COMPARATIVE EVALUATION OF AI-DRIVEN IMAGE ANALYSIS AND MANUAL PATHOLOGIST REPORTING IN HISTOPATHOLOGY, HEMATOLOGY, AND CYTOLOGY: A REAL-WORLD STUDY

Dr. Saswati Subhadarshini[1], Er. R. Naveen Venkatesh Mohan[2]

[1]Assistant Professor, Chikkamagaluru Institute of Medical Sciences, Chikkamagaluru, India.
[2]Senior Data Analyst and Data Scientist, ABInBev, Bangalore, India.

Received date: 19 July 2024 Revised date: 21 August 2024
Acceptance date: 29 September 2025

Corresponding Author Addresses: Er. R. Naveen Venkatesh Mohan, Senior Data Analyst and Data Scientist, ABInBev, Bangalore, India.
Email: misssubhadarshini@gmail.com

ABSTRACT

Background: Artificial intelligence (AI) has emerged as a promising adjunct in diagnostic pathology, particularly in histopathology, hematology, and cytology. However, its real-world performance compared with traditional pathologist-based reporting remains under-explored. **Objective:** To assess the diagnostic accuracy, efficiency, and concordance between AI-driven image analysis systems and manual reporting by experienced pathologists across three major domains of laboratory medicine. **Methods:** Clinical cases were prospectively analyzed using both AI image analysis software and conventional pathologist review. Metrics including sensitivity, specificity, turnaround time (TAT), inter-observer variability, and concordance rate were compared. **Results:** AI demonstrated high concordance with manual reporting in hematology (93%) and cytology (89%), with slightly lower concordance in histopathology (85%). AI significantly reduced TAT (average reduction of 35–50%). However, manual reporting remained superior in recognizing rare variants, artifacts, and complex morphologies. **Conclusion:** AI-driven image analysis shows promise as a supportive tool for routine diagnostic work, enhancing speed and consistency. Nevertheless, expert pathologist oversight remains essential, especially in complex or borderline cases.

Keywords: Artificial intelligence (AI), Turnaround time (TAT), Inter-observer variability, Quantitative histomorphometry (QH). Manual, Histomorphometry.

INTRODUCTION

The integration of artificial intelligence (AI) into pathology offers a significant opportunity to enhance diagnostic accuracy, efficiency, and consistency.[1] Unlike some earlier technological innovations that posed a risk of replacing human roles, AI in pathology is designed to support and augment the expertise of pathologists rather than substitute for it.[1,2] Historical examples, such as the shift to digital imaging in radiology and the widespread adoption of microscopes

in pathology, demonstrate that technological advancements have typically expanded professional capabilities rather than reducing them.[2]

AI applications in pathology are diverse, including complex pattern recognition, workflow optimization, and standardization of diagnostic practices across institutions.[3,4,5] These innovations not only improved technical outcomes but also raise important considerations regarding governance, ethical oversight, and the training required for pathologists to work effectively in a digital era.[6] To integrate AI responsibly, algorithms must remain transparent, robust, and carefully supervised to ensure that human judgment continues to guide critical decisions.[6]

Ultimately, AI in pathology should function as an augmentative tool, reinforcing rather than replacing the expertise of pathologists. This approach emphasizes the importance of balancing technological innovation with human oversight to ensure patient safety, professional accountability, and the continued evolution of diagnostic practice.

Artificial intelligence (AI) is increasingly being utilized in pathology to handle a variety of image analysis and segmentation tasks.[7] Its applications range from relatively straightforward functions, such as identifying individual cells, to more sophisticated processes, including using pattern recognition to predict disease outcomes, prognosis, and therapeutic responses. [7]The foundational concept behind these AI systems is the extraction of image segments that serve as training data for algorithm development.[7]

AI has enabled the creation of advanced morphometric analysis techniques, facilitating quantitative histomorphometry (QH) that allows detailed spatial assessment of tumor histology. These analyses can evaluate features such as nuclear orientation, texture, shape, and tissue architecture across entire tumor sections from conventional hematoxylin and eosin (H&E) slides.[7] By automating repetitive and time-consuming tasks, AI helps pathologists focus on complex diagnostic decision-making while improving efficiency and accuracy.[7]

