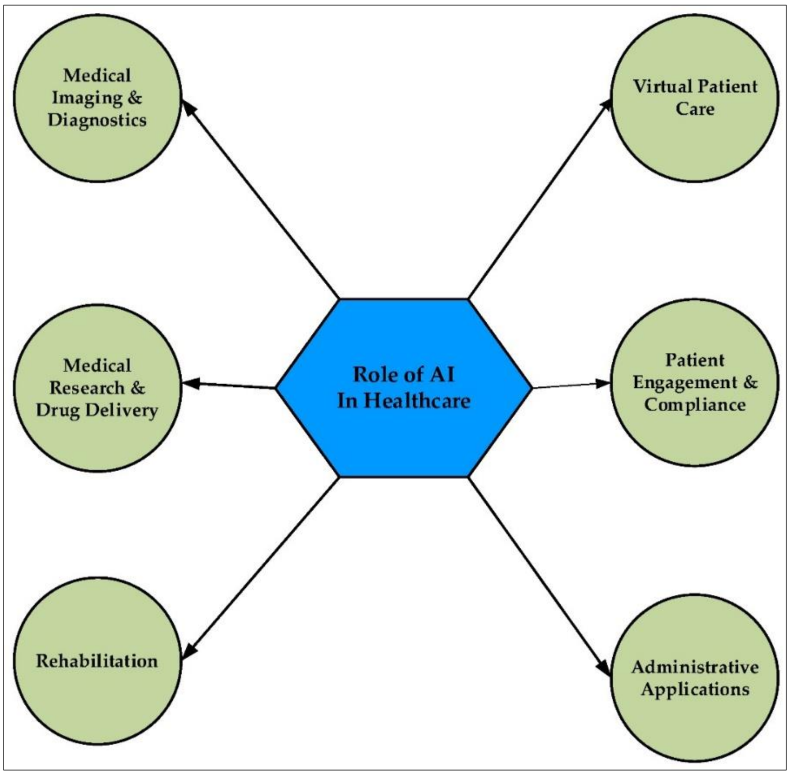

Figure 1 A Review of the Role of Artificial Intelligence in Healthcare

2. Background and Related Work

Medical Artificial Intelligence (AI) has changed dramatically over the past decades from simple rule-based systems to complex, data-driven predictive models. With today's AI technologies now mature enough to be increasingly deployed into clinical practice, interest not only is rising in optimizing predictive performance but also in rendering such systems transparent, reproducible, and interpretable in a clinical context. This section provides an in-depth review of existing research and trends in three broad categories: history of AI in healthcare, predictive modelling for healthcare, and large-scale deployment of explainable AI (XAI) in clinical environments.

2.1. History of AI in Healthcare

Deployment of AI in healthcare was pioneered by research on expert systems during the 1970s and 1980s. These initial systems, such as MYCIN and INTERNIST-I, employed symbolic logic and rule-based reasoning. These systems attempted to mimic the human clinician's ability to make decisions by encoding expert knowledge into strict preprogrammed rules. While these systems held promise for use in diagnostic support, their rigidity and the issue of adding new knowledge through their updates bound them. Rule-based systems were taking enormous amounts of time-consuming maintenance and were unable to learn from experience and were, hence, short to cope with the complexity and heterogeneity of clinical data.

The advancement of machine learning (ML) and more so recently of deep learning (DL) transformed the technology towards data-driven as opposed to knowledge-based solutions. These new techniques use massive reservoirs of structured and unstructured health care information—e.g., electronic health records (EHRs), imaging, laboratory testing, and genomic data—to predict and identify relationships. Machine learning algorithms like support vector machines, decision trees, and ensemble techniques started to replace traditional statistical models in activities like early disease identification, outcome prediction, and image interpretation.

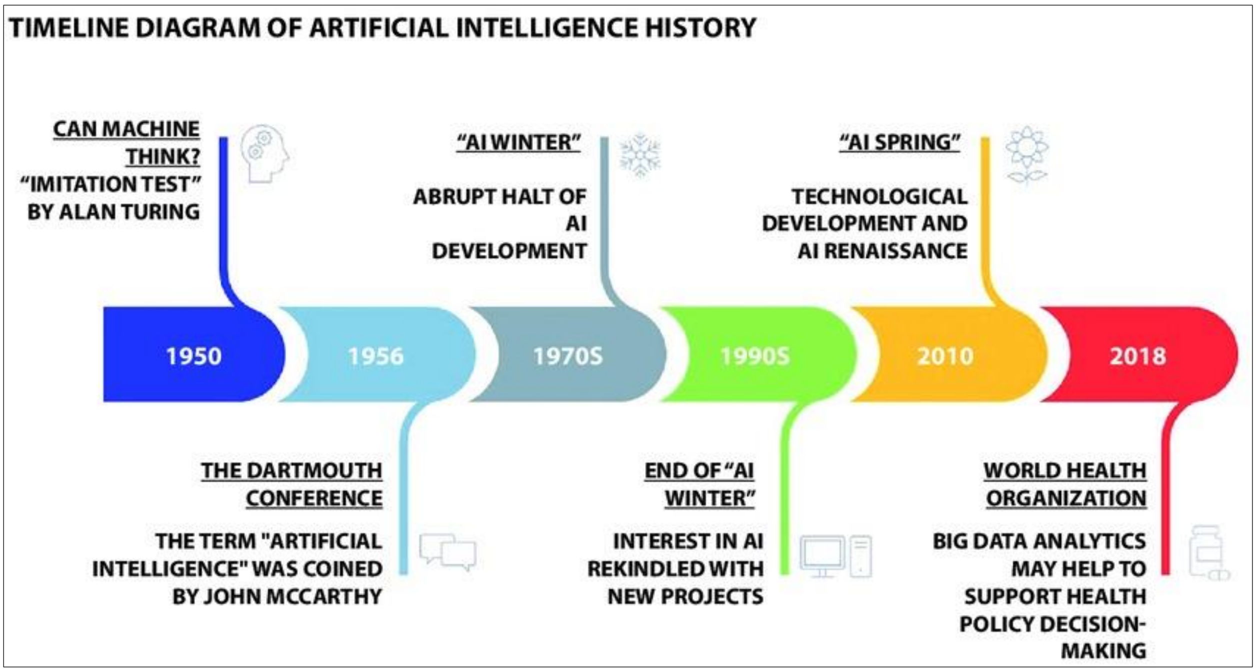

Figure 2 Timeline diagram showing the history of artificial intelligence

The tempo of contribution of AI in healthcare accelerated with deep learning. CNNs, RNNs, and transformer models provided the avenues for record-setting innovation in clinical notes natural language processing, medical imaging, and genomics. Through such models, there is a likelihood that it is possible to obtain very fine things from data, which would inevitably go undetected by humans, with the potential of pre-empting earlier diagnosing and prognosis. As the computational power grew and health data became increasingly available, AI began to be used more and more in clinical decision support, population health management, and precision medicine.

While all these advances are encouraging, their "black-box" nature also created serious concerns regarding their explainability, consistency, and security, particularly when used in dangerous fields like medicine. In line with this reality, greater emphasis is now placed on building models of AI, which not only predict but explain, are ethics-compliant and clinically relevant.

2.2. Predictive Modeling in Medicine

Predictive modeling is increasingly becoming the backbone of AI in medicine, with the potential to predict patient outcomes and guide clinician decision-making. Building such models predicts the probability of future occurrences—disease progression, hospital readmission, or treatment response—based on historical and present patient information.

Early models employed logistic regression and Cox proportional hazards, which were simple to interpret but had low ability to fit high-dimensional nonlinear interactions. More recent machine learning methods, such as gradient boosting machines, random forests, and neural networks, have shown better predictive ability by fitting weak patterns in high-dimensional data. These models are being used more and more universally in clinical practice.

The most commonly used application of predictive modeling is readmission risk prediction. Models are used by hospitals to identify which patients have the highest chances of readmission within 30 days of discharge in a bid to intervene with these patients via such interventions as post-discharge phone call, medication reconciliation, or post-discharge. Likewise, sepsis detection algorithms also monitor vital signs, laboratory findings, and clinical observations to alert clinicians to the onset of sepsis earlier than when it will become clinically apparent. These early warnings have the potential to reduce mortality and delay by a tremendous amount.

