



Mini Review



Digital and Artificial Intelligence-based Pathology: Not for Every Laboratory – A Mini-review on the Benefits and Pitfalls of Its Implementation

Iris Z. Shen¹ and Lanjing Zhang^{2,3,4*}

¹The Winsor School, Boston, MA, USA; ²Department of Chemical Biology, Ernest Mario School of Pharmacy, Rutgers University, Piscataway, NJ, USA; ³Department of Pathology, Princeton Medical Center, Plainsboro, NJ, USA; ⁴Rutgers Cancer Institute of New Jersey, New Brunswick, NJ, USA

Received: January 23, 2025 | Revised: March 19, 2025 | Accepted: March 20, 2025 | Published online: April 03, 2025

Abstract

Background and objectives: With the increasing use of artificial intelligence (AI) in diagnostics, AI algorithms have shown great potential in aiding diagnostics. As more of these algorithms are developed, there is overwhelming enthusiasm for implementing digital and artificial intelligence-based pathology (DAIP), but doubts and pitfalls are also emerging. However, few original or review articles address the limitations and practical aspects of implementing DAIP. In this review, we briefly examine the evidence related to the benefits and pitfalls of DAIP implementation and argue that DAIP is not suitable for every clinical laboratory. **Methods:** We searched the PubMed database using the following keywords: “digital pathology,” “digital AI pathology,” and “AI pathology.” Additionally, we incorporated personal experiences and manually searched related papers. **Results:** Ninety-two publications were found, of which 24 met the inclusion criteria. Many advantages of DAIP were discussed, including improved diagnostic accuracy and equity. However, several limitations of implementing DAIP exist, such as financial constraints, technical challenges, and legal/ethical concerns. **Conclusions:** We found a generally favorable but cautious outlook for the implementation of DAIP in the pathology workflow. Many studies have reported promising outcomes in using AI for diagnosis and analysis; however, there are also several noteworthy limitations in implementing DAIP. Therefore, a balance between the benefits and pitfalls of DAIP must be thoroughly articulated and examined in light of the institution’s needs and goals before making the decision to implement DAIP. Approaches for mitigating machine learning biases were also proposed, and the adaptation and growth of the pathology profession were discussed in light of DAIP development and advances.

Citation of this article: Shen IZ, Zhang L. Digital and Artificial Intelligence-based Pathology: Not for Every Laboratory – A Mini-review on the Benefits and Pitfalls of Its Implemen-

tation. J Clin Transl Pathol 2025;5(2):79–85. doi: 10.14218/JCTP.2025.00007.

Introduction

Digital and artificial intelligence-based pathology (DAIP) has grown rapidly in recent years.¹ This is largely attributable to advances in omics and machine learning algorithms,^{2–4} and the approval of whole-slide imaging equipment by the U.S. Food and Drug Administration.¹ Many institutions rightfully anticipate and prepare for the widespread adoption of DAIP in clinical practice, as its many advantages have been discussed, recognized, and accepted.^{5–12} However, few original or review articles address the limitations and practical aspects of implementing DAIP. It is noteworthy that artificial intelligence (AI)-based pathology is not the best fit for every clinical laboratory, as there are certain constraints and hurdles associated with its implementation and use, such as economic issues. We thus conducted a focused narrative review to discuss the benefits and pitfalls of implementing DAIP in clinical laboratories based on published works. We will also provide future directions on how DAIP should be used.

Search strategy and results

We searched the PubMed database with the keywords “digital pathology,” “implementation,” and “artificial intelligence” in October 2024 and again in January 2025. No time limit was set for the search. The search revealed 188 potentially relevant articles (Fig. 1).¹³ Filtering criteria were then applied, including article type (clinical study, clinical trial, meta-analysis, observational study, randomized controlled trial, review, and systematic review), article language (English), duplication, and species (human). After filtering and manual addition, 113 results were assessed for inclusion eligibility. We manually assessed each article for its relevance to DAIP implementation and included only those that fell within the scope of this review on DAIP implementation. A total of 53 articles met the inclusion criteria, and 39 additional relevant articles were manually included (Fig. 1).¹³ Therefore, 92 articles were reviewed and summarized in this review.

Keywords: Digital pathology; Artificial intelligence; Implementation; Costs; Bias; Medical legal; Large language models; LLMs.

*Correspondence to: Lanjing Zhang, Department of Pathology, Princeton Medical Center, Plainsboro, NJ 08536, USA. ORCID: <https://orcid.org/0000-0001-5436-887X>. Tel: +1-609-853-6833, Fax: +1-609-853-6841, E-mail: lanjing.zhang@rutgers.edu

