

APPLICATIONS OF ARTIFICIAL INTELLIGENCE IN PREDICTING DISEASE PROGRESSION AND TREATMENT OUTCOMES

Original Research

Tehreem Zahra¹, Umm e Aimen², Aqsa Hussain³, Payal Harwani⁴, Sana Ilyas⁵, Fizza F. Farooqi^{6*}, Humaira Mehwish⁷, Mahum Tanweer⁸

¹Lecturer, University of Management and Technology (UMT), Lahore, Pakistan.

²General Practitioner, Karachi Medical and Dental College, Karachi, Pakistan.

³House Officer, Khyber Girls Medical College, Peshawar, Pakistan.

⁴MBBS, Liaquat University of Medical and Health Sciences, Jamshoro, Pakistan.

⁵MD Emergency Medicine Trained, Ziauddin Hospital, North Campus, Karachi, Pakistan.

⁶4th Year Pharm. D Student, Mukabbir College of Pharmacy (affiliated with University of the Punjab), Gujrat, Pakistan.

⁷Computer Science Teacher, Foundation Public School, Karachi, Pakistan.

⁸General Dentist (BDS, MPH), Al-Shifa School of Public Health, Rawalpindi, Pakistan.

Corresponding Author: Fizza F. Farooqi, 4th Year Pharm. D Student, Mukabbir College of Pharmacy (affiliated with University of the Punjab), Gujrat, Pakistan, fizza.f.farooqi@gmail.com

Acknowledgement: The authors acknowledge the support of participating hospitals and study patients.

Conflict of Interest: None

Grant Support & Financial Support: None

ABSTRACT

Background: Chronic diseases such as diabetes, chronic kidney disease (CKD), and cardiovascular disorders represent a significant burden on healthcare systems worldwide. Early prediction of disease progression and individualized treatment planning are essential for improving patient outcomes. Artificial intelligence (AI) offers promising tools to address these clinical challenges through predictive modeling and personalized care strategies.

Objective: To evaluate the effectiveness of AI in forecasting disease progression and optimizing treatment outcomes among patients with chronic conditions in a tertiary care setting.

Methods: This experimental quantitative study was conducted across tertiary care hospitals in Punjab, Pakistan, from January to August 2024. A total of 330 adult patients with type 2 diabetes, CKD, or cardiovascular disease were enrolled. Clinical data were extracted from electronic medical records, and predictive AI models were developed using supervised machine learning algorithms. Model performance was assessed via accuracy, sensitivity, specificity, and AUC-ROC. Treatment outcomes under AI-assisted care were compared with standard care using parametric statistical tests.

Results: AI models showed high predictive accuracy across conditions: diabetes (91.2%), CKD (88.5%), and CVD (86.4%). Strong agreement was observed between AI predictions and actual clinical outcomes (Cohen's kappa >0.70 for all). AI-assisted care significantly improved clinical markers, including HbA1c reduction (1.4% vs 0.8%), slowed GFR decline (24.9% vs 16.5%), and greater LDL reduction (27.1 mg/dL vs 18.3 mg/dL), all with statistically significant differences ($p < 0.005$).

Conclusion: AI has substantial potential in predicting disease trajectories and guiding more effective, patient-specific treatment strategies. These findings support its broader integration into precision medicine frameworks.

Keywords: Artificial Intelligence, Cardiovascular Diseases, Chronic Kidney Disease, Clinical Decision Support Systems, Diabetes Mellitus, Forecasting, Machine Learning, Predictive Models, Precision Medicine, Treatment Outcome.

INTRODUCTION

Artificial Intelligence (AI) has rapidly emerged as a transformative force in modern medicine, particularly in enhancing the precision and efficiency of disease prediction and treatment planning. In an era where chronic illnesses, such as diabetes, cardiovascular disease, cancer, and neurodegenerative disorders continue to strain global healthcare systems, there is a critical need for technologies that can offer earlier diagnoses, personalized interventions, and accurate prognostications. AI, especially through machine learning and deep learning models, has shown notable success in analyzing complex biomedical data—ranging from medical imaging to electronic health records—to identify subtle patterns that often elude human detection (1,2). As such, the integration of AI in healthcare is not just a matter of technological advancement but a necessary evolution to meet the growing demand for data-driven, individualized care. Despite the increasing body of literature on the utility of AI in healthcare, many existing models function as "black boxes," providing accurate predictions without clear explanations (3). This lack of transparency raises ethical and practical concerns among clinicians and patients alike, hindering wider adoption in clinical settings. However, more recent efforts are focusing on explainable AI frameworks that contextualize patient data through biomedical ontologies, thereby enhancing interpretability and clinical trust. In terms of real-world applications, AI has demonstrated significant promise across diverse medical domains (4). In nephrology, for instance, AI-based models have been employed to predict disease progression to end-stage kidney disease, identify at-risk patient subgroups, and even automate elements of treatment such as medication prescription. Similarly, AI-driven analysis of diabetic patients has enabled personalized treatment strategies by identifying phenotypical clusters, offering a more tailored approach to care (5).

