Vol No.1, Issue 3, 2025
ACADEMIA International Journal of Clinical & Translational Health

Application of Machine Learning Models to Predict Clinical Outcomes in Low-Resource Settings

Nauman Ali Ch [a]

Assistant Professor Public Health, Department of Public Health, University of the Punjab, Lahore, Punjab, Pakistan, (54000),(naumanali.iscs@pu.edu.pk)

Received: 21-06-2025 **Revised: 19-07-2025** **Accepted: 13-08-2025** **Published: 26-08-2025**

Corresponding Author: *Nauman Ali Ch*

DOI: 10.63056

ABSTRACT

Healthcare systems in low-resource environments are usually challenged by limited medical staff, poor diagnostic tools, and delayed clinical decision processes. These restrictions add to the risk of preventable morbidity and mortality. Recent developments in machine learning (ML) provide promising prospects to assist healthcare workers with the predictive analysis of the outcomes of clinical cases based on routine patient data. This study investigates the use of ML models (logistic regression, random forests, support vector machines and artificial neural networks) to predict key factors such as disease course, risk of death, race, response to treatment and risk of readmission to the hospital. Based on evidence in low and middle-income settings (LMIC), the main message of the paper is the potential of ML to help improve the accuracy of diagnostics, optimize triage and resource allocations. However, challenges such as the lack of a large quantity of data, poor quality of data, limited computational infrastructure and lack of technical expertise have impeded large-scale implementation. The study concludes that integration of ML models into clinical workflows could play a major role in improving patient outcomes that they say, in low resource setting futures, as long as ethical considerations, stakeholder training and context-specific model customisation are prioritized.

Keywords: Machine learning; clinical outcomes; low resource settings; predictive analytics; LMICs; healthcare technology; artificial intelligence; clinical decision support systems.

INTRODUCTION

Machine learning (ML) has become one of the most significant technological innovations in healthcare around the world, as it allows systems to process complex patterns and spot early signs of disease and support clinical decision-making (Topol, 2019). In the high income countries, machine learning (ML) algorithms are becoming more and more ingrained into diagnostic imaging, risk prediction models and individualised treatment schemes. However, the use of these technologies in the low-resource world, including rural hospitals, underserved areas, and low-income countries, is an evolving, but promising, area (Rajkomar et al., 2019). These regions often have huge structural challenges such as shortages of medical personnel, access to laboratory facilities, availability of electricity and access to advanced medical equipment. As such, health workers can hardly make on-time decisions leading to preventable complications and death. The opportunity to use ML to predict clinical outcomes is to fill these gaps by allowing healthcare providers to process standard clinical data and derive actionable insights.

Low-resource settings often use simple diagnostic indicators, simple paper-based medical records and limited electronic databases (WHO, 2020). Despite these limitations, they have vast untapped amounts of data from outpatient visits, community health workers and local health facilities. ML models can use these datasets in order to identify high-risk patients, forecast the course of the disease and therefore to recommend appropriate intervention actions. For example, ML has shown potential in evolving neonatal mortality, maternal complications, infectious disease epidemics, treatment failure for chronic diseases like tuberculosis and HIV (Esteva et al., 2019) and so forth. Predictive analytics allow comparison of vulnerable patients to be identified at an early stage, allowing triage decisions to be better informed and limited resources are used more wisely. In setting where one doctor may be treating thousands of patients, ML-based systems can serve as decision support systems, increasing the efficiency of the clinical setting.

The application of ML in the low-resource setting already showed promising results in several LMICs. For example, simple ML models have been used in predicting maternal mortality in rural sub-Saharan Africa using only antenatal care data, including logistic regression and decision trees (Obuh et al., 2021). Similarly neural networks have also been used for the classification of chest x-rays in order to detect tuberculosis in South Asia with performance similar to trained radiologists (Lakhani & Sundaram, 2017). These are just a few examples that show promise for ML to overcome in the diagnostic world, where expert clinicians are not always readily available. Moreover, the affordability of smartphones, cloud storage and open-source ML frameworks (e.g. TensorFlow, PyTorch, Scikit-learn) offer an opportunity to use that for scalable implementation without the need of high end computational resources (Abebe et al., 2022). This democratization of AI tools is of special interest to the development of sustainable health technologies adapted to resource deficient environments.

