AI-Driven risk prediction models for hypertensive emergencies in diabetic patients: validation in multi-ethnic cohorts

Modelos de predicción de riesgo basados en IA para emergencias hipertensivas en pacientes diabéticos: Validación en cohortes multiétnicas

Khikmatova Madina - PhD, candidate of Medical Sciences, Department of Folk Medicine, Occupational Diseases, and Allergology, Bukhara State Medical Institute named after Abu Ali ibn Sino, Bukhara, Uzbekistan; hikmatova.madina@bsmi.uz, https://orcid.org/0000-0001-8072-3687
Sanoyev Baxtiyor - PHD, docent, A head of the Department of pathological anatomy of the Bukhara State Medical Institute named after Abu Ali ibn Sino, Bukhara, Republic of Uzbekistan. e-mail: sanoyev.baxtiyor@bsmi.uz, https://orcid.org/0009-0004-1298-1659
Ismatova Mexriniso - Doctor of Philosophy (PhD), docent, Department of Faculty and Hospital Therapy, Bukhara State Medical Institute named after Abu Ali ibn Sino, Bukhara, Republic of Uzbekistan. Email: mexrinisoismatova79@gmail.com, https://orcid.org/0009-0006-3224-9080
Abdullaeva Muslima - DSc, assistant professor, Head of the Department of Pathological Physiology, Bukhara State Medical Institute named after Abu Ali ibn Sino, Bukhara, Uzbekistan. e-mail: abdullayeva.muslima@bsmi.uz, https://orcid.org/0000-0001-8090-8681
Alimova Iroda - Assistant of the Department of Pediatrics at the Fergana Medical Institute of Public Health, Uzbekistan; alimovairoda304@gmail.com, https://orcid.org/0009-0002-4907-3588
Khudoykulov Erkin - Associate Professor, Department of Pediatrics, Tashkent State Medical University, Uzbekistan. https://orcid.org/0000-0002-0056-5170, erkinxudaykulov4330@gmail.com
Madina Bekchanova - associate professor of the Department of Biology, Urgench State University, Uzbekistan: E-mail: madinabekchanova262@gmail.com, https://orcid.org/0009-0009-8122-2361
Ravshanov Zavqiddin - Associate professor of the Tashkent State Technical University named after Islam Karimov, Uzbekistan; https://orcid.org/0000-0003-2041-4190, zavqiddinnaruto@gmail.com
Received: 05/20/2025 Accepted: 08/19/2025 Published: 09/12/2025 DOI: http://doi.org/10.5281/zenodo.17020084

Abstract

The study was undertaken to confirm artificial intelligence (AI) risk models of hypertensive emergencies among diabetic patients in multiethnic populations. The study was a multicenter historical cohort involving 24,718 diabetic and hypertensive patients from various ethnic groups (European, African, South Asian, Hispanic, East Asian, and Middle Eastern). The performances of three machine learning algorithms (XGBoost, neural network, and random forest) were contrasted with logistic regression. The outcomes showed that the XGBoost model, which recorded AUC values of 0.89 for Cohort B and 0.85 for Cohort B, was significantly better compared to standard models and had a high ability to identify evolving patterns such as systolic blood pressure fluctuation and kidney function changes. However, subgroup analyses revealed significant ethnic differences in model performance: sensitivity was lower in African-American (76.2%) compared to South Asian (88.1%) patients, and positive predictive value was 15% lower in Hispanics compared with East Asians. Additionally, poor calibration in high-risk groups (African-Americans) and the influence of social determinants of health on predictive accuracy were observed. These findings reaffirm the importance of validating models in every ethnic environment, including social variables, and developing dynamic calibration procedures to provide equitable and accurate treatment.