Table 1 Comparison of Predictive Modeling Use Cases in Healthcare

Use Case	Data Source	Model Type	Clinical Outcome
Sepsis Prediction	Electronic Health Records (EHR), vital signs	Random Forest, Deep Neural Networks	Early detection of sepsis, timely intervention, reduced mortality
Cancer Detection (e.g., breast cancer)	Imaging data (e.g., mammograms), genomic profiles	Convolutional Neural Networks (CNN), Support Vector Machines (SVM)	Accurate tumor classification, early diagnosis, treatment planning
Hospital Readmission Risk	EHR, demographics, comorbidity indices	Logistic Regression, Gradient Boosting	Reduced readmission rates, improved discharge planning
Diabetes Onset Prediction	Lab results, patient history, lifestyle data	Decision Trees, Neural Networks	Preventive care planning, lifestyle intervention recommendations
Heart Disease Risk Stratification	EHR, ECG, blood test results	Ensemble Models, Bayesian Networks	Risk categorization, personalized care strategies
ICU Mortality Prediction	Time-series ICU data, lab tests	Recurrent Neural Networks (RNN), XGBoost	Enhanced critical care decisions, optimized resource allocation
Stroke Prediction	Clinical history, MRI scans, genetic markers	Logistic Regression, Deep Learning	Early intervention, reduced long-term disability

A second critical use case is cardiovascular event prediction. EHR data, wearable data, and imaging protocols are integrated in predictive models that predict risk for a given individual to experience stroke or heart attack. Prediction facilitates advance care planning prior to the event and individualized treatment plans.

Because predictive models are so much advertised, their value to practice is no longer a matter of performance, but of usability within clinical workflow and provision of actionable insights. Above all, perhaps, unless predictive models can be understood and are not black boxes, or unless they make counterintuitive predictions, clinicians will not trust, and will not act on, them. This is about the ever more critical imperative of explainability and transparency of AI-driven prediction systems.

2.3. Explainable AI (XAI) in Clinics

Therapeutic use of AI systems requires performance perfection more than that; explainability will provide safety, trust, and accountability. Explainable AI, or XAI, is a set of methods and tools which makes clear what an AI model really does by explainability in simple language. In medicine, where literally life-changing decisions are made, it is required in interpretability from clinician trust, as well as regulation.

A number of frameworks have therefore been advanced with the aim of facilitating post-hoc explanations of black box models. SHAP (SHapley Additive exPlanations) provides global and local explanation view by determining each contribution of the features to a single prediction. LIME (Local Interpretable Model-agnostic Explanations) locally approximates high-complexity models around a specific prediction via an interpretable proxy model to help clinicians understand why a model has come to a specific decision. These are model-agnostic and can be applied to operate on an enormously broad class of algorithms.

Aside from all of these methods of feature attribution, counterfactual explanation is getting a lot of attention. Counterfactual explanations tell us about marginal changes to input features that would have caused the model to behave differently. A readmission risk-prediction model, for example, can tell us that better medication compliance or decreasing blood pressure would have resulted in a lower risk score. Not only are counterfactual explanations feasible but also actionable, and doctors can provide particular attention to modifiable risk factors.

Although XAI is so promising, it is far from easy to deploy it in the clinical setting. Interpretability is achieved at the expense of model simplicity and accuracy. In addition, explanations need to be adjusted to suit the agendas of different stakeholders-administrators, patients, clinicians, and regulators-differing in levels of demand and technical expertise.

An additional issue is that flawed or overly simplistic explanations will obscure the true behavior of the model they purport to explain.

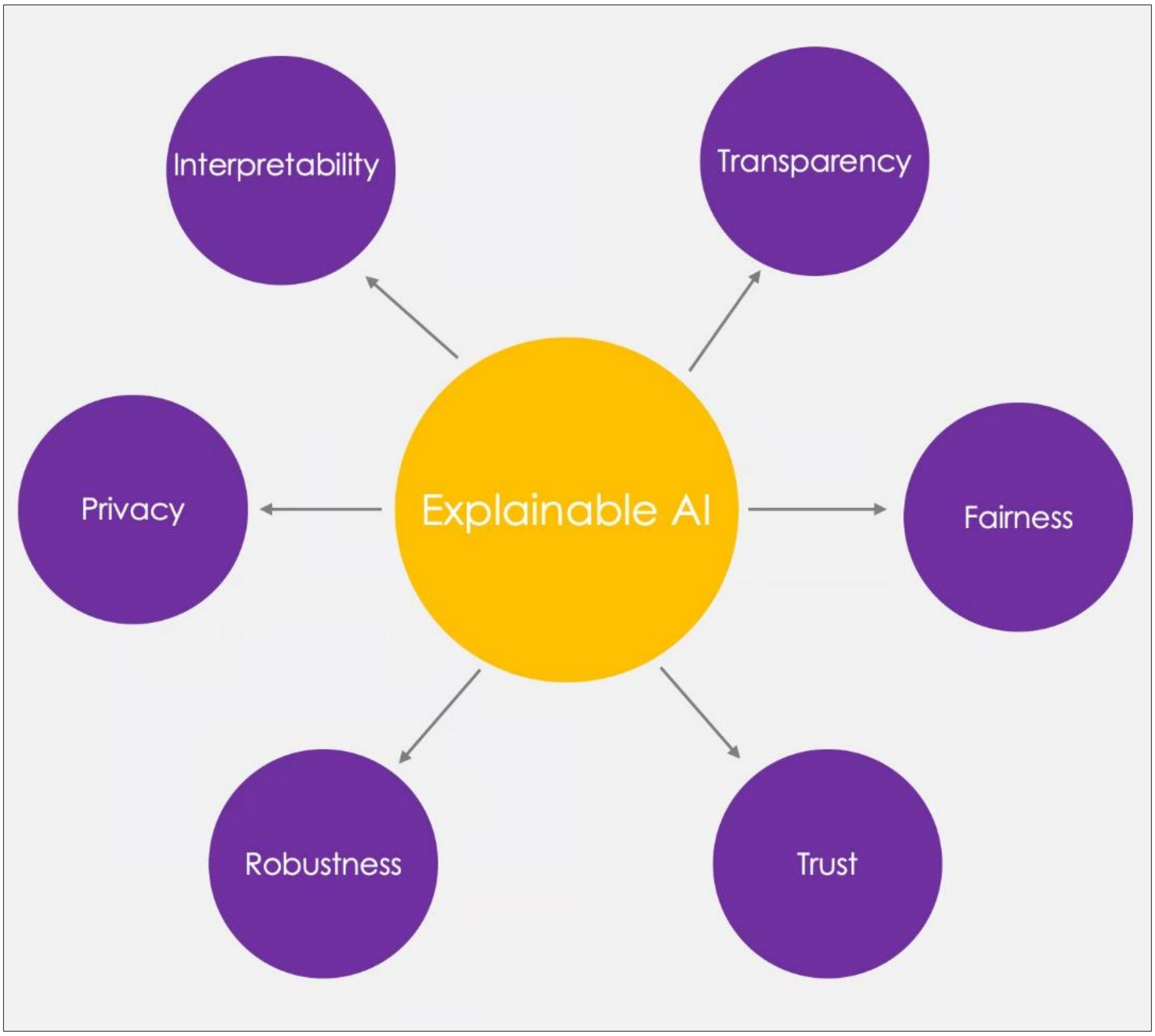

Figure 3 Explainable AI, LIME & SHAP for Model Interpretability

To tackle these challenges, recent studies have focused on combining XAI techniques with model development pipelines such that human-in-the-loop procedures may be triggered in which clinicians are able to view, criticize, and manipulate predictions from models. Even regulatory authorities such as the U.S. Food and Drug Administration (FDA) and the European Medicines Agency (EMA) are beginning to issue guidance on the application of explainable AI in health care, again citing the requirement for explainability when used in real-world applications.

3. Methodologies for Predictive Modeling

Predictive medicine modeling is advanced technology using ample and diverse data to predict clinical outcomes, identify patterns of risk, and guide individual therapies. Uniform methodology throughout many points in the pipeline is to be expected from predictive modeling. Big data curation and preprocessing, from the implementation and choice of top-of-the-line machine and deep learning tech to how to best decide whether a model holds, are nodes like these. They are all methodological issues that make sure the models are clinically actionable, stable, and interpretable in the sense that the results received are not statistically valid but workable in the day-to-day operations of clinics.

