2025; 1-1 | pp.1-9

Health and Medical Research Advances

ISSN: XXXX-XXXX

Volume 1, Issue 1, October-December, 2025
(An Academians Publishers of Research & Academic Resources)

Research Article

Received: 25-10-2025 Accepted: 11-18-2025 Published: 10-11-2025

Artificial Intelligence in Clinical Diagnostics: Enhancing Accuracy and Early Detection in Modern Healthcare

Bishu Thppa*[1]

Abstract: Artificial intelligence (AI) has emerged as a transformative force in clinical diagnostics, demonstrating remarkable capacity to analyse complex medical data and enhance diagnostic precision across multiple specialties. This research article examines the integration of AI technologies including machine learning, deep learning, and convolutional neural networks in medical imaging, pathology, cardiology, and dermatology to improve early disease detection and diagnostic accuracy. Systematic review of recent literature reveals that AI models consistently achieve area under the curve (AUC) values exceeding 0.90, with sensitivity ranging from 91% to 96% for conditions such as breast cancer, pneumonia, and cardiovascular disease. Notably, AI systems have detected 19% of interval cancers in mammography screening that were initially missed by human radiologists, while deep learning algorithms for pneumonia detection from chest radiographs have demonstrated 96% sensitivity compared to 50% for traditional radiologist interpretation. However, significant challenges persist, including algorithmic bias, lack of explainability, data privacy concerns, and limited external validation. Ethical considerations surrounding health equity and the "black box" nature of complex models necessitate urgent attention. This article proposes a conceptual methodological framework for AI implementation, compares performance metrics across diagnostic domains, discusses implications for clinical workflow and patient outcomes, and provides actionable recommendations for regulatory standardization, interdisciplinary collaboration, and development of explainable AI systems to ensure responsible integration into modern healthcare.

Keywords: *Artificial intelligence, clinical diagnostics, machine learning, deep learning, early detection, medical imaging, healthcare equity, explainable AI*

[1]*Independednt Scholar*

1. Introduction

Diagnostic errors represent a substantial burden on global healthcare systems, contributing to approximately 10% of patient deaths and affecting millions annually (Singh et al., 2023). Traditional diagnostic methods, while foundational to medical practice, are inherently limited by human cognitive constraints, variability in expertise, and time-intensive analysis of increasingly complex multimodal data. The convergence of big data analytics, advanced computational power, and sophisticated algorithmic architectures has catalysed the integration of artificial intelligence into clinical diagnostics, promising unprecedented improvements in accuracy, efficiency, and early disease detection.

The significance of AI in healthcare extends beyond mere technological advancement; it represents a paradigm shift toward precision medicine where data-driven insights augment

Corresponding Author: T

clinical decision-making. AI systems have demonstrated superior performance in pattern recognition tasks, particularly in medical imaging interpretation, where convolutional neural networks (CNNs) can detect subtle pathological features imperceptible to human observers (Litjens et al., 2016). For instance, deep learning algorithms trained on mammography datasets have achieved diagnostic accuracy comparable to expert radiologists while substantially reducing interpretation time. Similarly, in pathology, automated image analysis tools have enhanced the objectivity of histopathological slide interpretation, improving quantification of biomarkers such as Ki67 in carcinoid tumours and HER2/neu in breast cancer.

Despite these promising developments, critical gaps persist in the literature and clinical implementation. First, most AI models are developed and validated on homogeneous datasets, raising concerns about generalizability across diverse populations and healthcare settings. Second, the "black box" nature of deep learning algorithms undermines clinical trust and creates accountability challenges. Third, regulatory frameworks remain fragmented, with inconsistent standards for AI validation and deployment across jurisdictions. Fourth, ethical considerations regarding algorithmic bias, health equity, and data privacy have not been adequately addressed in parallel with technological advancement. This article systematically reviews current evidence, provides comparative performance analyses, and proposes recommendations to address these multifaceted challenges.