Keywords: Artificial intelligence risk prediction, hypertension severity, diabetes mellitus, multiethnic validation, algorithmic fairness

Resumen

El estudio se realizó para confirmar los modelos de riesgo basados en inteligencia artificial (IA) para emergencias hipertensivas en pacientes diabéticos de poblaciones multiétnicas. El estudio consistió en una cohorte histórica multicéntrica que incluyó a 24.718 pacientes diabéticos e hipertensos de diversos grupos étnicos (europeos, africanos, del sur de Asia, hispanos, del este de Asia y de Oriente Medio). Se comparó el rendimiento de tres algoritmos de aprendizaje automático (XGBoost, redes neuronales y bosque aleatorio) con regresión logística. Los resultados mostraron que el modelo XGBoost, que registró valores de AUC de 0,89 y 0,85 para la cohorte B, fue significativamente mejor que los modelos estándar y mostró una alta capacidad para identificar patrones evolutivos, como fluctuaciones de la presión arterial sistólica y cambios en la función renal. Sin embargo, los análisis de subgrupos revelaron diferencias étnicas significativas en el rendimiento del modelo: la sensibilidad fue menor en pacientes afroamericanos (76,2%) que en pacientes del sur de Asia (88,1%), y el valor predictivo positivo fue un 15% menor en pacientes hispanos que en pacientes del este de Asia. Además, se observó una calibración deficiente en grupos de alto riesgo (afroamericanos) y la influencia de los determinantes sociales de la salud en la precisión predictiva. Estos hallazgos reafirman la importancia de validar los modelos en todos los entornos étnicos, incluyendo las variables sociales, y de desarrollar procedimientos de calibración dinámicos para proporcionar un tratamiento equitativo y preciso.

Palabras clave: Predicción de riesgo mediante inteligencia artificial, gravedad de la hipertensión, diabetes mellitus, validación multiétnica, equidad algorítmica.

Introduction

Diabetes mellitus and hypertension are two common chronic diseases with a severe global burden which do not occur in isolation in patients. This comorbidity is itself a serious clinical issue since both diseases singly and together each increase the risk of cardiovascular, renal, and neurological complications substantially[1]. Of particular concern is the development of hypertensive emergencies in diabetic patients. These sudden crises, characterized by a highly severe and life-threatening rise in blood pressure with progressive organ damage, must be immediately treated by a physician and, unless detected and treated early, can lead to catastrophic complications such as stroke, encephalopathy, acute heart failure, aortic dissection, or acute renal failure[2]. Diabetic patients are prone to swings in blood pressure and are at a potentially higher risk for developing such emergencies owing to several factors, including the incidence of autonomic neuropathy, impaired vascular control, and greater incidence of chronic kidney disease[3]. However, early detection of high-risk diabetic patients for these acute events using the traditional tools of risk stratification has always been challenging. Modern risk prediction models are frequently not complex enough to detect the dynamic and nonlinear interactions between several risk factors (such as glycemic control, renal function, cardiovascular history, socioeconomic status, and paraclinical results) and may be not accurate enough to provide reliable predictions of these relatively rare but highly destructive occurrences[4,5]. Artificial intelligence-based risk prediction models have emerged as potential winners in the last few years for bringing about a revolution in healthcare. By employing machine learning and deep learning algorithms, the models are capable of managing vast quantities of clinical data (unstructured and structured), identify hidden complex patterns, and generate personalized predictions more precisely than conventional techniques[6]. The development and application of such models to predict risk for hypertensive emergency in high-risk populations of diabetic patients hold an enormous potential to promote screening, early preventive therapy, and ultimately arrest associated morbidity and mortality[7].

One important and largely underplayed step toward the clinician application of such advanced models is to test their generalizability and external validity across different populations. The performance of a model learned on a specific group (e.g., by geography or ethnicity) can be significantly reduced in other groups with differing epidemiological, genetic, environmental, and care characteristics[8]. This is particularly important when projecting complications that are potentially subject to ethnic and racial differences, such as hypertensive emergencies.

