WWW.IJRPR.COM

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Redefining Molecular Diagnostics: A Study on Human-AI Collaboration

Avani Thakare[1], Anushka Surdas[1], Rushika Dandare[2], Miheka Mali[2]

[a]First affiliation: MIT WPU , Kothrud, Pune-411038, India.
[b]Second affiliation: MIT ADT University, Loni Kalbhor, Pune-412201, India.

A B S T R A C T :

Artificial Intelligence (AI) and automation are also heavily deployed in most of the diagnostic pursuits, particularly in imaging specialties like radiology and pathology where they play a role in interpreting the imaging report. These workflow-improved and optimized turnaround times are particularly important in high-data-density workflows wherein speed of interpretation is of critical concern.

The increasing significance and sophistication of molecular diagnostics, much of the process downstream of sample handling remains largely manual. Although processes such as nucleic acid extraction and sequencing are progressively performed on automated platforms, the more comprehensive integration of artificial intelligence into these processes has not been fully achieved. There remains a large gap in the technological assistance for interpreting molecular information an aspect most important to providing accurate and timely results, particularly with increased demand for high-throughput testing.

This study integrates one academic high-throughput clinical genomics laboratory's workflow outcomes with literature review to evaluate the extent of accessible automation and AI-readiness in molecular diagnostics. Each phase of the workflow, from sample accessioning right through to amplification, sequencing, and reporting, was noted and categorized according to whether it is currently automated, manually performed, or ready for AI-assisted processes. Technologies such as KingFisher Flex, Illumina MiSeq, Oxford Nanopore, and Sanger sequencing platforms were discovered to efficiently process samples, yet downstream reporting and interpretation remain to a great extent based on human analysis, an indicator of the need for supportive technologies to ensure dependability and efficacy.

Keywords: Artificial Intelligence (AI), Molecular Diagnostics, Automation, High-throughput Testing, Sequencing Platforms, Interpretation.

1. Introduction

Artificial intelligence (AI) has revolutionized various clinical diagnostic techniques, particularly those dependent on image-based evaluation. In the management of rare diseases, AI plays a part in precision medicine by applying vast amounts of patient registry data to reveal possible associations and hence tailoring prevention, diagnosis, treatment, and monitoring to individualized genetic profiles [1] [16]. In digital pathology, dermatology, and radiology, deep learning algorithms are used to detect anomalies, segment lesions, and even create initial reports. Numerous studies have demonstrated that when AI is embedded within human-in-the-loop (HITL) systems, it can significantly enhance diagnostic efficiency without compromising clinical oversight [5] [6] [10]. These systems not only streamline routine assessment tasks but also reduce diagnostic fatigue and delays among clinicians. Traditional epidemiological tools are being significantly complemented by recent innovations especially using artificial intelligence (AI) and machine learning. AI-based model systems could improve pattern recognition of disease spread in populations and predictions of outbreaks in different geographical locations [7].

AI-based modeling systems have been found to indicate disease spread patterns and forecast outbreaks in various geographic regions. AI has progressed from experimental technology to clinical decision support systems in imaging diagnostics and pathology. Pathologists are now applying machine learning algorithms on a regular basis to determine tumor histology slides, detect features like nuclear pleomorphism and mitotic figures, and assist in classification and grading. As shown in Fig 1 [12], a broad array of AI approaches from convolutional neural networks to decision-support systems are now incorporated into routine diagnostic routines. This incorporation has streamlined turnaround times and enhanced reproducibility, showing what can be accomplished when AI is meaningfully integrated into diagnostic practice an attitude not yet broadly applied in molecular diagnostics. The application of AI as a component of the diagnostic process provides opportunities for innovative digital health and is simultaneously able to ensure enhanced patient outcomes [8]. While AI has achieved remarkable progress in imaging specialties, its clinical use remains heterogeneous across specialties. Interestingly, the two disease categories to which AI has been applied most intensively oncology and COVID-19 are both heavily dependent on molecular diagnostics.
Blood cancers like leukemia and lymphoma are generally evaluated through qPCR, flow cytometry, and next-generation sequencing (NGS), whereas testing for COVID-19 has largely relied on RT-qPCR and sequencing technologies. Ironically, though AI is broadly utilized to diagnose radiological and