3.1. Data preprocessing and collection

Data are the foundation for every predictive model project. In medicine, data are gathered from more than a single source with individual valuable information. Electronic medical records provide a convergence point with an integration of patient demographics, treatment, clinical history, diagnosis, and outcome. These data constitute the core of patient longitudinal trajectories and are linkable with imaging modality data, capturing diagnostic and prognostic data via imaging modalities like MRI, CT scans, and X-rays. Body sensors infuse the space of information by continuously monitoring such physiological metrics like movements, heartbeat, and sleeping habits in a manner that online monitoring of health based on one's clinical data is enabled. Genomics databases offer an immensely valuable facility through tracking of the patient's genetic variation that would be used in the calculation of disease susceptibility, drug sensitivity, and efficacy of treatment.

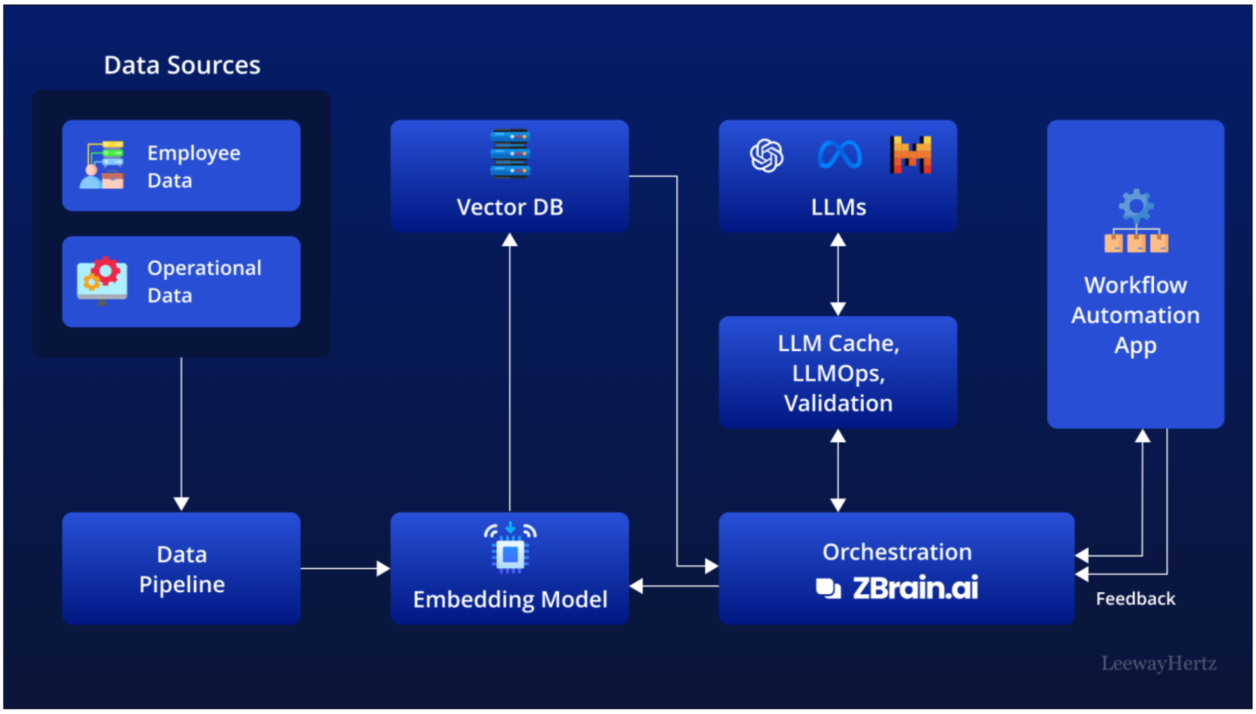

Figure 4 AI for workflow automation: Use cases, applications, benefits and development

These data sources are big, though, and it becomes hard for them to be processed by heavy preprocessing. The very first step is data cleaning with error detection and correction, inconsistencies, and outliers. For instance, wrong patient IDs or meaningless values in vital signs need to be detected and corrected. Normalization is the second operation, standardizing different size and unit data into one format for analysis. It is particularly required when combining data of different types such as genomic sequence and clinic data into a single homogeneous modeling platform. Nons occur values are the second pervasive issue, if not removed will bias model output or reduce performance. Missing value imputation techniques vary from basic like mean substitution to the sophisticated like k-nearest neighbors or multiple imputation by chained equations in filling and replacing missing values. Precise preprocessing not only enhances the data quality but also ensures that machine learning algorithms are fed correct, complete, and consistent inputs.

3.2. Machine Learning Approaches

Following. the. data. generation, the. second. requirement. is. selection. and. application. of. appropriate. machine. learning. algorithms. to. the. given. available. provided. prediction. task.

One of the default approaches used in healthcare predictive analytics due to the availability of data with outcomes is supervised machine learning. One of the widely tested but low-level algorithms used mainly for binary class situations like disease onset or non-onset prediction is logistic regression. The multicollinearity interpretability strength of logistic regression makes it the best choice for any medical problem. Random forest is a bagging technique where enormous numbers of decision trees are trained and the decision tree prediction output is averaged to give more predictive values. They are overfit-resistant and can handle high-dimensional data of enormous data sets and therefore are best suited to handle medical data that is complex. Support vector machines are another supervised algorithm, best suited to be run in high-dimensional space, and most suited to be utilized while building classification problems. Along with regular machine learning, deep learning is gaining prominence today since it can learn directly from raw input feature data and identify hard, non-linear patterns. Convolutional Neural Networks have been widely used in medical imaging processing such as cancer detection from radiographs or abnormality detection in pathological slides. Their cortex-inspired feature representation makes them learn spatial hierarchical representations of the features and are therefore computationally extremely effective when dealing with image processing.

Recurrent Neural Networks are, nonetheless, best suited to sequential data and have also been applied with great success to time-series analysis of wearable sensor or longitudinal Electronic Health Record data. Their capacity to learn temporal dependency can model disease progression as well as predict forward to forecast future occurrences.

Ensemble models were hypothesized to improve predictive performance by averaging multiple algorithms. Ensemble models capture individual learners' strengths and are susceptible to no weak points compared to them, but instead tend to provide a general and consistent rate of performance. Techniques such as bagging, boosting, and stacking were trained in ensemble modeling to enhance modeling compared to single-model solutions. A decision tree and a logistic regression model, together with a support vector machine, would perform better than either alone. The choice of the machine learning method relies primarily on data type, problem type to be predicted, and requirement for interpretability rather than accuracy.

3.3. Performance Measures

Precision-recall curves are enhanced for the class imbalance of the majority of the tasks such as prediction for a rare disease. The plots illustrate the correlation between true positive rate and positive predictive value, to what percentage the model is doing well in being good at good cases without generating a high rate of false positives. F1-score, as the harmonic mean of recall and precision, provides a single number that averages both and is of inestimable assistance when false negatives and false positives both have enormous applications to them, such as in medicine.

Table 2 Summary of Machine Learning Algorithms Used in Healthcare

Algorithm	Strengths	Weaknesses	Healthcare Use Cases
Logistic Regression	Simple, interpretable, effective for binary classification	Limited to linear relationships, less effective for complex patterns	Disease diagnosis (e.g., diabetes prediction), readmission risk
Decision Trees	Easy to interpret, handles both numerical and categorical data	Prone to overfitting, unstable with small data changes	Clinical decision support, treatment planning
Random Forest	High accuracy, reduces overfitting, robust to noise	Less interpretable than single trees, slower for large datasets	Predicting patient outcomes, identifying risk factors
Support Vector Machines (SVM)	Effective in high-dimensional spaces, works well with clear margins	Not easily interpretable, computationally intensive for large datasets	Cancer classification, medical imaging
k-Nearest Neighbors (k-NN)	Simple to implement, no training phase	Poor scalability, sensitive to irrelevant features	Disease classification, anomaly detection
Naive Bayes	Fast, works well with small datasets, handles missing data well	Assumes feature independence, can underperform with complex dependencies	Medical text classification, infection prediction
Gradient Boosting (e.g., XGBoost)	High predictive accuracy, handles various data types	Computationally expensive, less interpretable	Hospital readmission prediction, sepsis risk scoring
Neural Networks (Deep Learning)	Excellent for unstructured data (e.g., images, text), automatic feature extraction	Requires large datasets, low interpretability	Radiology, pathology image analysis, clinical natural language processing (NLP)

Calibration plots are also a critical element of model evaluation. A discriminatorially extremely proficient but poorly calibrated model will generate predicted probabilities that do not reflect the rate of observed events. Calibration plots plot observed frequency of events vs. predicted probability and make the risk estimates derived by a model meaningful and rational. It is particularly useful in clinical decision-making issues where an inappropriately calibrated model will mislead doctors and result in non-optimal treatment.