Despite these promising developments, there are a number of major challenges related to the use of ML models in low resource settings. One of the greatest concerns is that of data quality. Many hospitals do not have a standardized data entry process, data entry is incomplete or incorrect, and it results in incomplete or incorrect patient records (Sarkar et al., 2022). Poor quality data used to train ML models can result in unreliable model predictions, which can be harmful for making clinical decisions. Another challenge is the challenge of having local technical expertise to build, validate and maintain ML systems. Most healthcare workers in resource-poor settings lack formal education on AI, and no training programs are available (Adedeji et al., 2021). Ethical challenges - such as bias in AI models, data privacy, and patient consent - also need to be given careful consideration so that the application is safe and equitable.

Furthermore, ML models trained on data collected from high-income countries might not generalize well to low-resource settings because of variations in the prevalence of diseases, genetics, socio-economic conditions and healthcare practices (Rajkomar et al., 2019). Therefore, the use of context-specific development of models, local validation and continuous performance monitoring is critical. Integrating ML systems into the clinical workflow also demands the involvement of stakeholders, capacity building, and investing in basic digital infrastructure. With an absence of adequate implementation strategies, it is possible that the possibility of transforming healthcare through ML may simply fail.

Nevertheless, ML remains considered an important tool to enhance the capacity of health systems in low-resource settings. It has the potential to automate routine screening, assist community health workers, create early warning systems in an outbreak and inform intervention plans for public health agencies (WHO, 2020). By predicting clinical outcomes more accurately, ML technology can help save healthcare costs, prevent healthcare complications and eventually lives. As more efforts are made globally in the form of digital health initiatives, predictive ML models have a lot to contribute being able to contribute to universal health coverage and solve health inequities.

In summary the use of ML models to competence van PAS for clinical outcomes in residing low resource situations is forthcoming and demanding. While the development of new possibilities from advancing technology in the form of data-driven healthcare delivers new hope for more effective healthcare outcomes, it must come with robust implementation governance, reliable data pipeline, and a multidisciplinary approach of cooperation. Specifically, this paper sets out to review the potential, limitations and feasible pathways for embedding ML prediction models into the healthcare systems of low-resource environments as part of the burgeoning claims around global digital health innovation.

LITERATURE REVIEW

Machine learning (ML) applications in healthcare have been on a rapid ramp since the last decade with the solutions provided by the technology in the field of healthcare: disease detection, risk forecast, and clinical decision support. While high-income countries have seen massive research work and implementation of ML-driven healthcare solutions, the importance of ML in low-resource settings has only recently become a topic of discussion (Rajkomar et al., 2019). Such regions are faced with unique challenges, such as understaffed health facilities, lack of diagnostic tools and inconsistent electronic medical records, which make efficient clinical decision making difficult (World Health Organization [WHO], 2020). This literature review summarizes existing research on the application of ML for predicting clinical outcomes in low-resource settings, paying particular attention to the performance of the models, use cases, challenges, and the gaps in the context that still need addressing.

Predictive Analytics: Importance in Low Resource Settings

Clinical prediction models have an important effect on bettering the delivery of healthcare in settings where physicians, for example, have minimal diagnostic tools. Studies show predictive analytics could help reduce mortality, help improve the disease surveillance and help in early identification of risks (Topol, 2019). In low-income and middle-income countries (LMICs), ML-based predictive systems have been suggested to be cost-effective alternatives to high-technology medical equipment (Abebe et al., 2022). The capability of ML-based models to learn patterns suggested in routine health data, such as vital signs, demographic data, and lab results, makes them applicable in any type of hospital or rural clinics that lack sophisticated infrastructure.

Research done in Nigeria showed that ML models using basic antenatal care data were able to predict maternal complications with reasonable accuracy for early referral and intervention (Obuh et al., 2021). Similarly, in rural India, ML-based systems have been developed to predict neonatal sepsis based on clinical observational data alone to better the early detection of sepsis where laboratory testing is poor (Kumar et al., 2020). These examples indicate that using ML-based prediction is not only possible, but potentially transformative for environments where traditional diagnosis pathways are not only slow, but unfruitful or non-existent.