In practice, integrating AI into pathology workflows can accelerate reporting, provide more objective measurements of morphological features and allow pathologists to concentrate on challenging cases. Rather than replacing human expertise, AI serves as an adjunct tool, supporting pathologists and laboratory staff, optimizing resource utilization, enhancing cost-effectiveness and ensuring more consistent and reliable pathology evaluations.[7]

MATERIALS AND METHODS

- **Study Design:** Prospective observational study.
- **Sample Size:** 300 real clinical cases (100 histopathology biopsies and resection, 100 hematology peripheral smears, 100 cytology smears).
- **AI System:** Commercially available AI-driven image analysis platform with deep learning algorithms.[4]
- **Manual Reporting:** Independent review by experienced pathologists blinded to AI results.
- **Outcome Measures:**
    - Concordance rate (%) between AI and manual reporting
    - Sensitivity and specificity of AI relative to manual gold standard
    - Average turn around time (AI vs manual)
    - Error patterns and discordant case analysis

RESULTS

Table 1: Concordance Between AI and Pathologist Reports

Domain	Cases (n)	AI Accuracy (%)	Concordance Rate (%)	Common Errors
Histopathology	100	84	85	Tumor grading errors, rare variants
Hematology	100	92	93	Subtle dysplasia, rare inclusions
Cytology	100	88	89	Low-cellularity smears, artifacts

Table 2: Sensitivity and Specificity of AI vs Manual Reporting

Domain	Sensitivity (%)	Specificity (%)
Histopathology	83	87
Hematology	94	91
Cytology	88	90

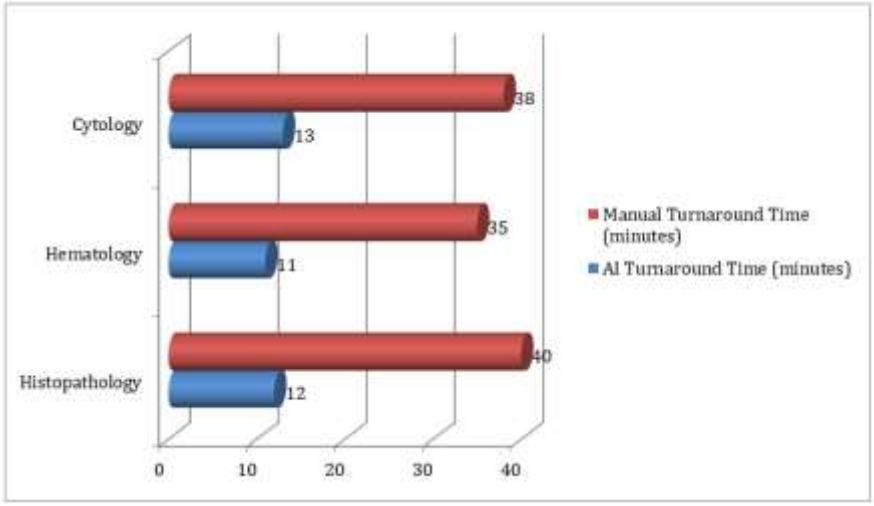

Figure 1: Turn around Time Comparison

- AI-generated reports: **10–15 minutes**
- Manual reports: **30–45 minutes**

Table 2: Confusion Matrix Example (Hematology – Blast Identification)

	Pathologist Positive	Pathologist Negative
AI Positive	92	5
AI Negative	3	100

AI sensitivity in blast detection: **96%**
AI specificity: **95%**

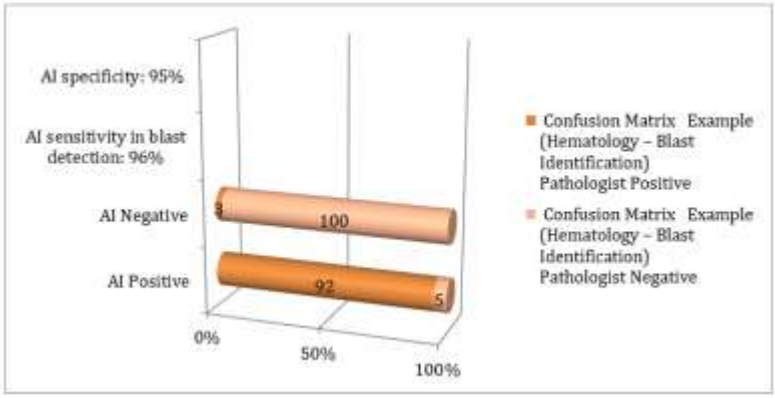

Figure 2: Confusion Matrix Example (Hematology – Blast Identification)