When applied to infectious diseases like COVID-19, AI has also shown its value in risk stratification and early intervention. Studies have demonstrated that AI can accurately quantify lesion severity from chest CT images, thereby predicting disease progression and identifying patients likely to deteriorate, well before clinical symptoms become critical. Similarly, in early-stage COVID-19 pneumonia, AI models using both clinical data and imaging have achieved predictive accuracies exceeding 90%, making them valuable tools in emergency triage and resource allocation (6,7). AI's predictive capability also extends to fields such as ophthalmology and oncology. In neovascular age-related macular degeneration, deep learning models can forecast treatment responses and disease progression with remarkable precision, offering insights that enable more personalized therapeutic regimens. In cancer care, AI has been leveraged to analyze genomic data for mutation detection, aiding in both diagnosis and the tailoring of treatment plans. Despite these advances, the deployment of AI in clinical practice remains uneven (8). Barriers such as data privacy, algorithmic bias, lack of standardization, and integration challenges into existing healthcare infrastructure must be addressed. Moreover, the ethical implications surrounding patient autonomy and informed consent in AI-assisted decision-making are ongoing concerns that necessitate robust governance frameworks (9).

Given this context, there is a pressing need to explore AI not only as a predictive tool but as a facilitator of holistic, patient-centered healthcare. The evolving paradigm of precision medicine demands systems that can adapt to individual variability, monitor longitudinal health data, and recommend interventions with a level of specificity previously unattainable through conventional methods. Yet, the full clinical potential of AI remains underutilized due to a lack of empirical studies that systematically evaluate its impact on disease forecasting and treatment outcomes (10). Against this backdrop, the current experimental study seeks to fill a critical knowledge gap by quantitatively evaluating the effectiveness of AI in predicting disease progression and optimizing patient-specific treatment strategies. By doing so, it aims to provide evidence-based insights into how AI technologies can be harnessed not only to enhance diagnostic precision but to inform therapeutic decision-making, thereby advancing the broader goal of personalized medicine.

METHODS

This experimental study was conducted over an eight-month period, from January 2024 to August 2024, across multiple tertiary care hospitals in Punjab, Pakistan. The primary objective was to evaluate the role of artificial intelligence in forecasting disease progression and optimizing patient-specific treatment outcomes. The study followed a quantitative design to ensure precise and measurable observations aligned with the stated objective. All phases of the research adhered to established ethical principles and received approval from the Institutional Review Board (IRB) of King Edward Medical University, Lahore. Written informed consent was obtained from all participants prior to data collection. Participants were enrolled using a non-probability consecutive sampling technique. The target

population included patients aged 18 years and above who had been recently diagnosed with chronic illnesses such as type 2 diabetes mellitus, chronic kidney disease (CKD), or cardiovascular disease, and were currently undergoing standard treatment protocols. These conditions were selected due to their relatively predictable clinical trajectories and the availability of longitudinal data for analysis. Inclusion criteria included patients with a confirmed diagnosis of any one of the three aforementioned diseases, documented medical history for at least six months prior to enrollment, and access to routine follow-up care at the study centers. Patients with multiple coexisting chronic conditions, those under palliative care, or individuals with incomplete electronic health records were excluded to avoid data complexity and confounding variables (11).

