Content available at: https://www.ipinnovative.com/open-access-journals

IP Journal of Diagnostic Pathology and Oncology

Journal homepage: https://jdpo.org/

Short Communication

Big data and the rise of AI-driven pathology

Indranil Chakrabarti[1]*, Sujaya Mazumder[1]

[1]Dept. of Pathology and Laboratory Medicine, All India Institute of Medical Sciences, Kalyani, West Bengal, India

Abstract

Diagnostic pathology has significantly evolved beyond its traditional reliance on optical microscopy. The integration of Big Data, artificial intelligence (AI), and digital tools is transforming the field by enabling comprehensive analysis of digital images, genomic information, and clinical data simultaneously. This advancement is particularly impactful in oncology, where pathomics—the intersection of image analysis and data science—extracts detailed molecular and clinical patterns from extensive collections of high-resolution slides and multi-omic datasets. AI has transitioned from a conceptual term to a practical instrument capable of automating complex tasks such as tumor grading and mutation prediction, which improves diagnostic accuracy and supports tailored therapeutic strategies. The infrastructure of whole-slide imaging coupled with cloud storage facilitates efficient data management and global collaboration. However, challenges remain regarding platform interoperability, data security, and clinical validation of AI workflows. Ethical considerations require addressing biases in AI training datasets, safeguarding patient privacy, and ensuring equitable access to these technologies across diverse healthcare settings. Big Data analytics enhance clinicians' ability to predict tumor behavior and treatment response by integrating various biological datasets. Consequently, the role of the pathologist has expanded to encompass combining clinical expertise with data science, interpreting complex information to inform patient management. This collaboration between expert human judgment and advanced computational methods marks a shift from descriptive pathology toward predictive diagnostics. The ongoing challenge lies in managing the increasing volume of data while maintaining a focus on improving disease understanding and patient outcomes.

Keywords: Big data, Artificial intelligence, Digital pathology, Precision oncology

Received: 21-09-2025; **Accepted:** 24-10-2025; **Available Online**: 29-10-2025

This is an Open Access (OA) journal, and articles are distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 License, which allows others to remix, tweak, and build upon the work non-commercially, as long as appropriate credit is given and the new creations are licensed under the identical terms.

For reprints contact: reprint@ipinnovative.com

1. Introduction

Diagnostic pathology has undergone profound evolution. Historically, pathologists relied predominantly on direct microscopic examination to interpret tissue specimens. Presently, the discipline is experiencing an informatics revolution whereby digital imaging, genomic profiling, and clinical datasets converge, fundamentally transforming diagnostic and therapeutic paradigms. This shift represents not merely incremental progress but a redefinition of disease conceptualization, particularly within oncology. The landscape of diagnostic pathology is being fundamentally transformed by the confluence of Big Data, artificial intelligence (AI), and digital technologies.[1,2]

2. Big Data in Pathology: A New Kind of Evidence

Pathology laboratories now generate an immense volume of data - the "Big Data" in Pathology - encompassing high-resolution digital whole-slide images, comprehensive genomic, transcriptomic, proteomic, and metabolomic profiles, immunophenotyping results, and extensive metadata drawn from electronic health records. The emerging field of pathomics—combining image analytics with data science—deconstructs these voluminous datasets to identify patterns linking tissue morphology and molecular as well as clinical attributes.[3]

These expansive digital repositories have facilitated the identification of histopathologic features previously indiscernible via conventional methods. Advanced deep

*Corresponding author: Indranil Chakrabarti
Email: drinch@rediffmail.com

https://doi.org/10.18231/j.jdpo.28467.1761719813
© 2025 The Author(s), Published by Innovative Publications.

learning architectures, notably convolutional neural networks, emulate human visual recognition while providing unparalleled scalability and fatigue resistance.[3,4]

3. AI: Turning Data into Action

Artificial intelligence is the missing link between this ocean of raw data and actual, useful answers. By training on huge collections of annotated slides and molecular data, AI models can automate tumor grading, count mitoses, and even predict genetic mutations. AI can now pull predictive signals right out of standard H&E slides, flagging spatial details tied to the tumor environment and how a patient might respond to treatment.[2,5]

Presently, the discipline is witnessing the advent of "foundation models," large, multimodal AI frameworks that integrate imaging, genomic, and clinical data streams. These innovations constitute a leap comparable to the early 2000s genomics revolution. Beyond expediting laboratory workflows, these systems empower pathologists to identify the pathology with heightened precision, facilitating personalized therapeutic selection.