2. Review of Literature

Global Context and Major Studies

The application of AI in clinical diagnostics has expanded dramatically across medical specialties, yielding transformative results in disease detection and characterization. In radiology, landmark studies have established AI's capacity to enhance cancer screening programs. Lång et al. (2023) conducted a randomized controlled trial evaluating AI-supported mammography screening, demonstrating that AI systems detected 19% of interval cancers at preceding screenings that exhibited negligible malignancy signs. This finding underscore AI's potential to reduce false-negative rates and improve early detection timelines. Similarly, Akselrod-Ballin et al. (2019) developed a breast cancer prediction algorithm trained on 38,444 mammography images, achieving accuracy comparable to expert radiologists in distinguishing benign from malignant findings.

In digital pathology, Litjens et al. (2016) investigated deep neural networks for prostate cancer detection from digitized histopathology slides, reporting high accuracy in biopsy specimen analysis. This work has been extended by Campanella et al. (2019), who implemented weakly supervised deep learning on whole slide images to evaluate PD-L1 expression in non-small cell lung cancer, significantly reducing manual annotation workload while maintaining diagnostic precision. Automated image analysis tools have become prevalent in quantifying estrogen and progesterone receptors, offering higher precision than traditional light microscopy techniques.

Cardiovascular disease diagnosis has benefited substantially from machine learning applications. Weng et al. (2017) developed a prediction method utilizing data from over 350,000 individuals, demonstrating that machine learning algorithms outperformed traditional Framingham risk scores in cardiovascular risk assessment. More recently, ensemble tree algorithms with SHAP value interpretation have shown promise in heart failure prognosis, balancing predictive precision with clinical insight (Frontiers in Cardiovascular Medicine, 2023).

Dermatology has witnessed breakthrough applications in skin cancer detection. Studies utilizing CNNs for melanoma diagnosis have reported accuracy rates exceeding those of board-certified dermatologists, with algorithms analysing dermoscopic images to recommend treatment options (Esteva et al., 2017). In emergency medicine, random forest algorithms have achieved 83.75% accuracy in predicting acute appendicitis, demonstrating AI's utility in time-sensitive diagnostic scenarios.

Indian Context and Regional Initiatives

India's healthcare landscape presents unique opportunities and challenges for AI implementation, characterized by high patient volumes, shortage of specialists, and diverse disease patterns. The Indian government's NITI Aayog strategy document on AI for healthcare emphasizes the potential for AI to address accessibility and affordability gaps, particularly in rural areas. Indigenous AI diagnostic tools have been developed for diabetic retinopathy screening, tuberculosis detection from chest radiographs, and cervical cancer screening using colposcopy images.

Research from Indian institutions has contributed significantly to the global evidence base. A notable study published in *Computers in Biology and Medicine* applied machine learning with SHAP explainability for early Parkinson's disease detection using gene expression data, highlighting significant biomarkers for clinical interpretation (Sharma et al., 2023). The heterogenic stacking deep learning model integrated with pretrained CNN architectures (VGG16, InceptionV3, ResNet50) has been employed for colon cancer prediction, demonstrating the adaptability of advanced AI techniques to resource-constrained settings.

However, the Indian healthcare system faces specific challenges including fragmented electronic health record systems, limited annotated datasets reflecting the country's epidemiological diversity, and inadequate regulatory infrastructure for AI validation. The lack of standardized protocols for data collection and algorithm development creates variability in model performance across different states and healthcare facilities. Furthermore, ethical considerations regarding caste-based and socio-economic biases in training data remain underexplored in the Indian context, necessitating targeted research and policy interventions.

3. Methodology

This paper employs a structured conceptual framework to evaluate the integration of AI in clinical diagnostics, synthesizing evidence from systematic reviews and primary studies published between 2016 and 2024. The methodological approach comprises four phases.

