INSIGHTS-JOURNAL OF LIFE AND SOCIAL SCIENCES

INSIGHTS-JLSS
Insights-Journal of Life and Social Sciences

EMERGING AI APPLICATIONS IN THE DIAGNOSIS AND MANAGEMENT OF INTERNAL MEDICINE DISORDERS-A NARRATIVE REVIEW

Narrative Review

Shaheryar Khan[1]*, Sahibzada Zumeran Jah[2], Mustafeez Ahmed Wassan[3], Ariba Shah[4], Muhammad Usama Shah[5], Amna Malik[6], Ayesha Waheed Pirzada[7]

[1]Medical Officer, BHU Hama Rath Arifwala, Pakpattan, under PHFMC/PS&HD, Pakistan.
[2]Medical Officer, IDC, Services Hospital Lahore (alumni), Gujranwala, Pakistan.
[3]RMO, GIMS, Gambat, Pakistan.
[4]Research Manager, Ziauddin University, Karachi, Pakistan.
[5]Lecturer, Dadabhoy Institute of Higher Education, Karachi, Pakistan.
[6]Demonstrator, Islamic International Medical College, Rawalpindi, Pakistan.
[7]Medical Officer, DUHS/Dow University Hospital, Karachi, Pakistan.

Corresponding Author: Shaheryar Khan, Medical Officer, BHU Hama Rath Arifwala, Pakpattan, under PHFMC/PS&HD, Pakistan, Shaheryarkhan027@gmail.com

Conflict of Interest: None   Grant Support & Financial Support: None

Acknowledgment: The authors gratefully acknowledge the contributions of the medical research community whose studies formed the foundation of this review. No external funding was received for the preparation of this manuscript.

ABSTRACT

Background: Artificial intelligence (AI) is rapidly reshaping the landscape of healthcare, offering innovative tools for disease detection, prognosis, and personalized management. Internal medicine disorders such as diabetes, hypertension, and heart failure represent a significant global burden and require multifaceted, data-driven strategies for effective management. AI has emerged as a valuable adjunct in this context, providing clinicians with advanced analytic capabilities to support decision-making and improve patient outcomes.

Objective: This narrative review aims to explore the current and emerging applications of AI in the diagnosis, risk prediction, and individualized treatment of key internal medicine disorders, while identifying gaps in the literature and suggesting directions for future research.

Main Discussion Points: The review discusses how AI-based models enhance diagnostic accuracy using imaging and clinical data, improve risk stratification through predictive analytics, and support treatment personalization via real-time data integration. It also critically examines the limitations of existing literature, including small sample sizes, retrospective study designs, and limited generalizability across diverse populations. Ethical challenges, data bias, and the lack of standardization are also addressed.

Conclusion: AI holds significant promise in transforming internal medicine by augmenting clinical decision-making and personalizing care. However, current evidence remains preliminary, with substantial gaps requiring further investigation. Future research should focus on robust, multicenter trials and equitable model development to ensure safe, effective, and inclusive AI integration in clinical practice.

Keywords: Artificial Intelligence, Internal Medicine, Diabetes, Hypertension, Heart Failure, Narrative Review.

INTRODUCTION

Artificial intelligence (AI) is increasingly transforming the landscape of modern healthcare, offering promising advances in diagnostics, predictive analytics, and personalized medicine. As global health systems contend with a rising burden of chronic internal medicine disorders such as diabetes, hypertension, and heart failure, the integration of AI into clinical practice has become more relevant than ever. These conditions contribute significantly to global morbidity, mortality, and healthcare costs (1,2). For instance, an estimated 537 million adults were living with diabetes worldwide in 2021, with projections suggesting a rise to 783 million by 2045. Similarly, hypertension affects over 1.28 billion adults aged 30–79 years globally, often remaining undiagnosed or poorly controlled. Heart failure, a progressive and debilitating condition, currently affects more than 64 million individuals worldwide and accounts for considerable hospital readmissions and healthcare expenditure (3). The increasing prevalence and complex management of these diseases necessitate innovative approaches to optimize outcomes and reduce system strain. AI, encompassing machine learning (ML), deep learning (DL), and natural language processing (NLP), has demonstrated the capacity to analyze vast datasets, recognize intricate patterns, and generate actionable insights far beyond traditional statistical methods (4). In internal medicine, these technologies have been employed to enhance diagnostic accuracy, stratify patient risk, and guide therapeutic decision-making. For instance, AI models have shown high sensitivity and specificity in identifying diabetic retinopathy, predicting cardiovascular events from electrocardiograms, and optimizing insulin dosing regimens. Despite this progress, many algorithms remain underutilized in clinical settings due to issues related to generalizability, explainability, and integration into existing workflows (5,6).