Their application against such parameters not only establishes their technical feasibility but also outlines the extent to which they can be applied. A technically feasible model must also be interpretable, justifiable, and integratable into clinician practices for making a positive difference. Thus measurement is quantitative and qualitative, not just numeric score format but to the degree in which the model operates in real practice settings, how clinicians use it, and how patients are affected by its predictions.

4. Explainability in Healthcare AI

The growing application of Artificial Intelligence (AI) in the health care system has brought with it unprecedented possibilities to enhance clinical decision support, diagnosis, and treatment planning. With AI algorithms becoming increasingly advanced—particularly through the common application of deep learning and ensemble methods—the challenge is making their output intelligible and credible. Explainability is not a nicety but a necessity in medicine. Clinical operators in the healthcare sector need to possess the ability of understanding, interpreting, and approving the decision of the AI technologies in the event that the decision has a direct influence on the outcome of the patients. Explainability of AI in healthcare has thus been of utmost importance with ethical, legal, and functional implications.

4.1. Need for Explainability

Healthcare AI definitely needs explainability as a result of the following reasons. Back of all this which is required is that clinicians must hold faith and know why AI outputs are so. That is not similar to some other field where risk shall be comparatively low. Medicine does have, of course, interventions that can change lives. Physicians carry immense responsibilities, and they cannot employ rubbish algorithms while prescribing drugs, making a diagnosis, or forecasting prognosis. Artificial intelligence systems that are black-box reasoning methods can be dismissed or even tampered with on grounds of ignorance.

Lastly, explainability is codified legislatively. Data protection laws such as those of the United States' Health Insurance Portability and Accountability Act (HIPAA) and the European Union's General Data Protection Regulation (GDPR) require transparency in implementing automated systems officially. GDPR, for instance, provides the "right to explanation," whereby automated decision-making individuals have the right to be explained in detail how the decisions were made. In medicine, it does not just have moral but also legal right to be informed of the route taken by an AI-based diagnosis or treatment plan. Explainability is not a theoretical notion but one of ethics and regulation towards safe and legal application of AI in the clinic.

4.2. Model Interpretation Techniques

In the quest for real-world explainability, different techniques towards interpretability have been found, moving towards global and local solutions. Global explainability has been characterized as understanding the model's general behavior and structure. It gives a glimpse of what most important is in aggregate for the entire set of data. Decision trees and feature importance rankings are two broad approaches to how one would arrive at those broad trends. Feature importance techniques rank the input features based on how much they contribute to model prediction. Clinically, for instance, that would be selecting most informative biomarkers or demographic variables most likely to discover disease risk models. Decision trees, particularly shallow and interpretable decision trees, are direct representations of decision logic, and doctors can trace the reasoning chain for a specific single output.

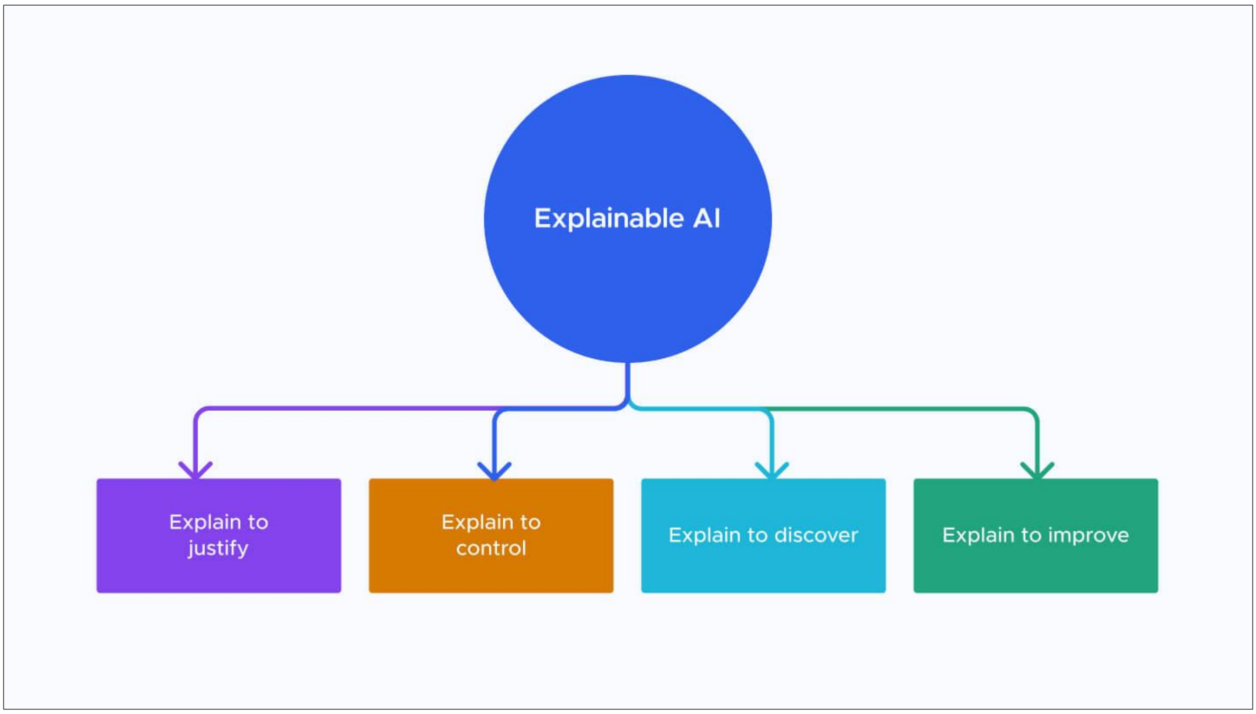

Figure 5 Local Explainability Approaches

Local explainability, however, is interested in explaining individual predictions. This is most clinically relevant where case-by-case explanation must be done. SHapley Additive exPlanations (SHAP), Local Interpretable Model-Agnostic Explanations (LIME), and Gradient-weighted Class Activation Mapping (Grad-CAM) are some of the methods that have been employed to analyze this function to some degree. SHAP employs methods from cooperative game theory to calculate a contribution of each of the input's features towards some particular prediction. LIME locally approximates richer, less interpretable models with proxy models that are easier to analyze and allows for graphically observing the effect of small perturbations in patient data on prediction. Grad-CAM, applied extensively in CNNs of radiologic images, identifies image areas most accountable for a classification result, e.g., indicating chest X-ray nodules or MRI scan lesions.

There is another crucial difference between model-specific and model-agnostic interpretability techniques. Model-agnostic methods can be used on any prediction model regardless of how it works and thus are generally great post-hoc tools. SHAP and LIME are model-agnostic tools and are popular as they can be used on any model. Model-specific methods are usable only for a given model architecture. Grad-CAM is a model-specific method used for CNNs. Model-specific solutions lowered transfer and increased explanatory power to generic from model-agnostic solutions. Clinical usage, model form, and requirement of explanation will guide use in clinical AI.

4.3. Human-AI Interaction

Clinical Explainable AI exceeds transparence algorithms and resets human-AI boundaries. Explanations allow clinicians to make use of AI systems as a partner rather than a black box with precision, transparence of explanations, and clinical usability of explanations. In doing so, they are able to provide collaborative decision-making wherein human judgment and machine learning are able to collaborate with each other. A system may predict early high-risk sepsis from suggestive indicators, for instance, and a clinician, receiving an insight into the presuppositions upon which the system has made a prediction so, could make predictive treatment choices in prudent discretion. Such complementarity proves especially useful in high-variability, time-sensitive situations such as emergency or intensive care.

Explainability is also relevant at the training and knowledge transfer phase. Explainable models are used by clinicians to gain observation data in blindness to routine clinical observation. Two-way learning facilitates improved AI system design and enhanced clinical practice. Explainability also pre-empts automation bias in case of biased over-trust with a lack of critical scrutiny of the algorithmic output. In explainable availability scenarios, physicians will ask questions, verify, and reject AI recommendations, thereby acting as a curb on mistakes.