The advent of generative AI models, akin to ChatGPT, has similarly inspired innovative applications within pathology.

Some of the notable models include:[2]

1. UNI /UNI2: Trained on over 200 million images derived from 350,000 whole-slide images, excelling in classification tasks
2. CONCH (Contrastive Learning from Captions for Histopathology): A multimodal model utilizing histopathology images and annotated medical text, efficacious in cancer subtyping and descriptive image captioning.
3. GigaPath (Microsoft): Developed using more than 170,000 slides sourced from 28 cancer centers across the United States
4. mSTAR (for Multimodal Self-taught Pretraining): Integrates gene expression, imaging data, and textual information.

AI Chatbots for Pathologists are now available to streamline the work for pathologists

1. PathChat: Integrates the UNI model with a language processing module and has been granted US FDA breakthrough device designation. It supports interactive image analysis and report generation via a conversational interface, enabling pathologists to engage dynamically with complex pathology datasets.
2. SmartPath: An advanced, end-to-end AI pathology system for cancer diagnosis, integrating morphology, molecular data, and clinical context. SmartPath-R1, a more recent iteration, adds reasoning capability through a multimodal large language model (MLLM) structure.

These AI assistants demonstrate advanced capabilities including autonomous planning, decision-making, and operational actions.[2]

4. Digital Pathology: Building the Data Backbone

The efficacy of AI-driven pathology relies fundamentally on digital slide acquisition. Whole-slide imaging complemented by cloud-based storage and computational resources establishes the infrastructure for AI deployment, telepathology, and international collaboration. The imperative to establish digital and AI-enabled pathology platforms (DAIP) is evident.[6]

By mid-2025, over 50 AI diagnostic tools have obtained regulatory clearance (CE-IVD or CE-IVDR), enabling clinical implementation. These systems automate diagnostic tasks such as HER2 immunohistochemical scoring, Gleason grading, and detection of lymph node metastases.[7] Nonetheless, challenges persist, including ensuring interoperability across heterogeneous systems, data security, and rigorous validation of algorithmic accuracy. The standardization, safeguarding, and verification of these mechanisms are essential to cultivate clinical trust.

5. Hurdles and Ethics: The Other Side of the Coin

The transition to this digital pathology era presents logistical and ethical complexities. Incompatibilities among digital platforms and inherent biases in training datasets risk perpetuating disparities in diagnosis. Moreover, managing extensive patient data necessitates stringent attention to privacy, data security, and equitable access. Addressing these issues requires the establishment of transparent regulatory frameworks for AI, international data-sharing agreements, and the inclusion of under-resourced regions.[1,6,7] Additional barriers, such as model generalizability, explainability ("black box" phenomena), legal and ethical implications, hidden costs, and implementation hurdles, remain to be resolved for widespread adoption.

6. Big Data and the Push for Precision Oncology

Precision oncology depends on connecting the dots—bringing together genetic, clinical, and tissue data to tailor treatments. Big Data analytics now enable the prediction of tumor behavior and treatment response, in some instances preceding the initiation of therapy. Studies indicate that AI-driven multi-omic analyses yield novel biomarkers and assist clinicians in stratifying patients according to risk, particularly in malignancies with complex biology such as breast and lung cancers. Established platforms including The Cancer Genome Atlas (TCGA), METABRIC, and ClinOmicsTrail exemplify the utility of combining clinical outcomes with gene expression and tissue features to optimize personalized cancer care.[8-11]

7. The Evolving Role of Pathologists: Beyond Microscopy

Modern pathologists assume roles extending well beyond conventional slide review. They function as data interpreters orchestrating the synthesis of quantitative image data and molecular findings to generate clinically meaningful insights. This interdisciplinary expertise encompasses in-depth comprehension of tissue architecture, sophisticated computational modeling, and molecular diagnostics. Consequently, the contemporary diagnostic pathologist emerges as a hybrid professional, adept in both biological sciences and data analytics, instrumental in advancing precision medicine and oncology.