Machine Learning Models Used in Prediction in Clinical Medicine

There are a number of ML models that were widely explored in Healthcare. Classical models such as logistic regression, decision trees and support vector machines (SVMs) are still popular because of interpretability and low computational needs - important in low-resource settings (Sarkar et al., 2022). Logistic regression has been used for the prediction of maternal mortality, postoperative complications and infectious disease with the use of minimal datasets (Obuh et al., 2021). Its transparency helps health workers see clearly what are the key risk factors.

Tree based models like Random Forests have similarly shown good predictive performance under several low-resource scenarios. For example, Random Forest algorithms have been applied to the prediction of HIV treatment failure in sub-Saharan Africa based on demographic and clinical variables and have been shown to be superior to traditional clinical scoring techniques (Musabende et al., 2021). These models glorious handle missing data relatively well, which is a significant pro if one considers where patient records are often incomplete.

More sophisticated models like artificial neural networks (ANNs) and deep learning models (e.g. convolutional neural networks) have demonstrated phenomenal accuracy in image-based diagnosis. Research conducted on tuberculosis detection based on chest X-rays in south Asia demonstrated that neural networks could match the sensitivity and specificity of radiologists and lessen the need for experts (Lakhani & Sundaram, 2017). However, deep learning also demands larger amounts of data and computational resources, which could restrict deep learning's usability in facilities with weak digital infrastructures.

ML Applications in Disease Diagnosis and Outcome Prediction

Maternal and Neonatal Health

Maternal and neonatal complications are significant factors of death in LMICs. ML models have been found to be effective with antenatal records, vital signs and delivery histories to predict postpartum hemorrhage, eclampsia, premature birth and neonatal mortality (Allison et al., 2018). A study conducted in Ethiopia, used ML to predict stillbirth with 13 routine collected variables or parameters and over 80% efficiency or accuracy (Dadi et al. 2020). These findings highlight the possibilities in a data-driven approach to maternal care when the medical expertise can be limited.

Infectious Diseases

Infectious diseases are still prevalent in areas of low resources, making predictive analytics an essential part of early diagnosis and control of infectious disease outbreaks. ML models have also been applied to the prediction of the presence of malaria infection based on the information from the environment and clinically observed patients, thereby identifying high-risk populations even before the signs of the disease appear (Kweka et al., 2020). The application of ML in the prediction of the progression of tuberculosis has also gained momentum. Deep learning models analyzing chest X-rays have been shown to improve the screening accuracy, and are in some instances deployed on mobile diagnosing units where they can be used to reach rural areas (Lakhani & Sundaram, 2017).

In HIV treatment, ML has been applied to predict virologic failure, treatment regimens, and adherence (Musabende et al., 2021). These models can help healthcare workers prioritize patients for follow up; thereby, improve the long-term outcome with limited resources.

Chronic Diseases and Mortality Prediction

The LMICs are becoming more exposed to chronic diseases like diabetes, hypertension and cardiovascular diseases. Early risk identification based on interpretation of simple indicators such as blood pressure, body mass index, glucose levels, has been done through predictive models when there is a shortage of laboratory facilities (Adebayo et al., 2022). ML-based mortality prediction models have been built in rural hospitals to predict the critically ill patients based on a simple triage data. An Rwanda study showed that Mortality prediction accuracy by ML-based triage systems was better than traditional clinical tools (Adedeji et al., 2021).

Challenges in Implementing ML in Low-Resource Settings

Data Quality and Availability

The poor quality of data is one of the key challenges. The quality of ML models is limited by incomplete records, irregular formats, and values that are not present (Sarkar et al., 2022). Most health facilities are still using paper-based records, which limit the accessibility of structured, digitized data. Also, larger representative datasets are lacking which is a problem with generalizability of models.

Infrastructure and Technical Capacity

Lots of low-resource environments do not have the digital infrastructure needed to run ML systems, such as reliable electricity, access to the internet, and computing power (WHO, 2020). There is also a lack of technical expertise; not many healthcare professionals are trained on AI systems, and hence implementation and maintenance are challenging (Abebe et al., 2022).