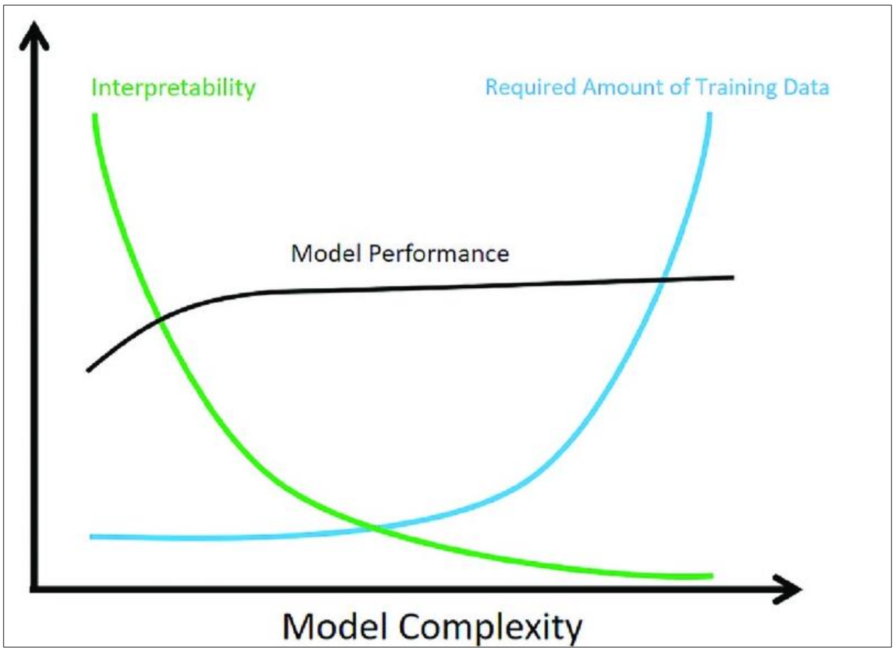

Figure 6 Machine learning model complexity and possible effects

Human-AI collaboration is the master narrative, and explainability is the terrain. Clinicians have been more inclined to embrace and utilize AI tools in practice when they can trace the origin of the tools. This is particularly so in new disease case management, atypical presentation, or atypical case where the AI predictions contradict the conventional medical intuition. Just such a situation is where an explainable system could be a bridge for life, one with evidence-based support for decision-making and therefore integratable with clinical judgment.

Second, explainability does align with patient-centered care. With more direct-to-patient functions being outsourced to AI, i.e., tele-diagnosis or health apps personalized to an individual's health, patients themselves must be able to see how their health data are being translated. Clinicians must be able to interpret AI output in a manner that is patient-meaningful. Transparency here is required for informed consent, shared decision-making, and maintaining the moral integrity of patient-clinician relationships.

5. Clinical Applications and Impact

Artificial Intelligence transformed the health sector by enabling diagnostic level activities enhanced to enable customized care and operational ease of efficiency. The use of Artificial Intelligence in medical practice not only transformed diagnosis but also transformed the functionality of the manner in which healthcare was being conducted, as much as efficiency is concerned. Artificial intelligence activities in different applications are some of the clinical applications, such as overall integration into disease diagnosis, personalized treatment regimes, and hospital finance operations. In the remainder of this paper, we describe further how AI accomplishes the above main areas, i.e., early disease detection and diagnosis, personalized treatment and prediction, and performance in operations with returns to the economy.

5.1. Early Disease Detection and Diagnosis

AI has done extremely well at diagnosing an extremely long list of diseases, even exceeding human capability in some cases. Its most unexpected application is likely in medical imaging. AI deep learning technology has been able to match the expertise levels of specialists in detecting diabetic retinopathy from a retinal scan. The technology is able to detect microaneurysms, hemorrhages, and other retinal defects with decent accuracy, and this enables on-time treatment and intervention. The technology is suited best for application in underdeveloped or far-flung areas where access to ophthalmologists is not feasible and therefore disease prevention and following blindness are made possible.

Similarly, in dermatology, AI has been employed to identify skin cancer like melanoma to the same level of accuracy as experienced dermatologists. Such machine learning algorithms having been trained on vast large sets of dermoscopic images will be able themselves to accurately classify benign versus malignant lesions with good confidence levels. This helps general practitioners and non-dermatologists to exclude the patients and refer them to dermatologists appropriately and start early treatments.

AI went even deeper in the COVID-19 pandemic period to assist in enabling rapid and accurate diagnoses from chest CT scans and X-rays. Computer software with AI detected viral lungs' behavioral patterns characteristic of their presence and allowed radiologists, thus, to scan enormous volumes of image studies in a matter of minutes. Where.presentation was early either subclinical or indefinite, these factors were back-up plans of ultimate significance in the future to be able to make a differentiation between COVID-19 and other respiratory infections and therefore to allow for proper isolation and control.

Overall, the ability of AI to analyze complex medical data and detect hidden patterns has established AI as a valuable tool for the early detection of disease. Not only does this capability improve patient care by enabling early intervention but also public health management by controlling the spread of infectious diseases.

5.2. Personalized Treatment and Prognostics

Other than diagnosis, AI is also leading the way when it comes to patient-specific treatment plans and disease progression prediction. Machine learning is used to predict drug response in patients from diverse data sources like genomic information, electronic medical history, lifestyle, and history. This is revolutionizing how the patient's treatment mode is being changed from one-size-fits-all to more enhanced and precise treatment regimen tailored solely to the patient's very own individualized biological and environmental profile.

In the oncology treatment, for example, AI technology can match genetic mutation of cancer cells to best drug or to best combination of drugs most likely to affect a patient. Besides maximizing chances of success, it reduces side effects and recovery times between the time of optimal solution identification and execution. Prognostic models also enable disease progression and survival prediction and hence assist patients and clinicians in decision-making based on best intensity of care and outcome desired.

Chronic disease management is another scenario where AI takes the front seat. In conditions like diabetes, congestive heart failure, and chronic obstructive pulmonary disease, readmission to hospital or an acute exacerbation is predictable based on real-time data on patients from wearable sensors integrated with past health information utilizing AI-based algorithms. The outcome enables pre-emptive treatment, lifestyle change, and drug adjustment prior to disease progression. By facilitating proactive chronic disease management, AI enhances the health condition of the patients and alleviates the healthcare systems' burden in the long term.

Table 3 Clinical Performance of AI in Disease Diagnosis

Condition	**AI Model**	**Accuracy**	**Sensitivity**	**Specificity**
Diabetic Retinopathy	Deep Convolutional Neural Network (CNN)	90.5%	91.2%	89.8%
Lung Cancer Detection	Ensemble Learning (e.g., Random Forest + SVM)	88.3%	86.7%	89.5%
Breast Cancer	Support Vector Machine (SVM)	93.1%	95.0%	91.2%
Skin Lesion Classification	Deep CNN (e.g., ResNet)	89.7%	87.9%	91.1%
COVID-19 (from CT scans)	Transfer Learning (e.g., VGG-19)	92.4%	94.3%	90.2%
Alzheimer's Disease	Recurrent Neural Network (RNN)	87.2%	85.5%	88.6%

AI is also leaving its stamp in the field of mental health. Predictive models are being created to screen out susceptible patients who can become depressed, anxious, or suicidal from behavioral patterns, word usage, and social media. Early detection leads to counseling and treatment before the onset of serious mental illness.

In brief, AI has the power to enable doctors to personalize treatment protocols and forecast disease courses with higher accuracy than is currently possible. This route to precision medicine has the potential to provide more, safer, patient-led care in numerous areas of medicine.

5.3. Cost-effectiveness and Operational Efficiency

Benefits of AI are also being implemented in design organizational in health facilities. AI technology is also changing hospitals by automating hospital patient triaging, enhancing resource planning streamlining, and applying predictive scheduling, saving much cost and improving efficiency.

Triagist systems are artificial computer systems powered by natural language processing, machine learning algorithms, to evaluate symptoms, history, and patient's vital signs and prioritize cases in order of urgency. It facilitates quick sorting by emergency rooms which enable the severe ones to be treated on priority and the non-severe ones to be treated with appropriate treatment. It enhances the clinical outcome and the patients do not have to wait for a long time, i.e., the patient has a good experience.