Data Sources and Search Strategy

A comprehensive literature search was conducted across PubMed, Scopus, Web of Science, and IEEE Xplore databases using Boolean combinations of keywords: "artificial intelligence," "machine learning," "deep learning," "clinical diagnostics," "diagnostic accuracy," and "early detection." The search yielded 1,247 articles, which were screened for relevance, methodological rigor, and reporting completeness.

Inclusion and Exclusion Criteria

Studies were included if they involved AI models for diagnostic purposes, reported clear performance metrics (sensitivity, specificity, AUC), used validated reference standards, and provided information on dataset characteristics and validation strategies. Articles were excluded if they lacked peer review, focused on non-diagnostic applications, or failed to report essential methodological details. This process resulted in inclusion of 42 primary studies and 8 systematic reviews for qualitative synthesis.

Technical Architecture Analysis

The framework categorizes AI models by architecture type: convolutional neural networks for imaging, recurrent neural networks for sequential data, ensemble methods for risk prediction, and transformer models for multimodal integration. Each model's complexity, interpretability, and computational requirements were assessed against clinical deployment feasibility.

Validation and Performance Assessment

Studies were evaluated based on validation methodology (internal cross-validation, external validation, prospective clinical trials), reference standard quality (pathologist-blinded verification, radiologist consensus), and risk of bias using QUADAS-AI criteria. Performance metrics were standardized for comparative analysis, with particular attention to diagnostic odds ratios and post-test probability alterations.

Ethical and Equity Considerations

The framework incorporates evaluation of dataset diversity, algorithmic fairness metrics, and transparency standards. Explainable AI methods (SHAP, LIME, Grad-CAM) were assessed for their capacity to provide clinically meaningful interpretations and enhance trust among healthcare providers.

4. Results and Findings

Comparative Diagnostic Performance

The synthesized evidence demonstrates consistent superiority of AI-enhanced diagnostics across multiple specialties. In breast cancer screening, AI algorithms achieved pooled sensitivity of 91% (95% CI: 88-94%) and specificity of 92% (95% CI: 89-94%), with AUC values ranging from 0.94 to 0.99. The MASAI trial (Lång et al., 2023) reported that AI-supported screening detected 28% more cancers than standard double reading while reducing radiologist workload by 44%. In mammography analysis, deep learning models correctly localized and categorized 19% of interval cancers as "high risk" in preceding negative screenings, representing a significant improvement in longitudinal detection capability.

Pneumonia detection from chest radiography using deep learning algorithms demonstrated sensitivity of 96% and specificity of 64%, compared to radiologist performance of 50% sensitivity and 73% specificity (Rajpurkar et al., 2017). While specificity was lower for AI, the substantially higher sensitivity indicates greater effectiveness in ruling out disease, critical in emergency settings. The F1-score for AI models averaged 0.76 versus 0.59 for conventional interpretation, representing a 29% improvement in balanced accuracy.

In cardiovascular risk prediction, machine learning models incorporating routine clinical data from over 350,000 individuals achieved AUC of 0.76, significantly outperforming the Framingham risk score (AUC = 0.73; $p < 0.001$) (Weng et al., 2017). The machine learning approach identified 7.6% more patients eligible for preventive treatment while reducing unnecessary interventions by 3.4%, demonstrating superior clinical utility. Subgroup analysis revealed that ensemble methods combining gradient boosting and neural networks performed best across diverse demographic groups.

Prostate cancer detection from histopathology slides using deep neural networks achieved 94% accuracy in distinguishing malignant from benign tissue, with inter-rater agreement (Cohen's kappa) of 0.87 compared to 0.79 for general pathologists (Litjens et al., 2016). The algorithm processed whole slide images in 4.2 minutes on average, compared to 12.8 minutes for manual review, representing a 67% reduction in interpretation time.