Ethical and Legal Concerns

ML models can be biased in training data, which results in discriminatory results. The challenge of data privacy, informed consent, and transparency of the algorithms need straightforward rules prior to expanding ML technologies (Rajkomar et al., 2019). The most sensitive groups that require ethical issues are the vulnerable ones where misuse of information may result to exploitation.

Generalizability of Models

Most ML models designed in high-income areas cannot be transferred to LMICs because of the differences in the patterns of diseases, genetic factors and socio-environmental conditions. It is crucial to include local validation and context-specific model training, but it is usually overlooked (Obuh et al., 2021).

Opportunities and Future Directions

Nevertheless, the world is turning to digital health faster, and scalable ML solutions can be an opportunity. The cost-effective deployment is made possible by cloud-based solutions and low-cost mobile technologies as well as open-source frameworks (Abebe et al., 2022). Data collection and capacity building can be carried out with the help of collaboration among governments, NGOs, and academic institutions. Moreover, explainable AI (XAI) is becoming an enticing trend to make clinicians more confident and guarantee the levels of transparency of predictions made by the ML.

The other technology is federated learning, which enables hospitals to learn common ML models without sharing their sensitive patients and maintain privacy but enhance the robustness of the model (Kairouz et al., 2021). These inventions would provide great improvements in the sustainability and ethical use of ML in low resource settings.

Current literature is very optimistic about how machine learning can transform low-resource health care environments by means of precise clinical outcome prediction, earlier diagnosis, and improved allocation of resources. Although several ML frameworks (such as logistic regression to deep neural networks) have shown positive outcomes in maternal health, infectious disease, and chronic disease management, notable obstacles still remain. Limitation of data, poor infrastructure, risk of ethics and limited generalizability of models are all substantial issues. According to the literature, the further development will be based on the context-dependent development of the models, capacity building, ethical governance, and multidisciplinary cooperation. ML has a great potential, and its successful application depends on the responsible and sustainable approach to the realities of the low-resource settings.

METHODOLOGY

Research Design

In this case, the research design adopted was quantitative and predictive based on secondary clinical data collected in healthcare facilities that operate in low-resource settings. The choice of design was based on the fact that machine learning (ML) methods demand numerical, structured information to generate dependable predictive modeling on clinical outcomes.

Study Population and Sample Size

The sample size of the dataset was 200 records of patients, which were gathered in primary and secondary healthcare units in Lahore. Data consisted of demographic patient data, clinical signs, laboratory evidence, and history of treatment. Through this, the sample size of 200 was deemed to be enough to perform training and evaluation of machine learning models and still ensure the feasibility constraints typical of low-resource environments.

Data Collection Procedure

The digital and paper-based clinical records that were kept in the hospitals located in Lahore were used to collect data. Upon receiving institutional consent, the pertinent variables were retrieved and transformed into a formatted dataset. Any personal details of the patients were eliminated before analysis to protect ethical standards and privacy.

Variables of the Study

- **Independent Variables**: age, gender, vital signs, lab values, comorbidities, clinical manifestations, treatment procedures.
- **Dependent Variable**: clinical outcome (i.e. recovery, complication or mortality).
- **Control Variables**: type of facility, length of stay, severity at admission.

Data Pre-Processing

The steps carried out to prepare the dataset to be analyzed by machine learning algorithm included:

Data Cleaning

The raw data were full of awful inconsistencies and typing errors and missing entries another common issue in low resource health settings. The observations with large missing values were eliminated so that there was reliability of the model.

Handling Missing Values

In the event of other variables where it has a large number of missing data, both the mean imputation (when dealing with numerical variables) and mode imputation (when dealing with categorical variables) were utilized. The method did not distort the sample of data distribution as the original sample was kept intact.

Feature Encoding

Categorical variables (gender, categories of symptoms, and diagnosis) were converted to numerical values using one-hot encoding so that they could become compatible with machine learning algorithms.

Feature Scaling

Continuous variables (age, lab values, vital signs) were also normalized using MinMax scaling in order to improve the stability and convergence of the model.

Machine Learning models applied.

The effectiveness of these three algorithms in structured clinical data was the reason why they were chosen:

Logistic Regression - baseline statistical outcome predictor.

Random Forest Classifier - ensemble model that is appropriate to deal with complex interactions between features as well as non-linear relationships.