One of the fields where AI has made itself felt prominently is in the administration of resources. Through predictive analytics, bed space, patient admission volume, and medical supply or staff needs can be predicted. Staff can be streamlined, inventory streamlined, and bottlenecks avoided to flow of patients on these. Streamlined and effective care keeps downtime to an absolute minimum.

AI is also being used to book surgery and out-patient procedures in the optimum manner ahead of time by cross-checking history and considerations real-time in the ED to make predictions regarding no-shows, procedure duration, and recuperation. Intelligence-assisted scheduling optimizes operating room utilization, delay is reduced, and cancellations are reduced, all of which are cost-efficient and more efficient health care provision.

Financially, it avoids billing discrepancies from occurring and insurance fraudulence. Computer systems can sort out treatments and claims for billing and detect discrepancies which are evidence of fraud or buggy billing. It enables health providers to recover lost revenues and satisfy regulations.

Administrative tasks such as documentation, coding, and reporting are also performed by AI in a manner that workload is shifted from the medical professionals so that they can devote more time to taking care of patients. Even electronic health records, clinical notes, even discharge summaries can be dictated using virtual assistants and speech recognition software with fewer hours and less errors.

6. Challenges and Limitations

Although Artificial Intelligence (AI) being applied in medicine is great with space for creativity, productivity, it is circumscribed based on a relationship of border and wholeness that must be approached with the highest level of cautiosness. Not only are they technology-wise advanced but deeply rooted in the social, ethical, and regulatory environment under which medicine is practiced. With AI technology used in clinical decision support, patient diagnosis, and treatment, reliability, fairness, and safety take center stage. The latter half of the book covers the most viable counterarguments to wide-ranging, long-term applications of AI in medicine, i.e., data bias and quality, ethics and regulation, and generalization and test problems in the real world.

6.1. Data Quality and Bias

The most difficult hurdle to the creation of highly trained and healthy medical AI models is availability, diversity, and representativeness of training data. There are no medical data available, biased, and in small volume. There are gaps due to the fact that there are enormous numbers of variables ranging from healthcare variability to compatibility problems in EHR standards and biased past medicine practice. When machine-learning training data sets fail to include some demographic groups—racial minorities, the very old, or patients with rare diseases, for instance—their models will learn gargantuan biases now and then. They'll resemble biased risk profiles, biased diagnoses, or biased treatment recommendations and can amplify or even worsen health inequities.

Second, even if data are representative and large, implicit sampling, organizational policy, or socioeconomic determinant biases can be baked into data capture. For example, a highly urban hospital learning pattern from data will not be applicable to rural or low-resource environments. Similarly, AI systems will learn inherent gender, race, or age bias in data from existing clinical history application contexts. These disparities emphasize the importance of data auditing, bias detection, and data enriching processes in validating certified AI models generating balanced and accurate health outcomes.

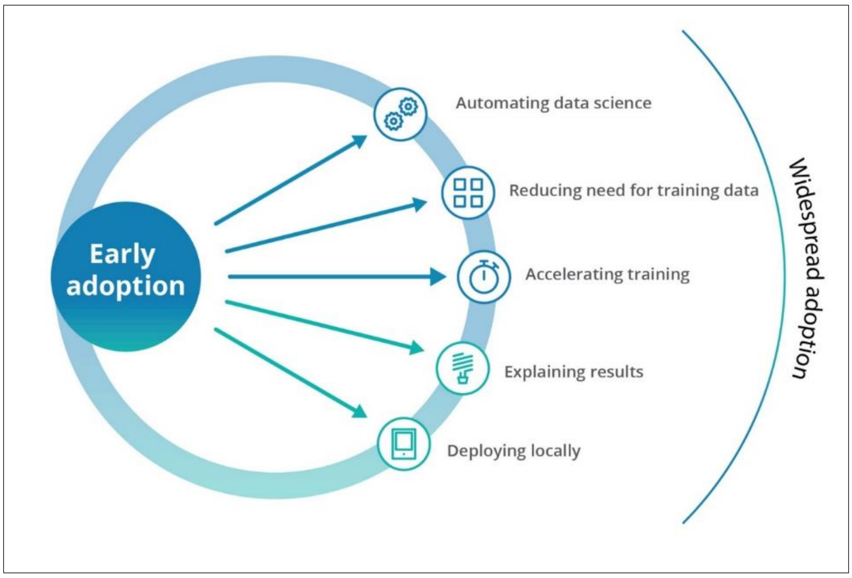

Figure 7 Barriers to AI dissemination and implementation in medicine

Noise, inaccuracy, and inconsistency of clinical data are also a quality issue with the data. Inconsistent input in capturing data, data loss in data capture, and format incompatibility are a few issues that cause most disturbance to training and model validity. Generalizability and precision of forecast models would be affected negatively if data are not preprocessed and cleaned correctly. Data quality issues are therefore a high-priority category in developing stable and interpretable AI systems across the health care continuum.

6.2. Regulation and Ethical Issues

With artificial intelligence playing an active role in medical decision-making, it is increasingly imperative that its use and development be out of the reach of regulation and ethics. Governments globally, including the U.S. Food and Drug Administration (FDA) and European Medicines Agency (EMA), have already begun to draft guidelines for computer programmes and medical devices incorporating AI. Sadly, regulatory systems still function in splinter mode and publicly accessible standards fall behind technology innovations.

Among the biggest questions before regulation is whose medical decisions are made by AI. If responsible AI diagnosis harms patients, no guide suggests that fault will be with developers, clinicians, or hospitals using technology. These issues must be solved by developing well-delineated boundaries of law that establish parameters of liability and who will bear responsibility for testing of AI systems before they are used in clinical practice.

Even on ethical grounds, using AI in treating cancer is problematic on the points of informed consent, patient autonomy, and transparency. Clinicians and patients both need to see why and how an AI program has arrived at its conclusion—particularly when such life-or-death choices are concerned like cancer diagnosis or organizing treatment for cancer. Explainable AI (XAI) is the answer. Explainable AI methods try to explain black-box models such that stakeholders will be in a knowledge state after listening to the explanation of the prediction. But for highly complex models, i.e., deep neural networks, explainability comes at a cost. Extremely brief explanations are clinically useless, and complete explanations are end-users' cognitive cost.

Moreover, following ethically sound AI development also involves ongoing stakeholder engagement including regulators, ethicists, patients, and clinicians. Open report practice, ethics review board governance, and participatory design can guarantee AI systems not only give the public good but also get developed in central healthcare values. Public trust will be lost for AI systems, and clinical worth will be forfeited unless maintained.

6.3. Generalization and Real-World Validation

The second substantial limitation of AI in medicine is that behavior on highly controlled experimental conditions is difficult to transfer to actual clinical practice. AI models are trained and tested using historical data from best-case scenarios in which data are heavily labeled, complete, and standardized. Clinical practice is less certain, more heterogeneous, and more dynamic though. These are concerns like biased patient groups, clinical practice guideline variation, and biased enrollment processes that have profound implications on model behavior in practice.

Generalizability is most when models are deployed in a large number of settings which were trained in a single setting of healthcare. The example is the case of training a machine learning model on hospital data based in America; it would provide weaker estimates when implemented in the European or Asian healthcare system based on varying patterns of disease, practice, diagnosis, etc. Setting dependency mentioned above compromises scalability and integrability through region-level database retraining or local model fine-tuning.

Global validation is the process of drawing conclusions about the safety, stability, and clinical utility of an AI system in real-world-use settings. It is not a trivial thing. It is all access to future data, end-user usage, and very tight control over privacy and ethics. In addition, following deployment the performance of the AI model—alternatively, even model performance following so-called market or drift model performance—ought to be tracked continuously by continuous performance drift tracking, data set shift, or long-term, unexpected effect.

All these on top of all the above failures. Frustration with implementing AI systems into deployable clinical workflow is on top of all the above failures. Even more precise models won't be applicable unless they're understandable and won't coexist in clinician decision processes. That's where the role of human-centered design and the addition of a number of rounds of usability testing throughout the course of AI development come into play. Only after AI has been implemented and applied in clinician workflow and prediction made understandable and interpretable can the systems start to influence clinical practice.