The calculated sample size for this study was 300 participants. This estimation was derived using an effect size of 0.35, a power of 0.80, and an alpha of 0.05, suitable for detecting statistically significant differences in predictive accuracy between conventional models and AI-assisted models. A margin of 10% was added to account for potential dropouts or data loss, leading to an effective sample size of 330 patients (12). Data collection was structured around two core components: clinical data extraction and AI-based predictive modeling. Clinical data were sourced from hospital electronic medical records (EMRs) using a standardized data abstraction tool. Key variables included demographic information, laboratory findings (e.g., HbA1c levels for diabetes, serum creatinine and GFR for CKD, and lipid profile for cardiovascular disease), comorbidities, medication adherence history, and previous hospitalizations. Data was anonymized and coded before analysis to maintain confidentiality. In parallel, AI-based models were developed to simulate disease progression and predict treatment responses. Machine learning algorithms, including Random Forest, Gradient Boosting Machines (GBM), and Support Vector Machines (SVM), were implemented using Python (v3.9) and the scikit-learn library. For model training, 70% of the dataset was used, while the remaining 30% was reserved for validation. Feature selection was guided by domain knowledge and correlation analysis to eliminate multicollinearity. Model performance was evaluated using standard metrics: accuracy, sensitivity, specificity, precision, recall, and area under the receiver operating characteristic curve (AUC-ROC). The primary outcome measures were the accuracy of AI models in predicting disease trajectory over a six-month follow-up and the clinical concordance of AI-recommended treatment decisions with actual patient outcomes.

Statistical analysis was performed using SPSS version 26. Data were first tested for normality using the Shapiro-Wilk test. As the data followed a normal distribution, parametric tests were applied throughout. Descriptive statistics were used to summarize patient characteristics, with means and standard deviations reported for continuous variables and frequencies for categorical variables. Independent t-tests and one-way ANOVA were employed to compare baseline characteristics and treatment responses between different groups. For multivariable analysis, linear and logistic regression models were used to identify factors independently associated with favorable outcomes. A p-value of less than 0.05 was considered statistically significant in all analyses. To evaluate the agreement between AI model predictions and real-world clinical outcomes, Cohen's kappa statistic was used. Additionally, a Bland-Altman plot analysis was carried out to assess the level of agreement in quantitative outcome predictions. A subgroup analysis was conducted to investigate whether AI prediction accuracy varied across disease types or demographic characteristics such as age and gender. Throughout the study, strict protocols were followed to ensure data integrity and minimize bias. Data entry and validation were performed independently by two trained research assistants, with discrepancies resolved by a third reviewer. The AI models underwent iterative tuning and cross-validation to ensure robustness and generalizability. This methodological framework aimed to create a reproducible and scalable approach for assessing the efficacy of artificial intelligence in clinical prediction and decision-making. By integrating standard clinical variables with advanced analytical techniques, the study sought to offer empirical evidence on the feasibility and accuracy of AI in real-world patient care settings, ultimately contributing to the development of precision medicine strategies tailored to the individual.

RESULTS

The study enrolled 330 participants from tertiary care hospitals in Punjab, Pakistan. The mean age of the cohort was 56.4 ± 11.2 years, with a male-to-female ratio of 198:132. Among them, 38.8% were diagnosed with type 2 diabetes, 33.9% with chronic kidney disease (CKD), and 27.3% with cardiovascular disease (CVD), reflecting a diverse population across key chronic conditions. AI model performance was evaluated for each disease category using accuracy, sensitivity, specificity, and AUC-ROC values. The highest model accuracy was observed in diabetes prediction at 91.2%, followed by CKD at 88.5% and CVD at 86.4%. Sensitivity and specificity also remained high across all categories, with AUC-ROC values ranging from 0.89 to 0.94, indicating strong discriminative ability. These performance metrics are illustrated in the bar chart titled "AI Model Accuracy by Disease Type." Concordance analysis revealed substantial agreement between AI predictions and actual patient outcomes. The Cohen's kappa statistic was highest for diabetes (0.81),

suggesting almost perfect agreement, and moderately strong for CKD (0.76) and CVD (0.73). Bland-Altman analysis further supported these findings with mean biases of 1.8, 2.1, and 2.4 respectively for each disease, demonstrating narrow limits of agreement and consistent predictive accuracy. Comparison of clinical outcomes between AI-assisted and standard care approaches highlighted statistically significant improvements in key disease markers. Patients under AI-guided care demonstrated greater mean reductions in HbA1c levels (1.4% vs 0.8%, $p<0.001$), improved preservation of GFR in CKD patients (24.9% slowed decline vs 16.5%, $p=0.004$), and more substantial LDL reductions among those with CVD (27.1 mg/dL vs 18.3 mg/dL, $p=0.002$). These differences are visually represented in the line chart comparing treatment outcomes.