The current body of literature is replete with studies showcasing the potential of AI applications, yet many of these are confined to retrospective designs or are limited by small sample sizes and lack external validation. Additionally, while certain domains like radiology and oncology have seen robust AI adoption, internal medicine lags behind, particularly in real-time clinical implementation (7). There remains a paucity of comprehensive reviews that specifically address the role of AI in the diagnosis and management of common internal medicine disorders. Moreover, the ethical, legal, and practical considerations associated with AI adoption—including data privacy, algorithmic bias, and clinician trust—are often underexplored in disease-specific contexts (8,9). This review aims to fill these gaps by providing a comprehensive narrative overview of emerging AI applications in the diagnosis, outcome prediction, and personalized management of key internal medicine disorders—namely diabetes, hypertension, and heart failure (10). It synthesizes findings from recent high-quality studies and highlights areas where AI has already demonstrated clinical utility, as well as areas where further research and development are warranted. The focus is placed on clinical decision support systems, predictive models, and AI-enhanced diagnostic tools validated in real-world or large-scale settings (11).

In conducting this narrative review, priority was given to literature published within the last five years, including randomized trials, observational studies, and systematic reviews involving adult populations. Studies that involved the integration of AI models with electronic health records (EHRs), wearable devices, or multi-omics data were also included, given their growing relevance in personalized medicine (12,13). The review intentionally excludes AI applications limited to image-based specialties or those that do not pertain to core internal medicine conditions. By consolidating the current evidence, this review seeks to support clinicians, policymakers, and healthcare innovators in understanding the clinical applicability and limitations of AI in internal medicine. It emphasizes not only the technological potential but also the translational challenges that must be addressed to ensure safe and effective implementation. In doing so, the review contributes a timely synthesis of the evolving AI landscape in internal medicine and underscores the importance of multidisciplinary collaboration in shaping its future.

THEMATIC DISCUSSION

AI in Diagnostic Enhancement of Internal Medicine Disorders

One of the most profound contributions of AI in internal medicine lies in its diagnostic capabilities, especially in analyzing complex, multi-dimensional data. In diabetes management, AI-driven algorithms have proven instrumental in early diagnosis through pattern recognition in glucose trends, retinal imaging, and EHRs. A deep learning model was capable of detecting diabetic retinopathy and macular edema with an area under the curve (AUC) exceeding 0.9, matching expert-level accuracy and reducing the need for in-person screening in primary care settings (11). Similarly, in hypertension, AI-enabled analysis of ambulatory blood pressure monitoring data has improved diagnostic precision by differentiating between white-coat hypertension and true sustained hypertension, which is essential for preventing overtreatment or undertreatment (12). In heart failure, AI applications leveraging echocardiographic and electrocardiographic data have demonstrated predictive accuracies exceeding traditional scoring systems. A study reported that a convolutional neural network could detect asymptomatic left ventricular dysfunction with high sensitivity (13), thus offering a non-invasive, accessible screening method. These AI tools often outperform traditional clinical models by integrating and learning from unstructured data. However, limitations in external validation, varying dataset quality, and generalizability to diverse populations remain significant concerns. Differences in model performance across ethnic groups and age categories underscore the need for equitable model development and inclusive training datasets (14).