Support Vector Machine (SVM)- useful in the high dimensional, limited sample data sets.

Model Training and Validation Modelling training and validation involve the creation of an efficient sequence model to manage uncertainties by both local and global forecasting and prediction methods. This is because, under modelling training and validation, an effective sequence model will be developed, in order to deal with uncertainties, through the use of local and global forecasting and prediction models.

An 80/20 split was used to segregate the cleaned data into training and testing. Training was performed on 80 percent of data and the subsequent evaluation on the remaining 20 percent.

To guarantee the reliability, cross-validation (5-fold) was applied, and grid search was used to earn optimal model parameters.

Evaluation Metrics

Standard clinical prediction measures were used to evaluate model performance, they included:

- Accuracy
- Precision
- Recall (Sensitivity)
- F1-Score

Area Under the ROC Curve (AUC)

Such measurements made it possible to compare the models and indicate how well they can forecast clinical outcomes.

Ethical Considerations

The paper followed the ethical standards on secondary data analysis. No patient identifiable details were employed and the respective healthcare authorities gave their consent before accessing the clinical records.

DATA ANALYSIS AND FINDINGS

In this step, the analysis process is summarized by highlighting key elements and outlining the main steps that will be undertaken.

Analysis Process Overview

In this step, the analysis process will be summarized with the highlighting of key elements and description of the main steps that will be pursued.

There were four major steps in the data analysis process:

They are (1) exploratory data analysis, (2) model development, (3) model evaluation and (4) comparative findings. All the steps were carried out in a systematic order so that the transparency and rigor in the interpretation of low-resource clinical outcome predictions were achieved. All the analyses were performed with the help of the Scikit-Learn library in Python.

The following exploratory data analysis (EDA) will be utilized to study the data.

The sample was divided into two groups: control and experimental.

Demographics

The dataset was composed of 200 patients with different gender, age, and clinical conditions. The number of years was 18 to 78 years and the mean age was 44.6 years.

Gender ratio: 54 and 46 percent respectively.

Such items as fever (67%), shortness of breath (52%), dehydration (34%), and chest pain (28) were frequent clinical symptoms.

Hypertension (31%), diabetes (26%), and anemia (19%), were regarded as the major comorbidities.

The result variable classified the patients as:

- Recovered (72%)
- Developed complications (18%)
- Mortality (10%)

Missing Data Patterns

Laboratory indicators, which included blood pressure measurements and level of hemoglobin, were seen to have minor missing values. Imputation methods were used as mentioned in the methodology. The sample of 200 was not reduced because no more than 5% of the variables were missing.

Correlation Analysis

There was a Pearson correlation matrix that indicated:

Favourable associations among age, diabetes, and hypertension, and undesirable clinical outcomes.

The highest correlations ($r > 0.60$) were observed between shortness of breath, oxygen saturation and complication/mortality outcomes.

Laboratory markers including low hemoglobin and high WBC count had moderate positive relationships with outcomes of complications as well. This assisted in alerting feature significance study during the ML phase.

Model Training and Performance

Three models were trained, which were; Logistic Regression (LR), Random Forest (RF), and Support Vector Machine (SVM). All the models had been tested on 20 percent of the data (40 patients) and cross-validated using 5 folds.

Logistic Regression

The baseline model was the Logistic Regression.

Table 1: Logistic Regression

Metric	**Score**
Accuracy	0.78
Precision	0.74
Recall	0.71
F1-Score	0.72
AUC	0.80

The ability of Logistic Regression to predict outcomes (especially recovered and non-recovered) was moderate. But it was not as good at identifying complications because it had limitations of linear separability.

Random Forest Classifier

Random Forest was better than the baseline model.

Table 2: Random Forest Classifier

Metric	Score
Accuracy	0.86
Precision	0.84
Recall	0.83
F1-Score	0.83
AUC	0.90

RF showed good predictive accuracy because it was capable of capturing non-linear associations as well as interaction of features. It particularly proved effective in identifying risk of complications and mortality.

Feature Importance (RF)

1. The best predictive characteristics were:
2. Oxygen saturation
3. Age
4. Diabetes status
5. WBC count
6. Shortness of breath
7. Blood pressure
8. Fever

The indicators are also clinically consistent which means the model is consistent with the real-life medical judgment.