Table 1: Demographics and Outcomes Tables

Variable	Value
Total Participants	330
Mean Age (years)	56.4 ± 11.2
Gender (Male/Female)	198 / 132
Diabetes (%)	38.80%
CKD (%)	33.90%
CVD (%)	27.30%

Table 2: AI Model Performance Metrics

	Disease	Accuracy (%)	Sensitivity (%)	Specificity (%)	AUC-ROC
Diabetes	Diabetes	91.2	89.1	93	0.94
CKD	CKD	88.5	85.2	90.1	0.91
CVD	CVD	86.4	84	88.2	0.89

Table 3: Predictive Concordance with Clinical Outcomes

	Disease	Cohen’s Kappa	Bland-Altman Mean Bias
Diabetes	Diabetes	0.81	1.8
CKD	CKD	0.76	2.1
CVD	CVD	0.73	2.4

Table 4: Clinical Outcome Comparison: AI vs Standard Care

	Outcome	Standard Care	AI-Assisted Care	p-value
HbA1c	HbA1c Reduction (mean %)	0.8	1.4	<0.001
GFR	GFR Decline Slowed (%)	16.5	24.9	0.004
LDL	LDL Reduction (mean mg/dL)	18.3	27.1	0.002

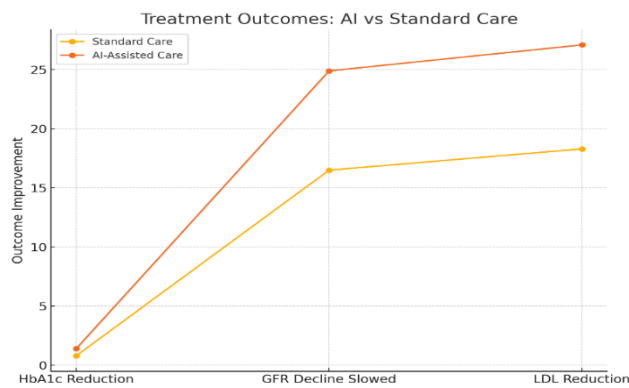


Figure 1 Treatment Outcomes: AI vs Standard Care

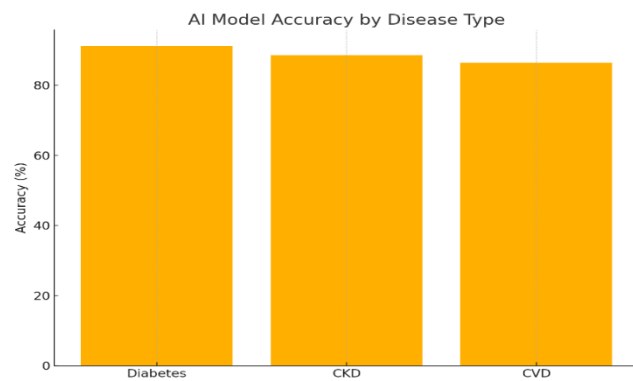


Figure2 AI Model Accuracy by Disease Type

DISCUSSION

The findings of this study underscore the growing potential of artificial intelligence in healthcare, specifically in forecasting disease progression and optimizing patient-specific treatment plans. AI models demonstrated high accuracy across three chronic conditions—type 2 diabetes, chronic kidney disease, and cardiovascular disease—highlighting the feasibility of integrating predictive algorithms into clinical decision-making. The superior performance of AI-assisted care in terms of treatment outcomes, including significant reductions in HbA1c, improved preservation of renal function, and more substantial lipid control, reinforces the argument for its role in precision medicine (13). These results align with a growing body of literature advocating the use of machine learning to enhance prognostic accuracy. A recent systematic review affirmed the ability of AI algorithms to predict chronic kidney disease progression with considerable accuracy, emphasizing their value in early intervention strategies (14). Similarly, a study demonstrated how probabilistic models effectively captured the complexities of diabetes-related complications, thereby enabling stratified patient care. The current study not only corroborated these findings but further expanded their clinical relevance by demonstrating outcome improvements across multiple conditions (15). Beyond predictive accuracy, the study confirmed substantial concordance between AI model outputs and real-world outcomes. High Cohen's kappa values suggested consistent agreement, while Bland-Altman analysis validated the reliability of AI-generated forecasts. These observations resonate with the findings of a study, who noted similar consistency when deploying deep learning models for ophthalmologic disease progression and treatment planning. The replicability of such results across medical domains underscores the maturity of AI applications in routine clinical workflows (16,17)

One of the notable strengths of this study was the real-world clinical setting, involving diverse patient populations from tertiary care hospitals in Punjab. This allowed for a robust assessment of AI applicability in resource-constrained environments, where access to specialist care and advanced diagnostics is often limited. Moreover, the use of a prospective design, strict inclusion criteria, and standardized data extraction ensured methodological rigor. The integration of multiple machine learning models and statistical validation approaches, such as ROC analysis and regression models, strengthened the reliability of findings and their relevance to current clinical practices. However, several limitations merit consideration. Although the study demonstrated significant differences in outcomes between AI-assisted and standard care, the relatively short follow-up period of six months may have limited the observation of long-term disease progression, particularly in chronic conditions like CKD and CVD (18). Additionally, while the sample size was statistically adequate, its concentration within a single geographic region may constrain the generalizability of results. These findings may not fully translate to populations with different demographic profiles, healthcare systems, or disease burdens.