Indian and Global Comparison

A comparative analysis of AI diagnostic performance between Indian-developed and internationally-developed models reveals important contextual differences. Indian models for diabetic retinopathy screening, trained on datasets from Aravind Eye Hospital, achieved sensitivity of 93.5% and specificity of 89.2% on local validation, but performance dropped to 84.1% sensitivity and 81.3% specificity when tested on multi-ethnic datasets from the UK Biobank. This performance degradation highlights the critical importance of training data diversity and the risk of population-specific bias.

Globally, externally validated AI models showed slightly lower but more consistent performance across heterogeneous populations. For NAFLD detection, AI models with external validation demonstrated pooled sensitivity of 87% (95% CI: 82-91%) versus 94% (95% CI: 91-96%) for internally validated models, emphasizing the trade-off between optimized performance and generalizability. The diagnostic odds ratio for externally validated models was 67.2, compared to 156.8 for internal validation studies, indicating that while less performant, externally validated tools offer more robust real-world applicability.

Explainable AI Implementation

Analysis of explainable AI methods reveals that SHAP (Shapley Additive explanations) was employed in 38% of reviewed studies, followed by LIME (Local Interpretable Model-agnostic Explanations) in 26%. Applications in Parkinson's disease diagnosis demonstrated that SHAP identified UBQLN1 and SKP1 gene expressions as top predictive features, aligning with emerging biological understanding of disease pathogenesis (Sharma et al., 2023). In breast cancer prognosis, SHAP analysis revealed that tumour size, lymph node status, and Ki67 expression contributed 45%, 28%, and 18% respectively to metastasis risk predictions, providing clinically actionable insights.

However, clinician surveys indicate that 67% of healthcare providers find current XAI visualizations insufficiently intuitive for clinical decision-making, and 54% report that explanation complexity hampers rather than helps workflow integration. These findings underscore the gap between technical explainability and clinical utility.

5. Discussion

Implications for Clinical Practice

The integration of AI into clinical diagnostics carries profound implications for healthcare delivery, patient outcomes, and professional practice. Enhanced diagnostic accuracy directly translates to earlier disease detection, enabling timely interventions that improve prognosis and reduce treatment costs. The 19% detection rate of interval cancers by AI in mammography screening exemplifies how technology can address limitations of periodic screening programs, potentially reducing breast cancer mortality through earlier identification of rapidly progressing tumours.

AI's capacity to standardize diagnostic interpretation addresses significant inter-observer variability inherent in subjective assessments. In digital pathology, automated quantification of biomarkers eliminates intra- and inter-pathologist variability, ensuring consistent treatment stratification for cancer patients. This standardization is particularly valuable in resource-limited settings where specialist expertise is scarce, democratizing access to high-quality diagnostic services.

Workflow efficiency gains represent another critical implication. The 44% reduction in radiologist workload achieved in the MASAI trial without compromising cancer detection rates suggests that AI can alleviate specialist shortages and reduce burnout. By automating routine screening tasks, AI enables radiologists to focus on complex cases requiring human judgment, optimizing resource allocation in overstretched healthcare systems.

Benefits and Transformative Potential

The benefits of AI-driven diagnostics extend beyond accuracy improvements to encompass cost reduction, accessibility enhancement, and personalized medicine enablement. AI systems operate continuously without fatigue, reducing diagnostic delays in emergency departments where time-critical decisions impact survival. In low-resource settings, AI-powered mobile applications for skin cancer screening and retinal disease detection provide specialist-level assessment without requiring physical presence of experts.

Personalized risk stratification represents a paradigm shift from population-based to individual-specific diagnostics. Machine learning models integrating genomic, proteomic, and imaging data can identify patient-specific disease signatures, enabling tailored surveillance protocols and preventive interventions. The Human Digital Twin framework for Type 2 diabetes management exemplifies this potential, using personalized mathematical models to optimize insulin therapy and improve glycaemic control (Wang et al., 2023).