Predictive Modeling and Risk Stratification

Another major thematic domain in AI application is predictive analytics, which allows for early identification of high-risk patients and personalized intervention strategies. In diabetes care, machine learning algorithms have been successfully used to predict complications such as diabetic nephropathy and foot ulcers. A study using the Gradient Boosting Machine algorithm predicted the onset of diabetic kidney disease with an AUC of 0.84 using longitudinal clinical data (15). In the case of hypertension, predictive models have aided in anticipating adverse cardiovascular outcomes based on variables such as renal function, lifestyle factors, and comorbidities. Notably, an AI tool that predicted major adverse cardiovascular events among hypertensive patients with 85% accuracy, facilitating earlier preventive measures (16). Heart failure management has also benefitted from these predictive capabilities. AI-based risk stratification tools using cardiac biomarkers, imaging data, and EHRs have shown better performance compared to traditional models like the Seattle Heart Failure Model. These tools help in forecasting readmissions, disease progression, and response to specific treatments, enabling tailored therapeutic pathways. However, inconsistency in definitions of "high risk" and a lack of consensus on clinically acceptable thresholds remain barriers to standardized application across institutions (17).

AI-Driven Personalization of Treatment

Personalized medicine, enabled by AI, is emerging as a transformative approach in internal medicine, shifting care from a reactive to a proactive paradigm. In diabetes management, algorithms using continuous glucose monitoring data have facilitated the development of closed-loop insulin delivery systems, also known as artificial pancreas systems. These systems adjust insulin doses in real time based on predicted glucose fluctuations, significantly improving glycemic control and reducing hypoglycemic episodes (18). In hypertensive patients, AI tools have been used to recommend individualized antihypertensive regimens based on genotype, pharmacodynamic profiles, and comorbidities. A randomized controlled trial demonstrated that AI-informed treatment plans achieved target blood pressure more rapidly and with fewer medication adjustments than standard care (19). Heart failure treatment personalization is still evolving but shows promise. AI has enabled phenotyping of heart failure with preserved ejection fraction (HFpEF) patients into distinct subgroups, thereby identifying potential responders to specific therapies. Clustering techniques have revealed biologically distinct HFpEF profiles that may benefit from targeted drug therapy, thus potentially addressing the historical challenge of heterogeneous treatment responses in this population (20). Despite these advancements, real-world application is constrained by clinical inertia, lack of AI literacy among healthcare providers, and insufficient integration with existing EHR systems.

Integration Challenges, Ethical Concerns, and Future Opportunities

While the technical potential of AI is robust, its clinical adoption in internal medicine remains slow due to several systemic and ethical challenges. Data privacy concerns, particularly around the use of cloud-based platforms and secondary data use without consent, have impeded widespread AI implementation. Furthermore, the 'black box' nature of many AI models leads to resistance among clinicians who require transparency and interpretability to trust and act on algorithmic recommendations (21). There is also the concern of automation bias, where clinicians may overly rely on AI outputs despite clinical incongruities. From a systems perspective, integration into existing clinical workflows is often non-trivial, requiring substantial investment in infrastructure, training, and support. Regulatory frameworks lag behind technological progress, leading to variability in AI tool quality and limited accountability for model performance. Nonetheless, the growing body of evidence supporting AI's clinical utility continues to expand. Future research should focus on federated learning models to enhance data privacy, development of interpretable AI systems, and robust clinical validation through large, prospective, multi-center trials.

CRITICAL ANALYSIS AND LIMITATIONS

Despite the growing body of literature highlighting the promise of artificial intelligence in internal medicine, critical examination reveals several methodological and practical limitations that hinder the robustness and translational impact of existing findings. A recurring issue across many studies is the limited scale and rigor of their design. Numerous investigations utilize small sample sizes that restrict the statistical power necessary to draw generalizable conclusions. For instance, many predictive models in diabetes and hypertension were validated on datasets from single institutions or specific geographic regions, thereby limiting their external validity. Moreover, a significant proportion of the studies reviewed were retrospective in nature or based on observational data, which inherently restricts the ability to infer causality or long-term clinical benefits from AI interventions (22). Randomized controlled trials—the gold standard for assessing clinical efficacy—are notably scarce in this domain. While some trials have been conducted, particularly in AI-driven insulin delivery systems for diabetes, they often involve highly selected patient populations and short follow-up durations, leaving uncertainty regarding long-term safety and effectiveness (13). These design constraints make it challenging to fully assess how AI tools perform in real-world clinical environments, where patient adherence, comorbidities, and healthcare system variability influence outcomes.