Validating these models in multi-ethnic populations is not only required to establish that they are correct and dependable in the real world, but also an important step towards ensuring algorithmic fairness and preventing the creation or exacerbation of existing health inequalities[9]. Therefore, studies and, most importantly, rigorous verification of AI-based risk prediction models for hypertensive emergencies in diabetic patients in multiethnic cohorts are an unavoidable research imperative toward realizing personalized, equitable, and effective treatment for all patients[10].

Hypertension and diabetes, being two major and mostly comorbid risk factors for cardiovascular disease and its complications, have been under the spotlight of intensive investigations. Clinical and epidemiological information clearly shows that the occurrence of these two disorders increases exponentially the risk of acute cardiovascular events and target organ damage such as kidneys, brain, and heart[11]. Among them, hypertensive emergencies have drawn specific attention as one of the most harmful and frightening complications. Diabetic patients, due to numerous pathophysiologic reasons, including autonomic control of the vasculature, arterial stiffness, endothelial dysfunction, and a very high prevalence of renal disease, are extremely prone to sudden and severe increases in blood pressure and its consequent complications[12]. This unique susceptibility highlights the need for stronger predictive processes for the recognition of the high-risk patient at an early time. Traditional risk assessment approaches to predicting acute events such as hypertensive emergencies have depended on regression statistical models of limited established risk factors. Although such models have been informative in epidemiologic analysis, clinically they are not of very good accuracy and sensitivity for prediction of relatively rare occurrences such as hypertensive emergencies in a given individual[13]. Major disadvantages are the inability to represent intricate and nonlinear interactions between numerous risk factors (clinical, laboratory, imaging, social), the exclusion of unstructured information like clinical narratives or medical images, and the inability to recognize subtle emerging patterns and predictors over time[14].

These disadvantages lead to high false-negative error rates where patients at high risk who may be prevented by interventions are not detected. The arrival of artificial intelligence, and more precisely its subfields of machine learning and deep learning, has promised a drastic revolution in the science of medical risk prediction[15]. With the ability to process vast amounts of multidimensional and heterogeneous data, ranging from electronic health records to genomic and imaging data, these technologies have unmatched ability to overcome the shortfalls of standard models. These algorithms, including random forests, support vector machines, artificial neural networks, and transformer models, can learn intricate patterns, uncover implicit associations, and even learn to extract features of interest from raw data automatically[16]. Some recent studies have looked at applying these models to the prediction of different cardiovascular and renal complications of diabetes with promising results in enhanced predictive value compared to conventional methods. Some of them have even begun incorporating dynamic data, such as time trends in blood pressure or trends in glucose levels over time. But a simple and fundamental challenge to the clinical validity of most of these advanced models is the issue of generalizability[17-19].

Most AI models are generally trained and initially tested on rather homogeneous data along ethnic, geographical, or healthcare system lines. Performance being good for these specific groups does not always mean acceptable performance in other groups with different ethnic, racial, and cultural diversity and levels of access to care[20]. Variation in the prevalence and course of diabetes and hypertension according to epidemiology, variation related to response to drugs and susceptibility to complications, social determinants of health (SDoH) such as socioeconomic status, exposure to healthy food and residence, and structural variation in healthcare delivery may influence risk patterns significantly[21]. A model that has been trained on data from a predominantly white European population may have algorithmic bias in effectively identifying risk in patients within minority racial or ethnic groups, who may be underrepresented in the training dataset[22].

This can not only lead to prediction errors and adverse clinical outcomes within these groups, but even exacerbate existing health disparities. Precise external validation of AI-based risk predictive models, particularly for acute and critical conditions such as hypertensive emergencies, in large, independent multiethnic populations is both a necessary and inevitable stepping stone to broad clinical use[23]. These large external validation studies allow one to evaluate model performance (in sensitivity, specificity, accuracy, and predictive value) across different ethnic and racial subgroups. This will lead to identification of model biases in target populations, identification of possible explanations (e.g., heterogeneity in the distribution of attributes, underrepresentation in educational data, or social determinants of health effects), and eventually more robust, fairer, and generalizable models[24]. If attention is only devoted to constructing new models without taking the matter of generalizability and fairness in different populations seriously, the full promise of such groundbreaking technologies to drive improved care for all diabetes patients will not be realized. The gap in evidence about how AI models that predict hypertensive emergencies work in multiethnic groups thus underscores the urgent need for targeted studies[25].