7. Future Directions

Along with the healthcare revolution, deployment of artificial intelligence (AI) into medical decision-making is growing by an order of magnitude. Nevertheless, realization of completely reliable and trusted AI systems remains pending in front of us. The most externally visible disciplines to establish themselves within the coming years are likely to set the tone for the innovation. While these fields are going to make AI systems efficient and powerful, they are also going to solve technical, operational, and ethical issues. Of them, the most influential ones that are going to leave their impact on future development are federated learning, causal inference, AI model clinical trials, and IoT and edge computing integration.

Federated learning is an emerging paradigm that is revolutionizing machine learning model training, particularly in data-sensitive healthcare systems. Centralization of data and consequent severe privacy issues render the deployment of conventional models redundant. This can be credited to federated learning as it possesses the ability to provide decentralized training by preserving data at a local level by the healthcare organizations and model parameters or updates that are passed via an encrypted network. This not only respects patient privacy but also allows training models using heterogenous data sources without data centralization. It empowers research institutions, physicians, and hospitals to collaborate on data without invading the data and ownership. Privacy-related regulatory laws such as HIPAA and GDPR are becoming increasingly restrictive, and hence federated learning is increasingly becoming more and more a vision of the future that still remains legal and ethical. Briefly, federated learning fosters model generalizability by interpolating data from diverse populations and sites of clinical acquisition, and therefore the ability to generalize well across multiple diverse patient populations. As technology continues to improve, actions will be inevitably focused on reducing communication overhead, enabling model convergence, and building efficient security defense against attacks by opponents.

The next wave of AI in health care is causal inference, the shift beyond present detection of statistical correlation to the detection and determination of actual causal relations. In most uses at present, AI models are very good at pattern recognition in broad bases but confuse causation and correlation in the middle. It can result in deceptive conclusions, especially when models are used to guide treatment or disease diagnosis. Causal inference attempts to achieve this by representational structure such as structural causal models and counterfactual logic. These allow us to disentangle the effect of a particular variable or intervention from confounding and hence have access to what would have occurred if things were otherwise. For example, rather than simply telling us that patients with a given biomarker will be at risk for a disease, a causal model tells us whether or not the course of the disease will be changed by changing the biomarker. This shift from descriptive to prescriptive analytics is exactly what is at the heart of precision medicine where therapeutic suggestion must be aimed at an individual causal pathway and not at population trends. This intersection between causal inference models and artificial intelligence will have to be tackled by multi-disciplinary efforts of data scientists, domain experts, and statisticians so that the models are clinically useful and statistically robust. Follow-up action would likely be on how to mechanize causality discovery and how to link conclusions from observation research and conclusions from experiments to provide more and relevant models.

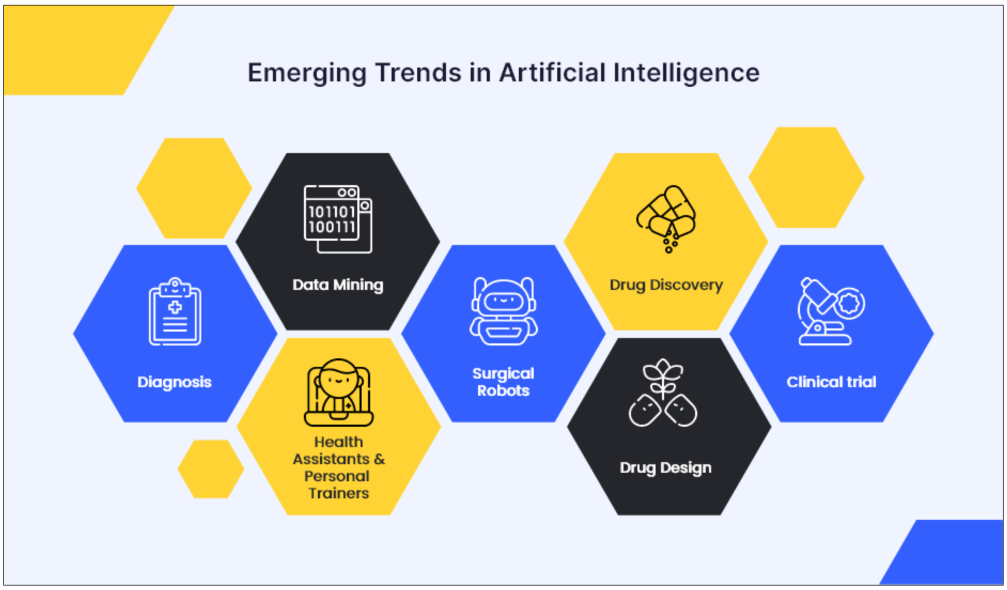

Figure 8 AI In Healthcare: AI Transformation the Healthcare Industry

One of the most important actions toward clinical use of AI technology is putting them through tough tests like drugs and medical devices. Clinical trials are something that has to be completed by AI algorithms so that their safety, effectiveness, and reliability can be tested before they are deployed in real-world healthcare centers. In comparison with conventional software, artificial intelligence-based systems to support medical decision-making would require to be consistent to apply across heterogeneous groups of patients and care settings. Formal test protocol development, prospective validation studies, randomised controlled trials, and post-marketing surveillance would be a part of it. Open publication of trial results and side effect will be necessary in an attempt to enable generation of confidence on the part of healthcare professionals and regulators. In addition, adaptive trial designs permitting learning and model update upon implementation can bridge the gap between static test environments and dynamic operational real-world environments. Strong regulatory models must adapt to include these newer components with potential to necessitate new classes or updates to existing pathways by regulatory agencies like the FDA. This, in the future, will be an exercise that takes place globally to standardize clinical trials procedures for AI software in order to provide more usage and worldwide cooperation.

Finally, integrating AI with Internet of Things (IoT) and edge computing holds out new possibilities for real-time health monitoring and adaptive treatment. Greater use of wearable sensors, remote patient monitoring platforms, and smart medical devices creates data streams that, when AI-infused, can yield hitherto unknown insights into patients' health. Edge computing, with its extension of data processing to the edge, removes latency and provides real-time analysis and response. It is especially valuable for use in such applications as intensive care vital sign monitoring, fall detection in elderly, or insulin level monitoring in diabetic adults. By allowing AI models to run on edge devices, health systems can render decision-making instantaneous, consistent, and cloud connectivity-independent. This also enhances data security as sensitive health data no longer has to be transmitted over public networks. Future space development will also continue to focus on reducing model size and power consumption to leave room for edge hardware constraints as well as further enhance the interoperability between artificial intelligence frameworks and medical IoT devices. Furthermore, self-adaptive learning algorithms based on periodically renewed data can react timelier and more tailored.

8. Conclusion

Artificial Intelligence (AI) deployment in health systems is one of today's most promising medical innovations. As technology evolves rapidly, the potential for clinical practice transformation increases daily. Two pillars are the cornerstones of such revolution: explainability and predictiveness modeling. They are the cornerstones of safe, effective, and sustainable implementation of AI in clinical practice. While predictive modeling enables clinicians to forecast patient outcomes, risk stratify patients, and individualize treatments, explainable AI (XAI) makes the predictions transparent, interpretable, and understandable—especially in high-stakes fields like human lives.

Predictive modeling has been shown to have unparalleled utility across various clinical problems in a broad spectrum of disease categories, including diagnosing early diseases, risk stratification, estimating treatment effects, and optimizing hospital resources. These models use large volumes of organized and unorganized health information to determine relationships and correlations that would generally be opaque to individual practitioners. Using data-driven wisdom, predictive models amplify the ability of practitioners to make timely, evidence-based choices, which are then translated into improved patient results. However, Predictive models must be carefully validated and closely monitored to ascertain if they generalize well to heterogeneous patient populations and shifting clinical environments.

All this added sophistication and uninterpretability of AI algorithms, specifically deep learning-based algorithms, is a primary barrier to their adoption. Even as excellent as they are at times, black-box models are opaque, and their opacity can collaborate with clinician and patient distrust. Uncertainty is ethical to start with anywhere AI systems have some control or even substitute human judgment. This is the very nature of human medicine, where, in such instances, explainable AI becomes not only a technical necessity but also an ethical necessity. XAI methods attempt to span this chasm by making the model's path to arrive at correct predictions or suggestions transparent. This supports clinician trust-building and patient comprehension and enables accountability in choosing. Beyond that, explainability is required for regulatory management so that AI systems satisfy safety, fairness, and reliability requirements. The explainability of AI has a lot in common with the clinical utility of AI. Reliable and transparent AI predictions will have more applications in the real world. For example, an AI system to predict the likelihood of sepsis for a patient in an ICU can be life-critical, but only if the reason why the warning was triggered has been started is it explainable to the medical professionals. Without transparency, clinicians will inevitably either disregard good advice or heed poor advice, neither of which is acceptable from a patient safety standpoint. Therefore, the intersection of predictive modeling with solid XAI platforms is the secret to unlocking the complete clinical promise of AI.