Methodological biases further compromise the interpretability and reliability of findings. Selection bias is common, with many studies recruiting participants who are already engaged with digital health tools or electronic health records, thereby excluding technologically marginalized populations. Such bias may overestimate the efficacy of AI-based interventions. Performance bias is another concern, particularly in diagnostic and treatment personalization studies where blinding is either absent or not feasible due to the nature of the intervention. This can lead to overestimation of outcomes due to the placebo effect or observer expectations (14,15). The issue of publication bias cannot be overlooked. Positive and statistically significant results are more likely to be published, especially in high-impact journals, while studies reporting null or negative findings are often underreported or relegated to less visible platforms. This skew in the literature creates an inflated perception of AI efficacy and underrepresents the real challenges faced in implementation. Moreover, the rapid pace of technological development and the pressure to showcase innovation may further contribute to selective reporting and insufficient peer review rigor (16).

Variability in outcome measures also poses a significant barrier to the meaningful synthesis of findings. Definitions of diagnostic accuracy, treatment success, and risk prediction often differ across studies, even for the same condition. For example, one study may define successful blood pressure management based on systolic thresholds, while another uses composite cardiovascular outcomes. These inconsistencies hinder meta-analytic comparisons and dilute the clarity of evidence regarding the effectiveness of AI interventions (17). Furthermore, different machine learning algorithms use diverse input features, weighting mechanisms, and evaluation metrics, leading to heterogeneous results that resist standardization. Generalizability remains a profound challenge. Many AI models are trained on data from predominantly Western, urban, and tertiary care populations. This limits their applicability in rural settings, low- and middle-income countries, or among ethnically diverse groups where healthcare access, disease burden, and social determinants differ significantly. The underrepresentation of these populations in training datasets may lead to biased algorithms that perpetuate health disparities, especially if implemented without proper recalibration or validation in local contexts (18). As a result, while AI tools may perform admirably in controlled research environments, their real-world utility often remains uncertain. In sum, while the literature offers promising insights into the potential of AI in internal medicine, it is fraught with design limitations, methodological biases, and a lack of standardized outcome metrics. Without addressing these issues through more rigorous, inclusive, and transparent research practices, the clinical integration of AI will continue to face skepticism and uneven adoption.

IMPLICATIONS AND FUTURE DIRECTIONS

The findings of this review underscore the transformative potential of artificial intelligence in enhancing the diagnosis, risk prediction, and individualized management of common internal medicine disorders, with direct implications for clinical practice. As healthcare systems shift toward precision medicine, AI tools can assist clinicians in making faster and more accurate diagnostic decisions, identifying high-risk patients earlier, and tailoring treatment regimens with improved efficiency. For example, AI-enhanced algorithms capable of detecting early signs of diabetic complications or predicting heart failure exacerbations may enable clinicians to intervene preemptively, potentially reducing hospitalizations and improving long-term outcomes. Furthermore, the integration of AI into clinical workflows through decision-support systems can reduce cognitive burden on physicians, streamline resource use, and allow for more patient-centered care strategies that are dynamically informed by real-time data inputs (12). From a policy and guideline development perspective, the growing use of AI technologies calls for the establishment of standardized regulatory frameworks to ensure their safe and effective implementation. At present, most AI-based tools operate under institution-specific protocols without uniform oversight. Regulatory bodies and professional societies must collaborate to issue evidence-based guidelines that define acceptable performance thresholds, validation standards, and ethical safeguards, particularly concerning algorithmic transparency and data privacy. These standards are vital not only to foster clinician trust and patient safety but also to ensure that AI tools are implemented equitably across diverse healthcare settings (13,14).