histological images in such diseases, the molecular diagnostic pathways for the same conditions are still untouched by AI. AI-based technology has not yet achieved the full potential for diagnosis in COVID-19, but it can still provide useful information to doctors if considered together with clinical findings and symptoms. False-positive findings with the AI model are likely possible in those patients diagnosed with COVID-19 who tested negative by PCR but were positive for COVID-19 [14]. This reflects a notable translational divide AI has been able to cross at the disease stage but not the molecular diagnostic tools employed in diagnosing them.

These advancements in technology have yet to make a complete translation into the molecular diagnostic sphere. Even though all these protocols utilize advanced, data-driven tools like qPCR, digital PCR, Sanger sequencing, and NGS, the majority of the interpretive steps of these protocols are still labor-intensive.

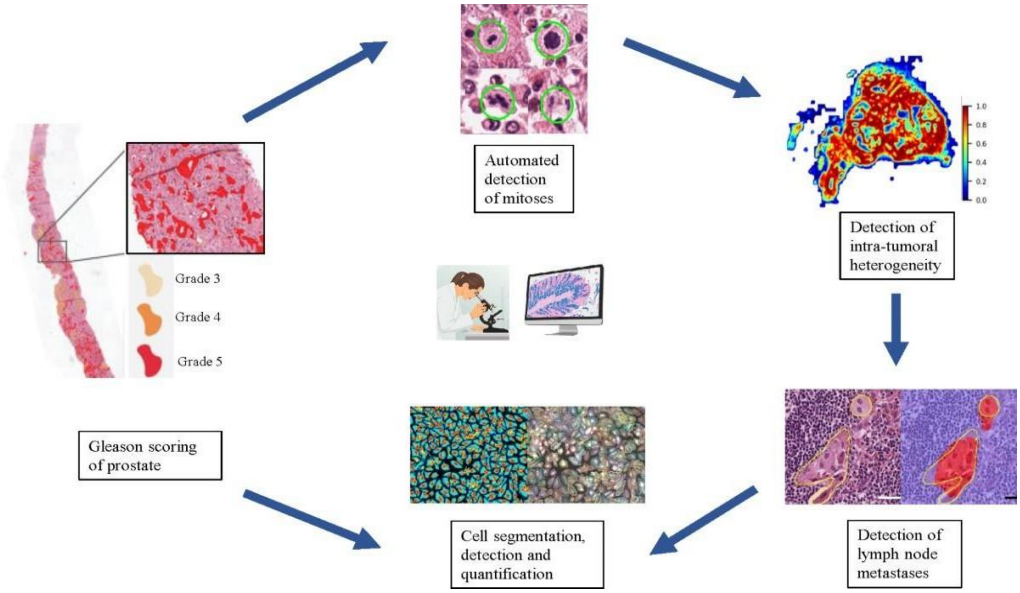


Fig 1: Depiction of Artificial intelligence (AI) and machine learning approaches currently used by pathologists to analyze images from tumor [12].