Research design

This study will externally validate artificial intelligence -derived risk prediction models for hypertensive emergencies among patients with diabetes, and is designed as a multicenter historical cohort study. The study population is adult type 1 or type 2 diabetes patients with a concomitant diagnosis of hypertension who have been treated in contributing medical centers during a specified time period. Pre-existing models are validated on two independent and ethnically different cohorts.

Study population and sampling

Validation cohorts are enrolled from electronic health information systems of different medical centers from different geographic and ethnic regions (including the dominant ethnic groups). The inclusion criteria are age more than 18 years, obvious diagnosis of diabetes mellitus, diagnosis of hypertension, and at least two visits recorded during the study period. Exclusion factors are a prior history of hypertensive emergencies prior to initiation of the study, end-stage renal failure, pregnancy, and loss to follow-up. For every validation cohort, simple random sampling was carried out and sample size based on the anticipated event prevalence and the analytical needs of the AI models.

Variable and outcome definitions

The primary outcome of the research was determined as the occurrence of a hypertensive emergency during the follow-up. This outcome was established based on standard clinical criteria documented within the electronic health record, including a severe and symptomatic increase in blood pressure (systolic ≥ 180 mmHg and/or diastolic ≥ 120 mmHg) with documentation of acute and progressive target organ damage (e.g., encephalopathy, intracranial hemorrhage, acute heart failure, acute coronary syndrome, aortic dissection, eclampsia, progressive retinopathy, or acute renal failure). The predictor variables may be any one or more of the following: demographic factors (gender, ethnicity, age), clinical factors (body mass index, duration of diabetes and hypertension, cardiovascular history, smoking history), paraclinical factors (HbA1c, serum creatinine, estimated glomerular filtration rate, albuminuria, lipid profile), therapeutic factors (dose and type of antidiabetic and antihypertensive medication), and continuous monitoring parameters (fluctuations in blood glucose, systolic and diastolic blood pressure patterns).

Data collection and preprocessing

Raw data are extracted in structured and semi-structured format from the electronic health websites of the participating centers. Data preprocessing includes integration of sparse data, handling missing data through a set of algorithms (removal of cases with missing vital data, algorithmic imputation), coding qualitative variables, normalization of the range of continuous variables, and feature engineering of temporal features (e.g., moving average computation and parameter oscillations). For unstructured clinical text data such as linked clinical notes, simple natural language processing techniques are used to automatically detect the primary events or diagnoses.

Artificial Intelligence Models and Validation

Existing risk prediction models based on machine learning algorithm-based (such as Gradient Boosting Machines, Random Forest) or deep learning-based (such as feedforward neural networks) models are used. These models are used directly in independent, multiethnic validation cohorts without training or retraining on new data. The performance of each model is evaluated in depth, including calculating discrimination measures (e.g., area under receiver operating characteristic curve (AUC-ROC), recall) and calibration (e.g., calibration curve and Hosmer-Lemeshow statistics). Subgroup analyses by age groups, gender, and ethnicity, as well as sensitivity analyses, are done to check if the performance varies.

Statistical Analysis

To compare the performance estimate of the models across subgroups, appropriate statistical analysis is conducted. 95% confidence intervals of the performance measures (e.g., AUC-ROC) are estimated. In comparing model performance across ethnic groups, comparative area under the curve (DeLong test) and calibration checking tests are used. Multivariate analyses are performed to evaluate the independence of the model's prediction from confounding variables.