Technology development alone is insufficient to deploy AI in medicine effectively. Additional root system systemic, organizational, and ethical barriers must be addressed. Availability and data quality are persistent problems. Healthcare data collections frequently are disjunctive, absent, or biased and may result in biased model output and amplification of health inequities. Creating representative, high-quality processes is the priority for equitable AI system construction. Additionally, testing must be conducted in operational clinical environments to facilitate AI tool generalizability and robustness. This involves continuous end-user-developer collaboration to validate the model's clinical utility and consistency with changing healthcare demands.

There are also ethical issues with the use of AI in medicine. Informed consent, patient confidentiality, algorithmic bias, and the risk of over-reliance on machine systems are some of the problems that must be managed seriously. Explainable AI solves some of these issues because it allows transparency and accountability when decisions are made. Effective use of AI also requires robust governance frameworks, open rules, and active patient and public engagement.

Lastly, the successful deployment of AI in medicine necessitates interdisciplinarity. Data scientists can give technical expertise towards model building and training, clinicians can offer clinical experience and knowledge, and policymakers can offer regulatory and ethical input to make adoption appropriate. Together, these partners can design AI systems that are not just cutting-edge but also safe, dependable, and grounded in the very principles of the practice of medicine.

Compliance with ethical standards

Disclosure of conflict of interest

If two or more authors have contributed in the manuscript, the conflict of interest statement must be inserted here.

References

[1] Jiang F, Jiang Y, Zhi H, Dong Y, Li H, Ma S, Wang Y, Dong Q, Shen H, Wang Y (2017) Artificial intelligence in healthcare: past, present, and future. Stroke Vasc. Neurol 2:e000101

[2] Yang CC, Yang H, Jiang L (2014) Postmarketing drug safety surveillance using publicly available health consumer contributed content in social media. ACM Trans Manag Inf Syst 5:1

[3] Nasajpour M, Pouriyeh S, Parizi RM, Dorodchi M, Valero M, Arabnia HR (2020) Internet of things for current COVID-19 and future pandemics: an exploratory study. J. Healthc. Inform Res 4:325–364. https://doi.org/10.1007/s41666-020-00080-6

[4] Xu J, Glicksberg BS, Su C, Walker P, Bian J, Wang F (2021) Federated learning for healthcare informatics. J. Healthc. Inform Res 5:1–19

[5] Liu Z et al (2020) A deep reinforcement learning approach for type 2 diabetes mellitus treatment. IEEE Int. Conf. Healthc 2020:1–9. https://doi.org/10.1109/ICHI48887.2020.9374313

[6] Krieg SJ, Robertson DH, Pradhan MP, Chawla NV (2020) Higher-order networks of diabetes comorbidities: disease trajectories that matter. IEEE Int. Conf. Healthc 2020:1–11. https://doi.org/10.1109/ICHI48887.2020.9374390

[7] Zhu T, Li K, Chen J et al (2020) Dilated recurrent neural networks for glucose forecasting in type 1 diabetes. J. Healthc. Inform. Re 4:308–324. https://doi.org/10.1007/s41666-020-00068-2

[8] Gulshan V, Peng L, Coram M et al (2016) Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. JAMA 316(22):2402–2410. https://doi.org/10.1001/jama.2016.17216

[9] Jia Y, Burden J, Lawton T, Habli I (2020) Safe reinforcement learning for sepsis treatment. IEEE Int. Conf. Healthc 2020:1–7. https://doi.org/10.1109/ICHI48887.2020.9374367

[10] Yu C, Ren G, Liu J (2019) Deep inverse reinforcement learning for sepsis treatment. IEEE Int. Conf. Healthc 2019:1–3. https://doi.org/10.1109/ICHI.2019.8904645

[11] A. Jazayeri, M. Capan, C. C. Yang, F. Khoshnevisan, M. Chi, and R. Arnold, 2019 “Network-based modeling of sepsis: quantification and evaluation of simultaneity of organ dysfunctions,” Proceedings of ACM Conference on Bioinformatics, Computational Biology, and Health Informatics, Niagara Falls, September 9 – 10

[12] Hu P, Li S, Huang Y, Hu L (2019) Predicting hospital readmission of diabetics using deep forest. IEEE Int. Conf. Healthc 2019:1–2. https://doi.org/10.1109/ICHI.2019.8904556

[13] Xue Y, Klabjan D, Luo Y (2019) Predicting ICU readmission using grouped physiological and medication trends. Artif Intell Med 95:27–37. https://doi.org/10.1016/j.artmed.2018.08.004

[14] Kim H, Liang OS, Yang CC (2020) Detecting potential adverse drug reactions of preschool ADHD treatment using health consumer-generated content. IEEE Int. Conf. Healthc 2020:1–6. https://doi.org/10.1109/ICHI48887.2020.9374395

[15] Kavuluru R, Rios A, Tran T (2017) Extracting drug-drug interactions with word and character-level recurrent neural networks. IEEE Int. Conf. Healthc 2017:5–12. https://doi.org/10.1109/ICHI.2017.15

[16] H. Yang and C. C. Yang, 2016 “Discovering drug-drug interactions and associated adverse drug reactions with triad prediction in heterogeneous healthcare networks,” Proceedings of IEEE International Conference on Healthcare Informatics, Chicago, IL, October 4 – 7

[17] Yang H, Yang CC (2015) “Using health consumer contributed data to detect adverse drug reactions by association mining with temporal analysis,.” ACM Trans. Intell. Syst. Technol 6:4

[18] Zhao M, Yang CC (2018) Exploiting OHC data with tensor decomposition for off-label drug use detection. IEEE Int. Conf. Healthc 2018:22–28. https://doi.org/10.1109/ICHI.2018.00010

[19] Yang CC, Zhao M (May 2019) Mining heterogeneous networks for drug repositioning using phenotypic information extracted from social media and pharmaecutical databases. Artif Intell Med 96:80–92

[20] Yonatan E, Rosenberg AA, Schuster A, Bronstein AM, Yaniv Y (2021) “Meeting the unmet needs of clinicians from AI systems showcased for cardiology with deep-learning–based ECG analysis,.” Proc. Natl. Acad. Sci 118(24):e2020620118. https://doi.org/10.1073/pnas.2020620118

[21] Prashan Madumal, Tim Miller, Liz Sonenberg, and Frank Vetere. 2019. A grounded interaction protocol for explainable artificial intelligence. In Proceedings of the 18th International Conference on Autonomous Agents and MultiAgent Systems (AAMAS ’19). International Foundation for Autonomous Agents and Multiagent Systems, Richland, SC, 1033–1041.

[22] Miller T (2019) Explanation in artificial intelligence: insights from the social sciences. Artif Intell 267(2019):1–38

[23] Viton F, Elbattah M, Guérin J-L, Dequen G (2020) Heatmaps for visual explainability of CNN-based predictions for multivariate time series with application to healthcare. IEEE Int. Conf. Healthc 2020:1–8. https://doi.org/10.1109/ICHI48887.2020.9374393

[24] Q. Vera Liao, Daniel Gruen, and Sarah Miller. 2020. Questioning the AI: informing design practices for explainable AI user experiences. In Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (CHI ’20). Association for Computing Machinery, New York, NY, USA, 1–15. DOI:https://doi-org.ezproxy2.library.drexel.edu/https://doi.org/10.1145/3313831.3376590

[25] Yu K, Bean AL, Kohane IS (October 2018) “Artificial intelligence in healthcare”, Nature. Biomed Eng 2:719–731

[26] Deo RC (2015) Machine learning in medicine. Circulation 132:1920–1930

[27] Lindsell CJ, Stead WW, Johnson KB (2020) Action-informed artificial intelligence—matching the algorithm to the problem. JAMA 323(21):2141–2142. https://doi.org/10.1001/jama.2020.5035