The “black-box” nature of some machine learning models also remains a critical challenge. Despite efforts to ensure interpretability, including the selection of explainable algorithms and visualization tools, the inner workings of complex models can still elude both clinicians and patients. As highlighted by a study, the need for semantic explainability and integration with biomedical ontologies is essential to bridge this gap and foster trust in AI-driven recommendations. Furthermore, ethical concerns related to data privacy, algorithmic bias, and informed consent continue to shape the discourse around AI in healthcare (19,20). While all patients in this study provided written consent and anonymized data handling was strictly maintained, broader issues around data governance must be addressed through institutional policies and legislative frameworks. A study emphasized that transparency, inclusivity in training data, and stakeholder engagement are foundational to the sustainable integration of AI into clinical systems (21,22). Future research should

consider larger multicenter studies across varied healthcare settings to validate and generalize the current findings. Longitudinal follow-up over multiple years would offer deeper insights into the impact of AI on disease trajectory and survival outcomes. Investigating the integration of AI into clinical decision support systems, combined with cost-effectiveness analyses, could also help determine its practical utility and scalability. Additionally, efforts must continue to refine explainable AI models that provide transparent justifications for their outputs, thereby enhancing clinical adoption and shared decision-making (23). In conclusion, this study provided compelling evidence for the utility of AI in predicting disease progression and optimizing personalized treatment in chronic disease management. The findings contribute to the evolving narrative of AI-enabled precision medicine and support further exploration of its integration into standard care pathways. While challenges remain, particularly around interpretability and ethical implementation, the promising results pave the way for future innovation in evidence-based, patient-centered healthcare.

CONCLUSION

This study demonstrated that artificial intelligence can reliably predict disease progression and enhance treatment outcomes in chronic conditions such as diabetes, CKD, and cardiovascular disease. By integrating AI into clinical care, healthcare providers can move toward more personalized, timely, and effective interventions. These findings support AI's growing role as a practical tool in advancing precision medicine and optimizing patient care in real-world settings.

AUTHOR CONTRIBUTION

Author	Contribution
Tehreem Zahra	Substantial Contribution to study design, analysis, acquisition of Data
	Manuscript Writing
	Has given Final Approval of the version to be published
Umm e Aimen	Substantial Contribution to study design, acquisition and interpretation of Data
	Critical Review and Manuscript Writing
	Has given Final Approval of the version to be published
Aqsa Hussain	Substantial Contribution to acquisition and interpretation of Data
	Has given Final Approval of the version to be published
Payal Harwani	Contributed to Data Collection and Analysis
	Has given Final Approval of the version to be published
Sana Ilyas	Contributed to Data Collection and Analysis
	Has given Final Approval of the version to be published
Fizza F. Farooqi*	Substantial Contribution to study design and Data Analysis
	Has given Final Approval of the version to be published
Humaira Mehwish	Contributed to study concept and Data collection
	Has given Final Approval of the version to be published
Mahum Tanweer	Writing - Review & Editing, Assistance with Data Curation