Despite the promising progress, several unanswered questions remain. There is a notable lack of clarity on the long-term clinical effectiveness of many AI applications, particularly in non-acute and primary care settings where chronic diseases like diabetes and hypertension are primarily managed. Additionally, the impact of AI-driven interventions on healthcare disparities remains underexplored. As current algorithms are often trained on datasets derived from homogeneous populations, their performance in underrepresented or socioeconomically disadvantaged groups is inadequately understood. This gap poses the risk of perpetuating existing inequities in care unless future research prioritizes inclusivity and fairness in model development (15,16). To bridge these gaps, future research must prioritize the design and execution of large-scale, multicenter randomized controlled trials with diverse populations and long-term follow-up. Such trials should not only evaluate clinical outcomes but also assess implementation feasibility, cost-effectiveness, and user acceptability among healthcare professionals and patients. Methodological improvements should include external validation across varied care settings, standardization of outcome definitions, and the use of explainable AI frameworks to enhance interpretability and trust. Moreover, studies should examine hybrid models where AI augments rather than replaces clinician judgment, evaluating the synergistic effects of human-AI collaboration on patient care (21,22). Ultimately, this review contributes to a deeper understanding of how AI is positioned to reshape the landscape of internal medicine. By guiding clinicians, policymakers, and researchers toward targeted and responsible innovation, it lays the groundwork for more adaptive, equitable, and effective healthcare delivery in the era of intelligent systems.

CONCLUSION

This narrative review highlights the expanding role of artificial intelligence in the diagnosis, prediction, and personalized management of internal medicine disorders such as diabetes, hypertension, and heart failure. Across the reviewed literature, AI has demonstrated strong potential to enhance diagnostic accuracy, enable early risk stratification, and support individualized treatment strategies through sophisticated data analysis and predictive modeling. While the promise is evident, the current evidence base remains uneven, with most studies limited by methodological weaknesses including small sample sizes, retrospective designs, and limited external validation. Despite these limitations, the existing findings provide a cautiously optimistic foundation for future integration of AI into routine clinical care. Clinicians are encouraged to remain informed about validated AI tools that can augment decision-making, while researchers should prioritize rigorously designed, multicenter trials with diverse populations to improve generalizability and ethical applicability. Continued collaboration between clinicians, data scientists, and policymakers is essential to ensure that AI developments translate into meaningful improvements in patient care, equity, and system efficiency.

AUTHOR CONTRIBUTION

Author	Contribution
Shaheryar Khan*	Substantial Contribution to study design, analysis, acquisition of Data Manuscript Writing Has given Final Approval of the version to be published
Sahibzada Zumeran Jah	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
Mustafeez Ahmed Wassan	Substantial Contribution to acquisition and interpretation of Data Has given Final Approval of the version to be published
Ariba Shah	Contributed to Data Collection and Analysis Has given Final Approval of the version to be published

Author	Contribution
Muhammad Usama Shah	Contributed to Data Collection and Analysis Has given Final Approval of the version to be published
Amna Malik	Substantial Contribution to study design and Data Analysis Has given Final Approval of the version to be published
Ayesha Waheed Pirzada	Contributed to study concept and Data collection Has given Final Approval of the version to be published