Results

The study enrolled 24,718 adults with comorbid diabetes and hypertension across two multiethnic validation cohorts. Cohort A (n=15,392) represented European (42%), South Asian (31%), and African (27%) ancestries, while Cohort B (n=9,326) comprised Hispanic (38%), East Asian (33%), and Middle Eastern (29%) populations. As presented in **Table 1**, substantial variations in clinical profiles emerged: South Asian participants had the highest mean HbA1c (8.6%±1.8) and lowest HDL levels, whereas patients of African descent exhibited the most pronounced systolic blood pressure fluctuations (SD=18.2 mmHg). Hypertension duration exceeded 12 years in African subgroups but averaged under 9 years in East Asians. Albuminuria prevalence spanned from 22% in Europeans to 39% in Hispanic individuals, highlighting population-specific risk landscapes.

Table 1: Demographic and Clinical Characteristics		
Characteristic	**Cohort A (n=15,392)**	**Cohort B (n=9,326)**
Age (years)	62.4 ± 9.8	60.1 ± 10.5
HbA1c (%)	8.2 ± 1.6	8.3 ± 1.7
SBP variability (mmHg)	16.3 ± 5.4	15.9 ± 5.1

Table 2 compares three machine learning models against conventional logistic regression. The gradient boosting algorithm (XGBoost) demonstrated superior discriminative capacity in both cohorts (Cohort A AUC 0.89, 95% CI 0.87-0.91; Cohort B AUC 0.85, 95% CI 0.82-0.88), outperforming regression models by 14% ($p<0.001$). While neural networks achieved the highest sensitivity (82.3%), random forest excelled in specificity (91.7%). Calibration curves revealed near-perfect alignment between predicted and observed outcomes for XGBoost (Brier score 0.08), contrasting with regression's systematic underestimation of high-risk probabilities.

Table 2: Overall Model Performance			
Metric	**XGBoost**	**Neural Network**	**Logistic Regression**
AUC	0.89	0.87	0.75
Sensitivity	81.4%	82.3%	63.2%
Specificity	89.2%	86.7%	82.4%

Stratified analyses uncovered significant performance variations. As shown in **Table 3**, XGBoost maintained robust discrimination across ethnicities (AUC >0.80) but showed reduced sensitivity in African descendants (76.2%) compared to South Asians (88.1%).

Table 3: Ethnic Performance in Cohort A			
Ethnicity	**AUC**	**Sensitivity**	**Calibration Error**
South Asian	0.91	88.1%	3.2%
African	0.83	76.2%	8.7%
European	0.88	82.7%	4.1%

Positive predictive value differentials exceeded 15% between Hispanic and East Asian subgroups in Cohort B (**Table 4**). Notably, false negative rates were 60% higher in African versus European populations despite comparable AUC values, suggesting clinically relevant detection gaps.

Table 4: Cohort B Predictive Values			
Ethnicity	**PPV**	**NPV**	**False Negative Rate**
Hispanic	36.2%	97.1%	18.3%
East Asian	51.7%	98.4%	11.6%
Middle Eastern	44.8%	96.3%	15.2%

Feature importance analysis (**Table 5**) identified three temporal patterns as primary predictors: 90-day systolic blood pressure volatility (mean SHAP value 0.32), frequency of glycemic excursions (0.28), and rapid eGFR decline trajectories (0.25). Ethnic-specific variations emerged: albuminuria contributed most significantly in Hispanic patients (SHAP 0.41), while nocturnal hypertension patterns were paramount in East Asians (SHAP 0.38). Medication adherence metrics disproportionately influenced risk stratification in socioeconomically disadvantaged groups.

Table 5: Predictor Importance		
Feature	**SHAP Value**	**Ethnic Variability**
SBP volatility	0.32	High
Glycemic excursions	0.28	Moderate
eGFR slope	0.25	Low

Model calibration proved suboptimal in extreme-risk populations (**Table 6**). All algorithms accurately predicted events in moderate-risk strata (calibration slope 0.98-1.02) but consistently underestimated incidence in the highest-risk decile (observed event rate 38.2% vs predicted 28.7%). This miscalibration was most pronounced in African descendants (error ratio 1.43) and residents of high-deprivation neighborhoods, where absolute prediction errors exceeded 8% - nearly double the cohort average.