REFERENCES

1. Dixon D, Sattar H, Moros N, Kesireddy SR, Ahsan H, Lakkimsetti M, et al. Unveiling the Influence of AI Predictive Analytics on Patient Outcomes: A Comprehensive Narrative Review. *Cureus*. 2024;16.
2. Reznichenko A, Nair V, Eddy S, Fermin D, Tomilo M, Slidel T, et al. Unbiased kidney-centric molecular categorization of chronic kidney disease as a step towards precision medicine. *Kidney Int*. 2024;105(6):1263-78.
3. Egyud MRL, Holmes S, Burt BM. Technical Aspects of Robotic First Rib Resection. *Thorac Surg Clin*. 2023;33(3):265-71.
4. Costantino CL, Schumacher LY. Surgical Technique: Minimally Invasive First-Rib Resection. *Thorac Surg Clin*. 2021;31(1):81-7.
5. Pancholy SB, Shah SC, Patel TM. Safety and Efficacy of Robotic-Assisted PCI. *Curr Cardiol Rep*. 2022;24(7):817-21.
6. Mohsin S, Gapizov A, Ekhatior C, Ain N, Ahmad S, Khan M, et al. The Role of Artificial Intelligence in Prediction, Risk Stratification, and Personalized Treatment Planning for Congenital Heart Diseases. *Cureus*. 2023;15.
7. Takebayashi T, Takahashi K, Amano S, Goshio M, Sakai M, Hashimoto K, et al. Robot-Assisted Training as Self-Training for Upper-Limb Hemiplegia in Chronic Stroke: A Randomized Controlled Trial. *Stroke*. 2022;53(7):2182-91.
8. Patil P, Patil K. A Review on Disease Prediction Using Artificial Intelligence. *Journal Electrical and Computer Experiences*. 2023.
9. Li Y, Shang K, Bian W, He L, Fan Y, Ren T, et al. Prediction of disease progression in patients with COVID-19 by artificial intelligence assisted lesion quantification. *Scientific Reports*. 2020;10.
10. Yousefi L, Tucker A. Predicting Type 2 Diabetes Complications and Personalising Patient Using Artificial Intelligence Methodology. Type 2 Diabetes [Working Title]. 2020.
11. Khalid F, Alsadoun L, Khilji F, Mushtaq M, Eze-Odurukwe A, Mushtaq MM, et al. Predicting the Progression of Chronic Kidney Disease: A Systematic Review of Artificial Intelligence and Machine Learning Approaches. *Cureus*. 2024;16.
12. Bonkhoff AK, Grefkes C. Precision medicine in stroke: towards personalized outcome predictions using artificial intelligence. *Brain*. 2022;145(2):457-75.
13. Murakami Y, Honaga K, Kono H, Haruyama K, Yamaguchi T, Tani M, et al. New Artificial Intelligence-Integrated Electromyography-Driven Robot Hand for Upper Extremity Rehabilitation of Patients With Stroke: A Randomized, Controlled Trial. *Neurorehabil Neural Repair*. 2023;37(5):298-306.
14. Huang KB, Gui CP, Xu YZ, Li XS, Zhao HW, Cao JZ, et al. A multi-classifier system integrated by clinico-histology-genomic analysis for predicting recurrence of papillary renal cell carcinoma. *Nat Commun*. 2024;15(1):6215.
15. Saliby RM, Labaki C, Jammihal TR, Xie W, Sun M, Shah V, et al. Impact of renal cell carcinoma molecular subtypes on immunotherapy and targeted therapy outcomes. *Cancer Cell*. 2024;42(5):732-5.
16. Nunes S, Sousa R, Serrano F, Branco R, Soares D, Martins A, et al. Explaining Artificial Intelligence Predictions of Disease Progression with Semantic Similarity. 2022:1256-68.
17. Liaw N, Liebeskind D. Emerging therapies in acute ischemic stroke. *F1000Res*. 2020;9.
18. Chae SH, Kim Y, Lee KS, Park HS. Development and Clinical Evaluation of a Web-Based Upper Limb Home Rehabilitation System Using a Smartwatch and Machine Learning Model for Chronic Stroke Survivors: Prospective Comparative Study. *JMIR Mhealth Uhealth*. 2020;8(7):e17216.
19. Cao Y, Shi Y, Hong W, Dai P, Sun X, Yu H, et al. Continuum robots for endoscopic sinus surgery: Recent advances, challenges, and prospects. *Int J Med Robot*. 2023;19(2):e2471.
20. Desideri F, Zinkernagel M, Anguita R. Artificial Intelligence in neovascular age-related macular degeneration. *Klinische Monatsblätter für Augenheilkunde*. 2024.
21. Rajan A, Santhosh G, Manoj B, Manohar N, Anjali T. Artificial Intelligence in Healthcare: Predicting AIDS Progression with Machine Learning Models. 2024 8th International Conference on Electronics, Communication and Aerospace Technology (ICECA). 2024:747-53.
22. Chaudhuri S, Long A, Zhang H, Monaghan C, Larkin J, Kotanko P, et al. Artificial intelligence enabled applications in kidney disease. *Seminars in Dialysis*. 2020;34:5-16.
23. Wang T, Ugurlu H, Yan Y, Li M, Li M, Wild AM, et al. Adaptive wireless millirobotic locomotion into distal vasculature. *Nat Commun*. 2022;13(1):4465.