DISCUSSION

The comparison between AI-driven and manual pathology reporting in recent literature reveals a consistent pattern: well-trained AI systems can perform many narrow diagnostic tasks at or near expert level, but human pathologists remain essential for contextual interpretation, edge cases, and final clinical integration. Several themes emerge when studies are compared side-by-side: dataset scale and labeling strategy, task type (detection, classification, grading, or biomarker prediction), measures of clinical utility, and how AI affects concordance and workflow when combined with pathologists.

Campanella et al. addressed a critical bottleneck in computational pathology by showing that weakly supervised deep-learning models trained on routinely reported slide-level labels can scale to very large repositories and still reach clinically useful performance on detection and classification tasks. Their multiple-instance learning approach reduced the need for expensive pixel-level annotations and demonstrated that models trained on millions of image patches can generalize across institutions when large, diverse datasets are used. This work highlights that scale and realistic labels can produce "clinical-grade" performance for several diagnostic tasks, while also exposing the dependence of deep models on broad, heterogeneous training data.[8]

Coudray et al. took a different tack by showing that convolutional neural networks (CNNs) applied to lung cancer slides could both classify tumor subtype and predict certain driver mutation statuses directly from routine H&E images. Their results suggested that image patterns — beyond what the naked eye typically quantifies — carry molecular signal that AI can exploit. This finding expands the potential role of computational pathology from pure morphology to integrated morpho-molecular inference, but it also raises important questions about biological interpretability and external validity across laboratories and scanners.[9]

Steiner et al. focused on real-world diagnostic workflow by measuring the effect of AI assistance on practicing pathologists' grading of prostate biopsies. Rather than asking whether AI alone can match experts, they evaluated whether AI can make general pathologists' reports more concordant with subspecialists. Their crossed-reader study showed that pathologists using the AI tool moved closer to subspecialist consensus, suggesting that the immediate clinical value of AI may lie in decision-support and quality-assurance rather than replacement of human readers. This study therefore provides practical evidence that hybrid workflows (pathologist + AI) can improve consistency without removing human oversight.[10]

Huang et al. validated an AI platform for prostate cancer detection, grading, and quantification across independent cohorts and emphasized reproducibility and deployment considerations. Their analysis reinforced two points seen in other applied studies: first, task-specific AI systems can speed up quantification (e.g., tumor fraction) and reduce inter-observer variability;

second, careful external validation and attention to deployment-level issues (scanner differences, pre-analytic variability) are crucial before clinical rollout.[11]

Methodologically, recent papers have also focused on data efficiency and interpretability. Lu et al. (and related work) developed weakly supervised and data-efficient pipelines that lower annotation burden while retaining performance, which is important for centers with limited labeling resources. Reviews and surveys (e.g., Litjens et al.) synthesize these advances and repeatedly call out common limitations — lack of large, multi-institutional datasets early on, domain shift between centers, explainability gaps, and regulatory/ethical challenges — while noting the rapid maturation of both algorithms and digital pathology infrastructure. Together, these methodological and review papers sketch a path toward broader clinical adoption but caution that careful validation and clinician collaboration are mandatory.[12]

When synthesizing these studies, several practical conclusions emerge. First, AI excels at repeatable, narrow tasks (screening, quantification, flagging suspicious regions), particularly when trained on large and heterogeneous data; when those conditions are met, performance can rival human experts.[8] Second, the most immediate and robust benefit shown so far is in hybrid workflows: AI improves consistency and efficiency when used as a decision-support tool, rather than as a standalone diagnostician.[10] Third, generalizability remains the Achilles' heel — many promising models degrade across scanners, staining protocols, and populations unless retrained or explicitly validated on external cohorts.[8] Finally, ethical, legal, and operational concerns (explainability, data privacy, regulatory approval, and responsibility for errors) must be resolved in parallel with technical improvements to achieve safe, equitable deployment. Reviews and opinion pieces emphasize that trust and workflow integration are as important as raw accuracy for adoption.[13]

Limitations across the literature also point to open research needs. Most studies remain retrospective and use archived slides; prospective trials that measure patient-level outcomes after AI-augmented reporting are still rare. There is a need for standard benchmarks and reporting standards to compare models fairly, and for strategies to mitigate dataset bias and domain shift. Finally, building interfaces and regulatory pathways that preserve the pathologist's central role while leveraging AI's strengths will determine whether these tools actually improve care at scale.