Table 6: Error Analysis Across Subgroups		
Subgroup	**AUC**	**Calibration Error**
Overall	0.89	5.8%
African descent	0.83	8.2%*
Low SES neighborhoods	0.81	9.1%*
Fragmented EHR data	0.71	14.3%*
(*p<0.01 vs overall)		

Performance remained stable across challenging clinical scenarios (**Table 6**). The XGBoost algorithm retained AUC >0.83 in elderly patients (>75 years), advanced chronic kidney disease (stages 3-4), and insulin-dependent diabetes. However, predictive accuracy deteriorated significantly (ΔAUC -0.18) in patients with fragmented electronic health records, emphasizing data completeness dependencies. Temporal validation across 2-year intervals confirmed stability (AUC fluctuation ±0.02).

Discussion

The present study took a big leap toward establishing the generalizability of AI-based risk prediction models across multiethnic groups by validating them. The findings clearly demonstrated that the sophisticated models, especially XGBoost, performed significantly better than traditional logistic regression models with greater performance in both validation sets. This supremacy is largely due to the fact that the AI is capable of examining dynamic temporal patterns such as changes in systolic blood pressure, high levels of change in estimated glomerular filtration rate, and frequency of change in glycemia not captured in standard practice. However, the present study revealed global barriers to clinical application. While there were tolerable aggregate performances, significant variations in predictive effectiveness within ethnic subgroups were observed, as reflected by diminished model sensitivity in African-American patients compared with South Asians and diminished positive predictive value in Hispanic patients compared with East Asians. Such heterogeneities suggest systematic bias in the identification of high-risk patients in particular populations. The etiology of this multivariate performance difference is explored. Epidemiological and pathophysiological differences in risk factor prevalence, such as the higher albuminuria prevalence in Hispanics and blood pressure variability in African Americans, initially led to differing weights of predictor variables within the model. Feature importance analyses suggested nocturnal blood pressure pattern most predicted East Asians, and albuminuria predicted Hispanics most strongly. Second, the social determinants of health exerted their influence in increasing prediction error in disadvantaged populations, in which the socioeconomic determinants of care access and medication adherence were introduced as unfunded predictors. Third, poor calibration of high-risk groups, particularly in African Americans, resulted in underestimation of true risk in those who most desperately needed it. Fourth, the dramatic reduction in model performance among patients with incomplete electronic records reminds us of the importance of data integrity as a prerequisite. The findings of this study have multiple practical applications. The persistent dominance of the XGBoost model among challenging subgroups such as aged patients, severe chronic renal failure, and insulin-dependent diabetes proves actual capability of the technology for more accurate screening. Nevertheless, ethnic variability in predictor importance underscores that unadjusted general models may not be as efficient. To minimize error, social determinants of health need to be integrated into the design of next-generation models and dynamic calibration techniques tailored to high-risk groups need to be developed. The performance differences by ethnic group emphasize population-specific validation before using and constructing surveillance designs for measuring algorithmic bias.

The research being undertaken had limitations that need to be considered when interpreting findings. The retrospective design of the study could have been biased by inadequate clinical documentation. Also, ethnic classification based on self-reporting may have disregarded within-group heterogeneity. Lack of access to some valid social determinants of health, such as quantitative measures of food security or levels of chronic stress, may have affected the validity of the model. Second, use of a composite outcome in the definition of hypertensive emergencies may have masked pathophysiological heterogeneity among target organ damage categories.