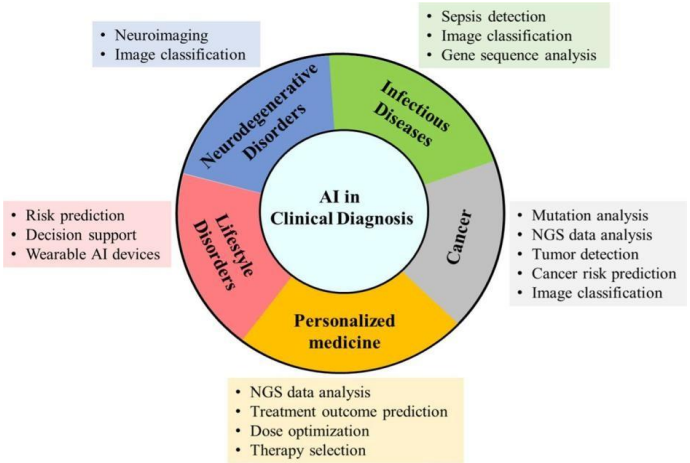
Whereas the upstream procedures like nucleic acid isolation (e.g., KingFisher Flex) and sequencing (e.g., Illumina MiSeq, Oxford Nanopore) have been efficiently automated, downstream interpretation such as Ct curves reading, melting peaks, electropherograms, and variant classification still depends on human interpretation in the majority of clinical labs [4]. In actual high-throughput molecular labs such as the one seen in this analysis, manual interpretation is still a main impediment. An example is the process of NGS data analysis, which typically includes command-line processing to create .bam and .vcf files, which are then inspected in genome browsers and checked base by base for mutations. Such a process is not only time-consuming and fatiguing for the brain but also subjected to inter-analyst error and misses.

Even ordinary qPCR tests involve manual review of melt and amplification curves, typically without the luxury of automated flagging or interpretation software.

Physicians also find it difficult to interpret raw molecular results like Ct values or electropherograms leading to delays and greater reliance on laboratory staff for interpretation. As molecular diagnostics moves into oncology, infectious disease, and rare disease markets, the quantity and complexity of samples increase exponentially. Without smarter support systems, diagnostic teams experience increasing burnout, longer turnaround times, and a growing risk for missed or delayed interpretations particularly in life-critical cases. While there are AI-based platforms like DeepVariant and VarSome, they are commonly limited to use in research environments or applied in isolated, non-integrated forms throughout the diagnostic pipeline [12].

Some of the specific uses of AI in diagnostics involve prognosis determination from tumor gene expression profiling, nucleic acid sequence analysis variant calling, and evaluation of variant pathogenicity.

Fig 2: Applications of AI in clinical diagnosis of various diseases [11].



Bioinformaticians and computer scientists are critically involved in formulating algorithms that help geneticists, pathologists, and molecular biologists analyze and combine laboratory and clinical information. In addition, AI systems have been investigated for applications like karyotype assignment, base calling, variant calling, variant evaluation, and tissue of origin identification in cancers with unknown primary [9]. The image highlights how AI-powered liquid biopsy, using blood-based biomarkers like circulating tumor DNA (ctDNA) and circulating tumor cells (CTCs), enables early and non-invasive cancer detection. Combined with next-generation sequencing and multi-omics data (genomics, transcriptomics, proteomics), AI and machine learning facilitate precise biomarker identification and personalized risk assessment.

This concurs with the work of [15], which highlights the revolutionizing capabilities of AI in improving diagnosis accuracy and personalizing cancer treatment according to genetic makeup. Concurrently, it is a legitimate concern to avoid the complete delegation of interpretive powers to AI. In contrast to imaging, which can usually depend upon recurring visual patterns, molecular diagnosis demands sophisticated clinical judgment, phenotype- genotype correlation, and capability to interpret indefinite or borderline results