8. The Next Frontier

Looking forward, Big Data has redefined pathology as a data-centric discipline rather than solely an image-based practice. The conventional microscope remains indispensable; however, the integration with machine learning tools that identify subtle morphological and molecular patterns augments diagnostic capacity. This symbiosis transforms pathology from a descriptive to a predictive science. The critical challenge lies in harnessing this exponential data growth without compromising the humanistic mission—to deepen disease understanding, enhance therapeutic interventions, and ultimately improve patient outcomes.

9. Conflict of Interest

None.

References

1. Yao L, Liu Y, Wang T, Han C, Li Q, Li Q, et al. Global trends of big data analytics in health research: a bibliometric study. *Front Med (Lausanne).* 2025;12:1456286. https://doi.org/10.3389/fmed.2025.1456286
2. Kwon D. How artificial intelligence is transforming pathology. *Nature*. 2025;641(1342):1342–4. doi:10.1038/d41586-025-01576-0.[cited 2025 Sept 3]. Available from: https://www.nature.com/articles/d41586-025-01576-0.
3. Ahuja S, Zaheer S. Advancements in pathology: Digital transformation, precision medicine, and beyond. *J Pathol Inform*. 2024;16:100408. doi: 10.1016/j.jpi.2024.100408.
4. Research and Markets. After the scan – The 'big data' vision of digital pathology to 2025 [Internet]. Medical Buyer; 2020 Nov 17 [cited 2025 Sept 4]. Available from: https://medicalbuyer.co.in/after-the-scan-the-big-data-vision-of-digital-pathology-to-2025/
5. Faber A. Digital pathology and AI highlights from ASCO 2025 [Internet]. Philadelphia: Proscia; 2025 Jun 16 [cited 2025 Sept 3]. Available from: https://proscia.com/digital-pathology-and-ai-highlights-from-asco-2025/
6. Shen IZ, Zhang L. Digital and Artificial Intelligence-based Pathology: Not for Every Laboratory – A Mini-review on the Benefits and Pitfalls of Its Implementation. *J Clin Transl Pathol*. 2025;5(2):79–85. doi: 10.14218/jctp.2025.00007
7. Bruns V, Benz M. Status and trends in digital pathology [Internet]. ESMO Daily Reporter; 2025 Oct 13 [cited 2025 Oct 15]. Available from: https://dailyreporter.esmo.org/news/status-and-trends-in-digital-pathology
8. Del Valle AC. Leveraging digital pathology and AI to transform clinical diagnosis in developing countries. *Front Med (Lausanne).* 2025 Sep 25;12:1657679. https://doi.org/10.3389/fmed.2025.1657679
9. Zhang H, Hussin H, Hoh CC, Cheong SH, Lee WK, Yahaya BH. Big data in breast cancer: Towards precision treatment. *Digit Health*. 2024;10. doi: 10.1177/20552076241293695
10. Nygren P. Precision cancer medicine 2025: some concerns. *Acta Oncol*. 2025;64:44604. doi: 10.2340/1651-226X.2025.44604.
11. Chang TG, Park S, Schäffer AA, Jiang P, Ruppin E. Hallmarks of artificial intelligence contributions to precision oncology. *Nat Cancer.* 2025;6(3):417–31. doi: 10.1038/s43018-025-00917-2

Cite this article: Chakrabarti I, Mazumder S. Big data and the rise of AI-driven pathology. *IP J Diagn Pathol Oncol*. 2025;10(3):144-146.