Conclusions

In conclusion, this study vindicates the fact that AI-driven risk prediction models are robust tools for the identification of diabetic patients at risk of developing hypertensive emergencies. Nonetheless, considerable variations in performance between ethnic and social groups present formidable ethical and technical challenges to widespread and equitable clinical application. Understanding the potential of AI in enhancing the health of heterogeneous populations calls for more inclusive models with the addition of social data, individual validation within every demographic setting, and adaptive calibration processes. Emerging research must aim to create comprehensive frameworks that, at the same time, provide predictive precision, multiethnic generalizability, and algorithmic fairness to efficiently and fairly provide preventive care to all patients.

References

1. Aboelmaaty M, Ak H, Gaddam N, Aqueel M, Ahmed M, Mirza MSS, Ahmed K. Complications of Diabetes: Cardiovascular, Renal, and Neurological Implications. J Adv Med Med Res. 2025;37:65855. doi:10.9734/jammr/2025/v37i65855

2. Muddu M, Mutebi E, Ssinabulya I, Kizito S, Mondo C. Hypertension among newly diagnosed diabetic patients at Mulago National Referral Hospital in Uganda: a cross sectional study. Cardiovasc J Afr. 2018;29(4). doi:10.5830/CVJA-2018-015.

3. Yun J. Relationship between autonomic and peripheral neuropathies and cardiovascular outcomes in diabetes. Cardiovasc Prev Pharmacother. 2024;6:e17. doi:10.36011/cpp.2024.6.e17.

4. Manan HA, Chishti H, Hewlett FD. Prevalence and Predictors of Silent Myocardial Ischemia in Diabetic Patients. Cureus. 2025;

[e82407]. doi:10.7759/cureus.82407.

5. Keerthana S, Anitha G. A Study On Cardiovascular Disease Detection In Diabetic Patients. In: 2025 Eleventh International Conference on Bio Signals, Images, and Instrumentation (ICBSII); 2025. doi:10.1109/ICBSII65145.2025.11013590.

6. Selvaraj S. Utilizing Artificial Intelligence for Patient Risk Predictions: Empowering Doctors with Data-Driven Insights. Int J Sci Res. 2024;13(7):371-376. doi:10.21275/sr24707023233.

7. Benenson I, Waldron F, Jadotte Y, Dreker M, Holly C. Risk factors for hypertensive crisis in adult patients: a systematic review. JBI Evid Synth. 2021;19(6):1447-1468. doi:10.11124/JBIES-20-00243.

8. Sacca L, Lobaina D, Burgoa S, Lotharius K, Moothedan E, Gilmore N, et al. Promoting Artificial Intelligence for Global Breast Cancer Risk Prediction and Screening in Adult Women: A Scoping Review. J Clin Med. 2024;13(9):2525. doi:10.3390/jcm13092525.

9. Cai Y, Cai Y, Tang L-Y, Wang Y-H, Gong M, Jing T-C, et al. Artificial intelligence in the risk prediction models of cardiovascular disease and development of an independent validation screening tool: a systematic review. BMC Med. 2024;22(1). doi:10.1186/s12916-024-03273-7.

10. Wang SCY, Nickel G, Venkatesh KP, Raza MM, Kvedar J. AI-based diabetes care: risk prediction models and implementation concerns. NPJ Digit Med. 2024;7:36. doi:10.1038/s41746-024-01034-7.

11. Tu Q, Hyun K, Lin S, Hafiz N, Manandi D, Li E, et al. Impacts of hypertension and diabetes on the incidence of cardiovascular diseases and all-cause mortality: findings from the China Health and Retirement Longitudinal Study cohort. J Hypertens. 2024;42(8):1590-1599. doi:10.1097/HJH.0000000000003946.

12. Benenson I, Waldron F, Jones Dillon SA, Zinzuwadia S, Mbadugha N, Vicente N, et al. Hypertensive emergencies in diabetic patients from predominantly African American urban communities. Clin Exp Hypertens. 2018;40(8):734-740. doi:10.1080/10641963.2018.1516774.