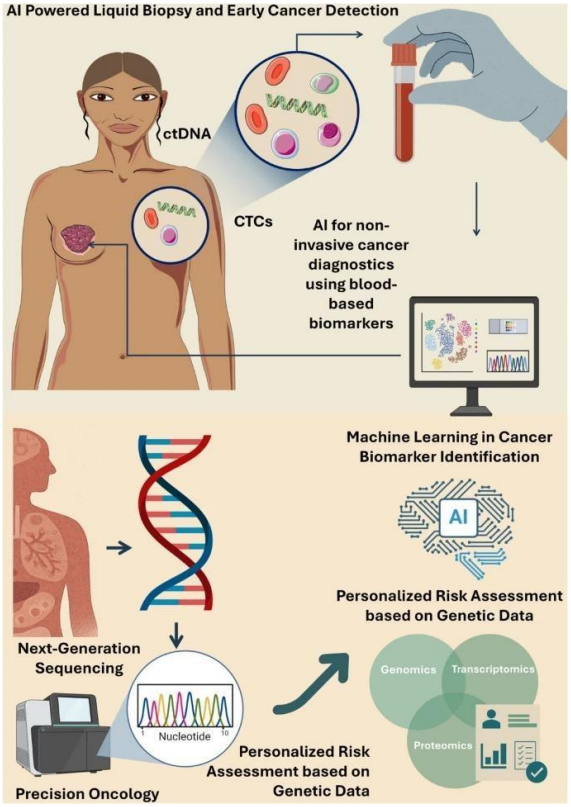


Fig 3: AI powered liquid biopsies and early cancer detection [15]

AI programs built upon past datasets are not at present capable of equaling the richness of clinical reasoning or the ethical acumen required to establish pathogenicity or diagnostic adequacy. Thus, the future of AI in molecular diagnostics is not to usurp human expertise, but to facilitate efficiency, minimize repetitive work, and aid in informed clinical decision-making. The literature increasingly highlights the value of explainable AI (XAI) and human-in-the-loop methodologies—models that notify, recommend, or rank findings without making ultimate clinical choices [10] [17]. In spite of increased recognition of the potential of AI, a holistic, task-oriented framework for integrating it into molecular diagnostics

pipelines are yet to emerge. Published literature tends to center on individual tools, particular pipelines, or algorithmic metrics, instead of providing systemic, workflow-driven advice for practical application. Consequently, most diagnostic labs still find themselves stuck between automated upstream processing and downstream labor-intensive, manual analysis, without a discernible strategy to integrate them. In response to this deficiency, this study simultaneously integrates a comprehensive literature review with on-site workflow observation in a high- throughput clinical genomics lab. We uniformly analyze each stage of the molecular diagnostics pipeline—ranging from sample accessioning and nucleic acid extraction through PCR interpretation, NGS analysis, and report generation—and group them into three zones: (1) automation-ready, (2) collaboration-ready, and (3) human-exclusive. Our goal is to suggest an operational and ethical model for adopting AI into molecular diagnostics—not to supplant clinical judgment, but to alleviate workloads, enhance consistency, and assist with quicker, safer diagnostic conclusions.

2. Methodology

The research used a twofold strategy involving the systematic observation of the workflow of an existing high-throughput clinical genomics laboratory as well as a focused review of current literature on the application of AI in diagnostics. The aim was to determine locations within the molecular diagnostics workflow where incorporation of AI is practical, useful, and ethically acceptable.

A) Workflow Mapping and Observation-

The diagnosis process in the high-throughput molecular lab was qualitatively observed using direct observation. The lab is an expert in clinical testing of infectious diseases, oncology, and genetic diseases through nucleic acid extraction platforms, quantitative PCR (qPCR), Sanger sequencing, and next-generation sequencing (NGS) technology platforms such as Illumina and Oxford Nanopore systems. The molecular pipeline in each step was evaluated for its:

- Level of automation
- Amount of human intervention
- Cognitive and interpretive burden
- Present or potential AI intervention

Each stage was captured through observation, i.e., sample accessioning, extraction, amplification, sequencing, analysis of data, and generation of report. Special attention was paid to recording interpretation stages such as variant analysis through .vcf files in genome browsers and task bottlenecks such as report generation and clinician engagement.

B) Literature Integration-

Recent peer-reviewed evidence (2019–2025) were combined to place the observations in context and in relation to international trends. The articles were chosen based on relevance to AI in molecular medicine, diagnostics workflow, human-AI collaboration, and explainable AI (XAI). Instead of algorithmic comparison alone, the literature was used to examine greater themes including:

- Effectiveness of AI imaging compared to underutilization of AI in molecular interpretation
- Human-in-the-loop design principles
- Barriers to AI adoption in high-stakes clinical diagnosis
- Current fragmentation of tools along molecular pipelines

C) Workflow Classification Model-

Consensus findings and review were used to classify each step in the molecular workflow into one of three zones:

- Automation-Ready: Rule-based, repetitive steps and suitable for complete automation. Example: nucleic acid extraction, thermocycling protocols, file formatting
- AI–Human Collaborative: Human-monitored steps where AI can contribute to speed, consistency, or pre-analysis. Example: variant prioritization, melt curve interpretation, electropherogram quality checking
- Human-Exclusive: Judgments that call for clinical judgment, integration of context, or ethics deliberation. Example: clinical significance determination of a variant, diagnosis conclusion, reporting communication

These three zones were then utilized to classify the decision-making into actionable points for safe and effective AI integration without jeopardizing the diagnostic integrity.