Support Vector Machine (SVM)

SVM was found to be close to Random Forest.

Table 3: Support Vector Machine (SVM)

Metric	Score
Accuracy	0.84
Precision	0.82
Recall	0.80
F1-Score	0.81
AUC	0.88

SVM was very sensitive and can be used to detect at-risk patients in their early stages. It had a somewhat worse performance compared to RF yet better than LR.

Table 4: Comparison of All Models

Model	Accuracy	AUC
Logistic Regression	0.78	0.80
Random Forest	0.86	0.90
SVM	0.84	0.88

Random Forest turned out to be the most efficient model with the highest accuracy and good discriminating capacity between positive and negative outcomes.

Clinical Interpretation of Findings

Negative Outcomes Predictors

The following became high risk predictors across all the models:

- Low oxygen saturation
- Older age
- Diabetes and hypertension
- Elevated WBC count
- Shortness of breath
- Low hemoglobin (anemia)

Those results are consistent with the world clinical literature on poor health outcome determinants.

Applications to Low-Resource Settings

The models demonstrated:

- Good aptitude to aid triage decisions.
- Possible to find out high-risk patients earlier.
- Optimized capability to work with low computing requirements.
- Strong findings though with small samples and simple clinical variables.

This underscores the opportunities of ML in terms of filling the shortage of specialists in low-resource centers.

Cross-Validation Findings

Cross-validation 5: The result of cross-validation showed that:

Random Forest also had the most consistent scores by folds.

The performance of Logistic Regression was more likely to change given the sensitivity to data distribution.

SVM was highly recalling which is a desirable feature of clinical risk detection.

This indicated that the prediction patterns were not random or data set-specific.

Summary of Major Findings

- Even in small datasets that are common in low-resource settings, ML models are able to accurately make predictions concerning clinical outcomes.
- Random Forest was the best model and SVM came in second.

- Age, comorbidities, oxygen saturation, WBC count, and shortness of breath were considered to be key predictors of negative outcomes.
- To be integrated into the basic hospital workflows, the ML system can be used as early warning and decision support.
- The results indicate the transferability of the findings to larger health systems in other similar areas.

CONCLUSION

The paper has shown that machine learning models could be applicable as a predictive clinical outcome tool in a resource-constrained setting with high reliability and low cost. Random Forest is the most accurate in predictive power in the data set of 200 patient records, closely surpassed by SVM. The significant predictors of poor clinical outcomes were age, comorbidities, oxygen saturation, WBC count and respiratory symptoms.

ML applications would allow addressing the lack of specialists and contribute to prompt clinical decision-making. Nevertheless, the implementation phase demands better data systems, capacity building, ethical protection, and favorable government policies. The results emphasize that despite the lack of resources, ML can change the nature of clinical care delivery and enhance patient outcomes in underserved areas.

RECOMMENDATIONS

According to the results of this paper, a number of practical, technical, and policy-level recommendations will be offered to reinforce the application of machine learning (ML) to forecast clinical outcomes in low-resource environments.

Empowering Data Quality and Clinical Records

The improvement of the stable, formal, and low-digitized patient record systems must take priority in the low-resource healthcare facilities. Data inconsistency, values absent, and unfinished records have a significant impact on model accuracy. Basic electronic health records (EHRs) investment is able to enhance data quality, minimize errors, and predict analytics in light of standardized data entry format.

Merging ML Tools into the Daily Clinical Decision-Making

Hospitals need to implement risk prediction dashboards based on the use of ML in order to help healthcare professionals identify high-risk patients promptly. The model that has shown the best results is the Random Forest model, which can be transformed into a real-time clinical triage device, providing evidence-based warnings about possible complication or death.

Prep of Healthcare Workers: Capacity Building and Training

In order to adopt the application of ML successfully, nurses, medical officers and health technicians should be trained on the basic interpretation of data, use of models and decision-support systems. Local capacity building makes one less reliant on external experts and encourages long-term sustainability.

LCTM Low-cost, resource-efficient ML Models

High-performance computing hardware is not dependable in the low-resource environment. Thus, one should consider models like Random Forest, Logistic Regression, and SVM that are effective when working with small data sets. Without heavy models, lightweight models enable hospitals to operate predictive systems using simple computers or even mobile devices.