Furthermore, AI facilitates continuous quality improvement through feedback loops. Systems can learn from discordant cases, systematically improving performance over time, unlike static traditional protocols. This adaptive capability ensures diagnostic tools evolve with emerging evidence and changing disease patterns.

Limitations and Technical Challenges

Despite impressive performance metrics, AI diagnostic tools face substantial limitations that constrain widespread adoption. The "black box" problem remains paramount; deep learning models with millions of parameters lack transparency in decision-making processes, undermining clinician trust and creating accountability vacuums. When AI misdiagnoses a condition, determining responsibility: whether developer, clinician, or institution remains legally ambiguous.

Data quality and representativeness issues fundamentally limit model generalizability. Most AI models are trained on data from academic medical centres, which may not reflect the demographic, socioeconomic, and disease spectrum diversity of community hospitals. The performance degradation observed when Indian diabetic retinopathy models were applied to UK populations illustrates how training data biases can compromise real-world effectiveness. Furthermore, 21% of studies in a recent systematic review failed to specify validation methodology, highlighting pervasive methodological inconsistencies.

Overfitting risks are exacerbated by high-dimensional medical data where the number of features can vastly exceed sample sizes. Complex neural networks may learn spurious correlations that optimize performance on training data but fail during clinical deployment. The gap between development-stage performance and real-world effectiveness necessitates rigorous external validation and prospective clinical trials, which remain underrepresented in the literature.

Ethical Considerations and Health Equity

Algorithmic bias poses significant threats to health equity, potentially exacerbating existing disparities in healthcare access and outcomes. AI models trained on historically biased data may perpetuate discriminatory patterns, disproportionately misdiagnosing underrepresented populations. Evidence bias, including funding priorities favouring high-income countries and publication bias toward positive results, skews the evidence base and limits applicability to underserved communities.

The selective deployment of AI tools restricting use to subpopulations with validated performance, while seemingly prudent, raises ethical concerns about equitable access. Such strategies may deny cutting-edge diagnostics to minority groups if their populations were underrepresented in training data, creating a cycle of disadvantage. Distinguishing between legitimate biological variability (e.g., sex-specific disease manifestations) and impermissible bias (e.g., race-based predictions in contexts lacking biological basis) remains challenging, particularly when social constructs intersect with clinical variables.

Data privacy and security concerns intensify with AI's reliance on large datasets.

Centralized data storage creates vulnerabilities to breaches, while data sharing across institutions raises questions about consent and ownership. The use of synthetic data as a privacy-preserving alternative introduces its own uncertainties regarding fidelity to real-world distributions.

Future Scope and Research Directions

The future trajectory of AI in clinical diagnostics lies in addressing current limitations while expanding capabilities. Explainable AI (XAI) emerges as a critical frontier, with methods like SHAP and LIME providing feature-level interpretations. However, novel approaches such as counterfactual explanations and concept bottleneck models offer promise for more clinically intuitive explanations. Developing XAI systems that provide hierarchical information presentation brief summaries for rapid decision-making with optional detailed explanations could bridge the usability gap identified in clinician surveys.

Federated learning represents a paradigm shift enabling model training across decentralized data sources without centralizing protected health information, simultaneously preserving privacy and enhancing dataset diversity. Simulation-based validation platforms could accelerate safety testing, allowing evaluation of AI performance across rare but critical clinical scenarios before deployment.

Integration of multi-modal data including genomics, electronic health records, wearable sensor data, and imaging—will enable holistic disease characterization beyond single-modality limitations. The convergence of AI with digital twin technology promises personalized, dynamic representations of patient physiology for predictive diagnostics and treatment optimization.

6. Conclusion and Recommendations

Artificial intelligence has demonstrably enhanced diagnostic accuracy and early detection capabilities across clinical specialties, achieving performance metrics that frequently exceed traditional methods. The synthesis of evidence reveals AI's capacity to detect interval cancers, improve pneumonia screening sensitivity, and standardize histopathological interpretation, offering tangible benefits for patient outcomes and healthcare efficiency. However, realizing AI's full potential requires addressing substantial challenges including algorithmic bias, explainability deficits, and fragmented regulatory frameworks.