3. Result

Observation of workflow in a high-throughput molecular diagnostics lab showed that automation of tasks was largely restricted to physical processes, with interpretive and post-analytical phases still performed manually. Every step in the molecular diagnostics pipeline, ranging from accessioning through report generation, was evaluated for the extent of automation and AI- readiness based on observed practices.

Manual sample accessioning was carried out by human staff, with ID verification and data entry done manually. Nucleic acid extraction was automated on KingFisher Flex instruments and manual in some cases. PCR setup and thermal cycling were automated according to standard protocols. Curve analysis and qPCR and melting data interpretation, however, were done manually by trained staff.

Additionally, closed-system molecular platforms such as BioFire FilmArray and Cepheid GeneXpert were observed to be in routine use for infectious disease detection. These platforms offer end-to-end automation — from nucleic acid extraction and amplification to real-time detection and result output — and are designed to minimize user input, reduce contamination risk, and deliver rapid syndromic panel-based results.

Data conversion of next-generation sequencing (NGS) from .fastq to .bam and .vcf files was performed using command-line bioinformatics tools. No Graphical User Interface (GUI) or pipeline-based automated systems were utilized. Variant analysis comprised the opening of .vcf files in genome browsers, where analysts manually scrutinized regions of interest for mutations. Sanger sequencing data was evaluated by visually inspecting electropherograms. Manual drafting and formatting were used in report generation, and interpretation and communication of patient reports were completely clinician-driven.tegrity.

Table 1: Classification of Workflow Tasks Based on Observed Practice and AI-Readiness

Workflow Step	Current Status	AI Integration Potential	Classification
Sample Accessioning	Manual	Moderate	AI–Human Collaborative
Nucleic Acid Extraction	Manual and Automated (KingFisher Flex)	Low	Automation-Ready
PCR Setup and Thermocycling	Automated	Low	Automation-Ready
qPCR Curve Analysis	Manual	High	AI–Human Collaborative
Melt Curve Evaluation	Manual	High	AI–Human Collaborative
BioFire FilmArray System and Cepheid GeneXpert System	Fully automated, closed system	Low (no interpretive flexibility)	Automation-Ready
Electropherogram Review	Manual	High	AI–Human Collaborative
NGS File Conversion	Manual (Linux CLI)	High	Automation-Ready
Variant Review (.vcf)	Manual	High	AI–Human Collaborative
Report Drafting	Manual	Moderate	AI–Human Collaborative
Final Interpretation	Manual	Low	Human-Exclusive
Report Communication	Manual	Low	Human-Exclusive

4. Discussion

The evidence presented here has an apparent dichotomy of processes in molecular diagnostics: whereas processes upstream such as extraction and thermocycling have been consistently automated, downstream processes—particularly data interpretation processes—remain largely dependent on manual processes. The gradient of automation causes operational inefficiencies, especially under high-volume diagnostic loads, and presents an opportunity yet to be tapped for artificial intelligence (AI) integration.

Automation-enabling tasks such as nucleic acid extraction and NGS file conversion were rule- based, procedural, and standardized. They were automatable since they are of low cognitive complexity and high repeatability. For instance, while KingFisher Flex automated extraction worked well, downstream file conversions (e.g., .fastq to .vcf) had to remain executed on the command line using Linux tools. Literature has indicated that such formatting pipelines may be automated securely with graphical interfaces or cloud computing without data integrity loss . The middle steps would reduce processing time, remove repetitive functions for bioinformatics personnel, and make workflow more streamlined.