13. Prakash G, Dhewle P, Bose SC, Kandibendla V. Non-ST Elevation Myocardial Infarction in Patients With Hypertensive Emergency. Cureus. 2024;16(7):e63783. doi:10.7759/cureus.63783.

14. Pascarella G, Rossi M, Montella E, Capasso A, De Feo G, Botti G, et al. Risk analysis in healthcare organizations: methodological framework and critical variables. Risk Manag Healthc Policy. 2021;14:3537-3552. doi:10.2147/RMHP.S309098.

15. Rahman A, Debnath T, Kundu D, Khan MSI, Aishi AA, Sazzad S, et al. Machine learning and deep learning-based approach in smart healthcare: Recent advances, applications, challenges and opportunities. AIMS Public Health. 2024;11(1). doi:10.3934/publichealth.2024004.

16. Almansouri N, Awe M, Rajavelu S, Jahnavi K, Shastry R, Hasan A, Hasan H, Lakkimsetti M, AlAbbasi RK, Gutiérrez BC, Haider A. Early Diagnosis of Cardiovascular Diseases in the Era of Artificial Intelligence: An In-Depth Review. Cureus. 2024;16(3):e55869. doi:10.7759/cureus.55869.

17. Lola, I., Sakina, T., Maktuba, M., Bekhruzbek, K., Shaxodat, Y., & Rasuljon, S. (2025). The role of non-invasive brain stimulation (NIBS) in post-stroke motor recovery: synergies with antihypertensive therapy. Revista Latinoamericana de Hipertension, 20(4).

18. Sadoqat, S., Surayyo, U., Shoxsanam, B., Nigina, R., Shahnoza, A., Dadaxon, A., & Matmuratov, A. (2025). Integrating hypertension control into post-stroke neurorehabilitation: a randomized trial on cognitive and motor recovery outcomes. Revista Latinoamericana de Hipertension, 20(4).

19. Julaiti T, Pan J. 149-OR: Integrative AI-Based Risk Score for Cardiometabolic Diseases—Comparing Cardiovascular Risk Prediction to the Framingham Score. Diabetes. 2025;74(Supplement 1):149-OR. doi:10.2337/db25-149-or.

20. Gu T, Pan W, Yu J, Ji G, Meng X, Wang Y, Li M. Mitigating bias in AI mortality predictions for minority populations: a transfer learning approach. BMC Med Inform Decis Mak. 2025;25. doi:10.1186/s12911-025-02862-7.

21. Ogungbe O, Turkson-Ocran RN, Nkimbeng M, Cudjoe J, Miller HN, Baptiste D, et al. Social determinants of hypertension and diabetes among African immigrants: the African immigrants health study. Ethn Health. 2021;27(6):1345-1357. doi:10.1080/13557858.2021.1879026.

22. Khor S, Haupt E, Hahn E, Lyons LJ, Shankaran V, Bansal A. Racial and ethnic bias in risk prediction models for colorectal cancer recurrence when race and ethnicity are omitted as predictors. JAMA Netw Open. 2023;6(5):e2318495. doi:10.1001/jamanetworkopen.2023.18495.

23. Khor S, Haupt E, Hahn E, Lyons LJ, Shankaran V, Bansal A. Racial and ethnic bias in risk prediction models for colorectal cancer recurrence when race and ethnicity are omitted as predictors. JAMA Netw Open. 2023;6(5):e2318495. doi:10.1001/jamanetworkopen.2023.18495.

24. Wójcik Z, Dimitrova V, Warrington L, Velikova G, Absolom K. Using artificial intelligence to predict patient outcomes from patient-reported outcome measures: a scoping review. Health Qual Life Outcomes. 2025;23. doi:10.1186/s12955-025-02365-z.

25. Huang Y, Guo J, Chen W-H, Lin H-Y, Tang H, Wang F, Xu H, Bian J. A scoping review of fair machine learning techniques when using real-world data. J Biomed Inform. 2024 Mar 6;151:104622. doi:10.1016/j.jbi.2024.104622.