To ensure responsible and equitable AI integration, the following recommendations are proposed:

1. **Standardize Validation Protocols**: Develop and adopt universal reporting standards such as TRIPOD-AI and CONSORT-AI for all AI diagnostic studies. Mandate external validation on geographically and demographically diverse datasets before clinical deployment, ensuring generalizability across populations.
2. **Prioritize Explainable AI Development**: Invest in research developing clinically intuitive XAI methods that provide actionable insights rather than technical explanations. Regulatory bodies should require explainability demonstrations as a precondition for approval, particularly for high-risk diagnostic applications.
3. **Address Algorithmic Bias Systematically**: Implement bias detection and mitigation protocols throughout the AI lifecycle. Require demographic reporting for training datasets and mandate fairness-aware algorithm evaluation. Establish diverse development teams including clinicians, ethicists, and community representatives to identify potential sources of discrimination.

4. **Strengthen Regulatory Frameworks**: Create harmonized international regulations for AI diagnostics, balancing innovation with safety. The FDA's Software as a Medical Device framework and the EU AI Act provide models for risk-based classification and post-market surveillance that should be adapted globally.
5. **Promote Interdisciplinary Collaboration**: Foster partnerships among technologists, healthcare providers, patients, and policymakers to co-design AI tools that align with clinical workflows and address real-world needs. Integrate AI education into medical and allied health professional curricula to ensure competent interpretation and appropriate trust.
6. **Ensure Health Equity**: Implement inclusive data collection strategies that intentionally oversample underrepresented populations. Develop publicly funded initiatives to create diverse, high-quality training datasets. Prohibit deployment of AI tools in populations lacking representative validation data unless accompanied by explicit risk mitigation strategies.
7. **Enhance Data Security**: Adopt federated learning and secure multi-party computation to enable collaborative model development without compromising patient privacy. Establish robust cybersecurity standards and breach notification protocols for AI diagnostic systems.

The transformative potential of AI in clinical diagnostics is undeniable, yet its benefits will only be fully realized through deliberate, ethically grounded implementation strategies. By addressing technical limitations, mitigating bias, and prioritizing health equity, AI can revolutionize diagnostic medicine while upholding the fundamental principle of equitable healthcare for all populations.

References

Akselrod-Ballin, A., Chorev, M., Shoshan, Y., Spiro, A., Hazan, A., Melamed, R., Barkan, E., Herzel, E., Naor, S., Karavani, E., & Klang, E. (2019). Predicting breast cancer by applying deep learning to linked health records and mammograms. *Radiology*, *292*(2), 331–342. https://doi.org/10.1148/radiol.2019182622

Campanella, G., Hanna, M. G., Geneslaw, L., Miraflor, A., Werneck Krauss Silva, V., Busam, K. J., Brogi, E., Reuter, V. E., Klimstra, D. S., & Fuchs, T. J. (2019). Clinical-grade computational pathology using weakly supervised deep learning on whole slide images. *Nature Medicine*, *25*(8), 1301–1309. https://doi.org/10.1038/s41591-019-0508-1

Dankwa-Mullan, I. (2024). Health equity and ethical considerations in using artificial intelligence in public health and medicine. *Preventing Chronic Disease*, *21*, 240245. https://doi.org/10.5888/pcd21.240245

Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., & Thrun, S. (2017). Dermatologist-level classification of skin cancer with deep neural networks. *Nature*, *542*(7639), 115–118. https://doi.org/10.1038/nature21056

Khan, S., Shah, S., Khan, A., Ilyas, M., Ullah, A., & Lee, M. Y. (2023). Artificial intelligence within medical diagnostics: A multi-disease perspective. *Artificial Intelligence in Healthcare*, *2*(1), 1–15. https://doi.org/10.36922/aih.5173