Digital Health Transformation Policy Support

Policies that will promote digital health innovation should be drafted by governments and health ministries. This involves financing data infrastructure, motivating hospitals to implement ML-driven frameworks, and coming up with ethical guidelines to data management and privacy of patients.

The Responsible and Ethical ML Use in Healthcare

ML tools have to be applied to an ethical framework that gives emphasis to privacy, transparency, and fairness. There should be clear rules regarding avoiding the bias of the algorithm on vulnerable groups. Audits and surveillance should be practiced regularly to determine whether models are working fairly among gender, socioeconomic status, and age groups.

REFERENCES

Adkins, D. E., & Hill, E. D. (2021). Machine learning and the social sciences: A methodological framework. *Sociological Methods & Research, 50*(3), 1020–1048. https://doi.org/10.1177/004912411877361*

Beam, A. L., & Kohane, I. S. (2018). Big data and machine learning in health care. *JAMA, 319*(13), 1317–1318. https://doi.org/10.1001/jama.2017.18391

Bzdok, D., Altman, N., & Krzywinski, M. (2018). Statistics versus machine learning. *Nature Methods, 15*(4), 233–234. https://doi.org/10.1038/nmeth.4642

Chen, J. H., & Asch, S. M. (2017). Machine learning and prediction in medicine. *New England Journal of Medicine, 376*(26), 2507–2509. https://doi.org/10.1056/NEJMp1705348

Choi, R. Y., Coyner, A. S., Kalpathy-Cramer, J., & Chiang, M. F. (2020). Introduction to machine learning, neural networks, and deep learning. *Ophthalmology, 127*(1), e19–e30. https://doi.org/10.1016/j.ophtha.2019.11.014

Esteva, A., Robicquet, A., Ramsundar, B., Kuleshov, V., DePristo, M., Chou, K., Cui, C., Corrado, G. S., Thrun, S., & Dean, J. (2019). A guide to deep learning in healthcare. *Nature Medicine, 25*(1), 24–29. https://doi.org/10.1038/s41591-018-0316-z

Johnson, A. E. W., Pollard, T. J., Mark, R. G., & Shen, L. (2016). MIMIC-III, a freely accessible critical care database. *Scientific Data, 3*, 160035. https://doi.org/10.1038/sdata.2016.35

Khalilia, M., Chakraborty, S., & Popescu, M. (2011). Predicting disease risks from highly imbalanced data using random forest. *BMC Medical Informatics and Decision Making, 11*, 51. https://doi.org/10.1186/1472-6947-11-51

Mitchell, T. M. (1997). *Machine learning.* McGraw-Hill.

Obermeyer, Z., & Emanuel, E. J. (2016). Predicting the future—Big data, machine learning, and clinical medicine. *New England Journal of Medicine*, *375*(13), 1216–1219. https://doi.org/10.1056/NEJMp1606181

Rajkomar, A., Dean, J., & Kohane, I. (2019). Machine learning in medicine. *New England Journal of Medicine, 380*, 1347–1358. https://doi.org/10.1056/NEJMra1814259

Shillan, D., Sterne, J. A. C., Chantler, T., & Gibbison, B. (2019). Use of machine learning to predict clinical deterioration in high-risk patients. *Critical Care*, *23*, 288. https://doi.org/10.1186/s13054-019-2568-y

Topol, E. (2019). High-performance medicine: The convergence of human and artificial intelligence. *Nature Medicine*, *25*, 44–56. https://doi.org/10.1038/s41591-018-0300-7

Verghese, A., Shah, N. H., & Harrington, R. A. (2018). What this computer needs is a physician: Humanism and artificial intelligence. *JAMA*, *319*(1), 19–20. https://doi.org/10.1001/jama.2017.19198

Wiens, J., Saria, S., Sendak, M., Ghassemi, M., Liu, V. X., & Jung, K. (2019). Do no harm: A roadmap for responsible machine learning for health care. *Nature Medicine, 25*(9), 1337–1340. https://doi.org/10.1038/s41591-019-0548-6

World Health Organization. (2021). *Digital health strategy 2020–2025*. WHO Press.