The majority of the functions were in the collaborative zone of AI-human, indicating high potential for assistive AI without usurping human judgment. These comprised qPCR curve analysis, melt curve interpretation, review of electropherograms, prioritization of variants, and report writing. In all of these, the data is patterned and arranged machine learning-ready but final interpretation must be rendered contextual by trained experts. For example, although AI models can detect amplification curve anomalies or flag bad-quality melt profiles as issues, assay validity determination is still dependent on sample history and control interpretation. Similarly, in variant review, pathogenicity scores are supplied by DeepVariant or VarSome, which need to be weighed against clinical history and well-established gene–phenotype correlations [4] [13].

The lack of AI in such collaboration spaces is most interesting given that the most common diseases COVID-19 and cancer are already imaging-based diseases where AI is ubiquitous. But in the same labs diagnosing the same diseases at the molecular level, tasks like .vcf review, gene panel validation, and electropherogram interpretation remain manual. This highlights a very critical translational gap: AI has reached disease level but not the support workflow thereof.

Human-exclusive tasks such as final diagnostic interpretation and communication must remain in human hands. These include clinical judgment, ethical responsibility, and nuanced discrimination that current AI programs are not programmed or certified to perform. In addition, communicating results to clinicians and patients requires interpreting uncertainty, connecting meaning, and often dealing with emotional responses areas where empathy and context are significant.

Ethically, AI adoption in common spaces needs to include explainable AI (XAI) and human-in- the-loop (HITL) philosophies [3]. XAI is a system of AI that provides understandable, comprehensible explanations for its output so that human subject-matter experts are aware of how a decision is reached [2]. This is the opposite of black-box models, which can indeed make accurate predictions but without transparency or visibility so that users cannot ascertain reliability or the why behind it. In molecular diagnostics, where the conclusions are a matter of life and death, excessive dependence on opaque black-box systems would erode trust and jeopardize clinical accountability. Therefore, interpretable and transparent AI models are essential to provide efficiency with safety.

Our tri-zonal model of classification offers a functional blueprint for AI incorporation into molecular diagnostics. Rather than demanding outright automation, it demands the hybrid approach: automate the routine, supplement the complex, and protect the interpretive. Not only does this model enhance efficiency and turnaround time, but also the validity of life-critical decisions.

5. Conclusion

Our analysis of current workflows reveals a clear mismatch between where automation has advanced and where it's still lagging in molecular diagnostics. While tasks like nucleic acid extraction and thermal cycling are well-covered by automated systems, much of the downstream work especially data interpretation, variant review, and report generation is still done manually. This slows things down, puts added pressure on staff, and makes it harder to scale up testing when volumes rise.

To address this, we introduced a tri-zonal framework that groups tasks into three categories: those ready for full automation, those that could benefit from AI working alongside humans, and those that should remain in human hands entirely. The goal isn’t to replace expertise, but to support it especially for the repetitive, pattern-heavy parts of the job.

Looking forward, there’s real potential to build tools that can flag issues in qPCR curves, pre- screen variants, convert sequencing files with minimal input, or even help draft reports. If these are developed thoughtfully integrated into existing lab systems and designed to explain their reasoning they could make a big difference. Hence, interpretation and final decision-making should still rest with trained professionals.

Molecular diagnostics is growing fast, especially in areas like cancer, genetic diseases, and infectious disease. If we want labs to keep up without compromising quality, smart, explainable automation isn’t just helpful it’s becoming essential.