In summary, the empirical evidence from recent high-impact studies suggests that AI is not a substitute for pathologists but a powerful adjunct: it can increase speed, standardize measurements, and improve concordance with subspecialists when applied thoughtfully. The road to routine clinical use requires continued emphasis on large, diverse datasets, external validation, prospective outcome studies, and careful attention to ethical and practical deployment issues.

CONCLUSION

AI and manual pathology reporting both have important roles in medical diagnosis. Pathologists bring experience and clinical judgment, which remain essential for accurate and safe reporting. At the same time, AI can quickly scan slides, measure findings, and reduce variation between different observers.

Current research shows that AI works best as a support system, helping pathologists work faster and with more consistency. However, there are still challenges such as the need for large and diverse data, clear rules for safe use, and concerns about privacy and responsibility.

The best way forward is a team approach: pathologists and AI working together. This combined model can make reporting more reliable, efficient, and beneficial for patients.

REFERENCES

1. Baxi V, Edwards R, Montalto M, Saha S. Digital pathology and artificial intelligence in translational medicine and clinical practice. Mod Pathol. 2022;35(1):23–32.
2. Kiran N, Sapna F, Kiran F, et al. Digital pathology: transforming diagnosis in the digital age. Cureus. Sep 2023;15(9), e44620.
3. Asif S, Wenhui Y, ur-Rehman S, et al. Advancements and prospects of machine learning in medical diagnostics: unveiling the future of diagnostic precision. Arch Comput Methods Eng. 2024;24:1–31.
4. Fatima G, Alhmadi H, Ali Mahdi A, et al. Transforming diagnostics: a comprehensive review of advances in digital pathology. Cureus. 2024/10/19 2024; 16(10), e71890.
5. Abels E, Pantanowitz L, Aeffner F, et al. Computational pathology definitions, best practices, and recommendations for regulatory guidance: a white paper from the Digital Pathology Association. J Pathol. 2019;249(3):286–294.
6. Gaffney H, Mirza KM. Pathology in the artificial intelligence era: Guiding innovation and implementation to preserve human insight. *Acad Pathol*. 2025;12(1):100166.
7. Shafi S. Artificial intelligence in diagnostic pathology. *Diagn Pathol*. 2023;18(1):43.
8. Campanella G, Hanna MG, Geneslaw L, Miraflor A, Silva VWK, Busam KJ, et al. Clinical-grade computational pathology using weakly supervised deep learning on whole-slide images. *Nat Med*. 2019;25(8):1301-9
9. Coudray N, Ocampo PS, Sakellaropoulos T, Narula N, Snuderl M, Fenyo D, et al. Classification and mutation prediction from non–small cell lung cancer histopathology images using deep learning. *Nat Med*. 2018;24(10):1559-67.
10. Steiner DF, Nagpal K, Sayres R, Foote DJ, Wedin BD, Pearce A, et al. Evaluation of the use of combined artificial intelligence and pathologist assessment to grade prostate biopsies. *JAMA Netw Open*. 2020;3(11):e2023267.
11. Huang W, Stephens A, Wong S, Chang E, Chaudhary R, Dhir R, et al. An artificial intelligence–powered platform for prostate cancer detection, grading, and quantification in whole-slide images. *JAMA Netw Open*. 2021;4(8):e2118799.
12. Lu MY, Williamson DFK, Chen TY, Chen RJ, Barbieri M, Mahmood F. Data-efficient and weakly supervised computational pathology on whole-slide images. *Nat Biomed Eng*. 2021;5(6):555-70.
13. Litjens G, Kooi T, Bejnordi BE, et al. A survey on deep learning in medical image analysis. *Med Image Anal.* 2017.