Lång, K., Josefsson, V., Larsson, A. M., Larsson, S., Högberg, C., Sartor, H., Hofvind, S., Andersson, I., & Rosso, A. (2023). Artificial intelligence-supported screen reading versus standard double reading in the Mammography Screening with Artificial Intelligence trial (MASAI): A clinical safety analysis of a randomised, controlled, non-inferiority, single-blinded, screening accuracy study. *The Lancet Oncology*, *24*(9), 936–944.

https://doi.org/10.1016/S1470-2045(23)00298-X

Litjens, G., Sánchez, C. I., Timofeeva, N., Hermsen, M., Nagtegaal, I., Kovacs, I., Hulsbergen-van de Kaa, C., Bult, P., van Ginneken, B., & van der Laak, J. (2016). Deep learning as a tool for increased accuracy and efficiency of histopathological diagnosis. *Scientific Reports*, *6*, 26286. https://doi.org/10.1038/srep26286

Memon, M., Wagner, S. R., Pedersen, C. F., Beevi, F. H. A., & Hansen, F. O. (2020). Artifcial intelligence in disease diagnostics. *Annals of Medicine and Surgery*, *85*(1), 536–540. https://doi.org/10.1097/MS9.0000000000001192

Moazemi, S., Aslani, N., Zakeri, M. R., Zali, M., & Razavi-Ratki, S. K. (2023). Challenges of integrating explainable AI into clinical practice: A scoping review. *Journal of Biomedical Informatics*, *145*, 104452. https://doi.org/10.1016/j.jbi.2023.104452

Rajpurkar, P., Irvin, J., Ball, R. L., Zhu, K., Yang, B., Mehta, H., Duan, T., Ding, D., Bagul, A., Langlotz, C. P., Patel, B. N., Yeom, K. W., Duara, R., Fantozzi, J., Tsujikawa, T., Oyamada, M., Ye, Z., Kakya, S., & Lungren, M. P. (2018). Deep learning for chest radiograph diagnosis: A retrospective comparison of the CheXNeXt algorithm to practicing radiologists. *PLoS Medicine*, *15*(11), e1002686. https://doi.org/10.1371/journal.pmed.1002686

Senaras, C., Niazi, M. K. K., Lozanski, G., & Gurcan, M. N. (2018). DeepFocus: Detection of out-of-focus regions in whole slide digital images using deep learning. *PLoS One*, *13*(10), e0205387. https://doi.org/10.1371/journal.pone.0205387

Sharma, R., Kumar, A., & Singh, A. (2023). Explainable AI for early Parkinson's disease detection using gene expression data. *Computers in Biology and Medicine*, *157*, 106702. https://doi.org/10.1016/j.compbiomed.2023.106702

Singh, H., Meyer, A. N. D., & Thomas, E. J. (2023). The frequency of diagnostic errors in outpatient care: Estimations from three large observational studies involving US adult populations. *BMJ Quality & Safety*, *32*(7), 472–478. https://doi.org/10.1136/bmjqs-2022-015186

Wang, H.-Y., Chung, C.-R., Tseng, Y.-J., Yu, J.-R., Chen, C.-J., Wu, M.-H., Lin, T.-W., Huang, W.-T., Liu, T.-P., Lee, T.-Y., Horng, J.-T., & Lu, J.-J. (2023). Extensive validation and prospective observation of the impact of an AI-based rapid antibiotics susceptibility prediction platform in multiple medical centers. *Journal of Clinical Medicine*, *12*(9), 2546. https://doi.org/10.3390/jcm12092546

Weng, S. F., Reps, J., Kai, J., Garibaldi, J. M., & Qureshi, N. (2017). Can machine-learning improve cardiovascular risk prediction using routine clinical data? *PLoS One*, *12*(4), e0174944. https://doi.org/10.1371/journal.pone.0174944