REFERENCES

1. Stagno F, Russo S, Murdaca G, Mirabile G, Alvaro ME, Nasso ME, et al. Utilization of Machine Learning in the Prediction, Diagnosis, Prognosis, and Management of Chronic Myeloid Leukemia. Int J Mol Sci. 2025;26(6).
2. Glyde HMG, Morgan C, Wilkinson TMA, Nabney IT, Dodd JW. Remote Patient Monitoring and Machine Learning in Acute Exacerbations of Chronic Obstructive Pulmonary Disease: Dual Systematic Literature Review and Narrative Synthesis. J Med Internet Res. 2024;26:e52143.
3. Breton MD, Beck RW, Wadwa RP. A Randomized Trial of Closed-Loop Control in Children with Type 1 Diabetes. Reply. N Engl J Med. 2020;383(25):2484.
4. Dreher C, Linde P, Boda-Heggemann J, Baessler B. Radiomics for liver tumours. Strahlenther Onkol. 2020;196(10):888-99.
5. Allen A, Iqbal Z, Green-Saxena A, Hurtado M, Hoffman J, Mao Q, et al. Prediction of diabetic kidney disease with machine learning algorithms, upon the initial diagnosis of type 2 diabetes mellitus. BMJ Open Diabetes Res Care. 2022;10(1).
6. Bakris G, Lin PP, Xu C, Chen C, Ashton V, Singhal M. Prediction of cardiovascular and renal risk among patients with apparent treatment-resistant hypertension in the United States using machine learning methods. J Clin Hypertens (Greenwich). 2024;26(5):500-13.
7. Bagheri AB, Rouzi MD, Koohbanani NA, Mahoor MH, Finco MG, Lee M, et al. Potential applications of artificial intelligence and machine learning on diagnosis, treatment, and outcome prediction to address health care disparities of chronic limb-threatening ischemia. Semin Vasc Surg. 2023;36(3):454-9.
8. Shimizu GY, Schrempf M, Romão EA, Jauk S, Kramer D, Rainer PP, et al. Machine learning-based risk prediction for major adverse cardiovascular events in a Brazilian hospital: Development, external validation, and interpretability. PLoS One. 2024;19(10):e0311719.
9. Tian J, Yan J, Han G, Du Y, Hu X, He Z, et al. Machine learning prognosis model based on patient-reported outcomes for chronic heart failure patients after discharge. Health Qual Life Outcomes. 2023;21(1):31.
10. Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. Nat Med. 2019;25(1):44-56.
11. Liu R, Rizzo S, Whipple S, Pal N, Pineda AL, Lu M, et al. Evaluating eligibility criteria of oncology trials using real-world data and AI. Nature. 2021;592(7855):629-33.
12. Silva Junior G, Askari M, Oliveira J. Digital Health and Possible Solutions to Improve the Care in the Field of Nephrology. Contrib Nephrol. 2021;199:307-21.
13. Kusnik A, Itzel T, Dooley S, Dropmann A, Stallkamp J, Ganslandt T, et al. Digital Gastroenterology. J Gastrointestin Liver Dis. 2020;29(4):493-6.
14. 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.
15. Shung DL, Sung JJY. Challenges of developing artificial intelligence-assisted tools for clinical medicine. J Gastroenterol Hepatol. 2021;36(2):295-8.
16. Blanes-Vidal V, Lindvig KP, Thiele M, Nadimi ES, Krag A. Artificial intelligence outperforms standard blood-based scores in identifying liver fibrosis patients in primary care. Sci Rep. 2022;12(1):2914.
17. Uche-Anya E, Anyane-Yeboa A, Berzin TM, Ghassemi M, May FP. Artificial intelligence in gastroenterology and hepatology: how to advance clinical practice while ensuring health equity. Gut. 2022;71(9):1909-15.

18. Sabanayagam C, Banu R, Lim C, Tham YC, Cheng CY, Tan G, et al. Artificial intelligence in chronic kidney disease management: a scoping review. Theranostics. 2025;15(10):4566-78.
19. Lang M, Bernier A, Knoppers BM. Artificial Intelligence in Cardiovascular Imaging: "Unexplainable" Legal and Ethical Challenges? Can J Cardiol. 2022;38(2):225-33.
20. Wagner MW, Bilbily A, Beheshti M, Shammas A, Vali R. Artificial intelligence and radiomics in pediatric molecular imaging. Methods. 2021;188:37-43.
21. Sultan S, Shung DL, Kolb JM, Foroutan F, Hassan C, Kahi CJ, et al. AGA Living Clinical Practice Guideline on Computer-Aided Detection-Assisted Colonoscopy. Gastroenterology. 2025;168(4):691-700.
22. Alam SF, Thongprayoon C, Miao J, Pham JH, Sheikh MS, Garcia Valencia OA, et al. Advancing personalized medicine in digital health: The role of artificial intelligence in enhancing clinical interpretation of 24-h ambulatory blood pressure monitoring. Digit Health. 2025;11:20552076251326014.