REFERENCES

1. Abdallah, S., Sharifa, M., Almadhoun, M. K. I., Khawar Sr, M. M., Shaikh, U., Balabel, K. M., ... & Oyelaja, O. T. (2023). The impact of artificial intelligence on optimizing diagnosis and treatment plans for rare genetic disorders. Cureus, 15(10), e46860. doi: 10.7759/cureus.46860

2. Biswas, A. A. (2024). A Comprehensive Review of Explainable AI for Disease Diagnosis. Array,100345. https://doi.org/10.1016/j.array.2024.100345

3. Budd, S., Robinson, E. C., & Kainz, B. (2019). A survey on active learning and human- in-the-loop deep learning for medical image analysis. Medical Image Analysis, 71, 102062. https://doi.org/10.1016/j.media.2021.102062

4. Dias, R., & Torkamani, A. (2019). Artificial intelligence in clinical and genomic diagnostics. Genome medicine, 11(1), 70. https://doi.org/10.1186/s13073-019-0689-8

5. Huang, S., Yang, J., Fong, S., & Zhao, Q. (2021). Artificial intelligence in the diagnosis of COVID-19: challenges and perspectives. International journal of biological sciences, 17(6), 1581. doi: 10.7150/ijbs.58855

6. Ilhan, B., Guneri, P., & Wilder-Smith, P. (2021). The contribution of artificial intelligence to reducing the diagnostic delay in oral cancer. Oral oncology, 116, 105254. doi: 10.1016/j.oraloncology.2021.105254

7. Malik, Y. S., Sircar, S., Bhat, S., Ansari, M. I., Pande, T., Kumar, P., ... & Dhama, K. (2021). How artificial intelligence may help the Covid‐19 pandemic: Pitfalls and lessons for the future. Reviews in medical virology, 31(5), 1-11. doi: 10.1002/rmv.2205

8. Mirbabaie, M., Stieglitz, S., & Frick, N. R. (2021). Artificial intelligence in disease diagnostics: A critical review and classification on the current state of research guiding future direction. Health and Technology, 11(4), 693-731. https://doi.org/10.1007/s12553-021-00555-5

9. O'Leary, T. J., O'Leary, B. J., & O'Leary, D. P. (2025). A Perspective on Artificial Intelligence for Molecular Pathologists. The Journal of Molecular Diagnostics: JMD, 27(5), 323. doi: 10.1016/j.jmoldx.2025.01.005

10. Prentzas, N., Kakas, A., & Pattichis, C. S. (2023). Explainable AI applications in the medical domain: A systematic review. arXiv preprint arXiv:2308.05411. https://doi.org/10.48550/arXiv.2308.05411

11. Pushkaran, A. C., & Arabi, A. A. (2024). From understanding diseases to drug design: can artificial intelligence bridge the gap? Artificial Intelligence Review, 57(4), 86. https://doi.org/10.1007/s10462-024-10714-5

12. Shafi, S., & Parwani, A. V. (2023). Artificial intelligence in diagnostic pathology. Diagnostic pathology, 18(1), 109.

13. Steiner, D. F., Chen, P. H. C., & Mermel, C. H. (2021). Closing the translation gap: AI applications in digital pathology. Biochimica et Biophysica Acta (BBA)-Reviews on Cancer, 1875(1), 188452. https://doi.org/10.1016/j.bbcan.2020.188452

14. Sun, J., Peng, L., Li, T., Adila, D., Zaiman, Z., Melton-Meaux, G. B., ... & Tignanelli, C. J. (2022). Performance of a chest radiograph ai diagnostic tool for covid-19: A prospective observational study. Radiology: Artificial Intelligence, 4(4), e210217. https://doi.org/10.1148/ryai.210217

15. Tiwari, A., Mishra, S., & Kuo, T. R. (2025). Current AI technologies in cancer diagnostics and treatment. Molecular Cancer, 24(1), 1-41. https://doi.org/10.1186/s12943-025-02369-9

16. Visibelli, A., Roncaglia, B., Spiga, O., & Santucci, A. (2023). The impact of artificial intelligence in the odyssey of rare diseases. Biomedicines, 11(3), 887. https://doi.org/10.3390/biomedicines11030887

17. Zöller, N., Berger, J., Lin, I., Fu, N., Komarneni, J., Barabucci, G., ... & Herzog, S. M. (2025). Human–AI collectives most accurately diagnose clinical vignettes. Proceedings of the National Academy of Sciences, 122(24), e2426153122. https://doi.org/10.1073/pnas.2426153122