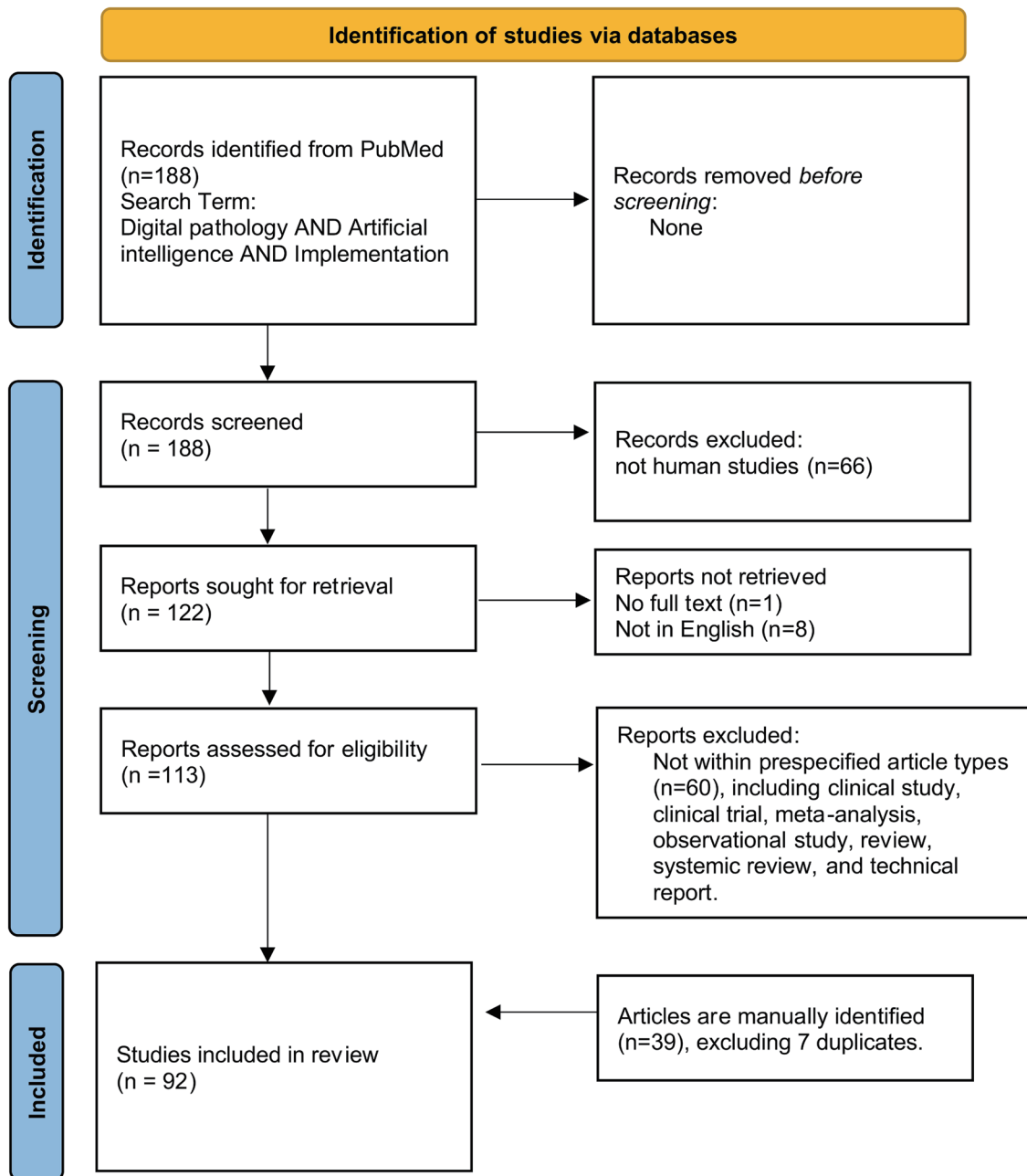


Fig. 1. Flowchart of the literature research using the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines.

Benefits of DAIP

Diagnostic accuracy and efficiency

While all types of AI, including those for pathological analysis purposes, are still relatively new and require further development, the use of DAIP demonstrates promising potential.^{1,5–12} DAIP diagnoses are faster and have reduced turnaround times due to the use of technology.¹⁴ Another study showed that, in a diagnostic setting, the accuracy rate of AI algorithms was 99% in identifying breast cancer metastasis in lymph nodes, compared with 81% for pathologists.¹⁵ These results were confirmed by other studies and in other cancers, including colon cancer, head and neck cancer,

and melanoma metastases.⁵ Other studies have shown that AI algorithms exhibit high sensitivity and low error rates in determining whether a disease is benign or malignant. Furthermore, AI algorithms can incorporate human feedback to improve classification performance and reduce the number of slides needed for retraining ("Human-in-the-loop").¹¹

DAIP can also improve pathology workflow and pathologists' efficiency.^{1,16,17} For example, DAIP can help quickly identify areas of interest on slides and assess markers for immunotherapies.^{18–20} It can also facilitate workload (re)allocation among pathologists.¹¹ Moreover, integration of DAIP with electronic medical record systems will make information retrieval much faster than using glass slides and reduce the

step of checking patient identification.^{1,11} Finally, a DAIP-enabled copilot may assist pathologists with differential diagnoses and in ordering the appropriate immunostains.²¹

In addition to image analysis, some large language models (LLMs) can help predict results based on clinical data (input). For example, an machine learning (ML) model named “survival quilts” can use the SEER database to predict cancer-specific mortality in patients with non-metastatic prostate cancers.⁶ LLMs can be continuously trained and learn from the data they are fed, suggesting that their predictive capabilities can improve over time as more data is provided.²² This adaptability proves that LLMs used for diagnostic and predictive purposes can remain relevant as information and practices evolve.

Specifically, LLMs can help extract important data from cancer reports to generate synoptic reports according to cancer protocols.^{23–25} For researchers, LLMs can be used to extract relevant information from free-text pathology reports for annotation and data curation.²³ They can also be used to produce patient-centered pathology reports, helping to better inform and educate patient.²⁶ The segmentation of unstructured pathology reports will become possible using LLMs, which will help accurately extract critical clinical information and classify nodal statuses.^{24,27} Thus, these applications of LLMs can reduce workloads and improve work efficiency for clinicians and researchers.

Convenience and accessibility

DAIP is particularly attractive to clinical laboratories due to its convenience and accessibility. Because of its digital interface, collaborations between pathologists are made much more accessible. Pathologists can share cases (i.e., digital slides) remotely and collaborate with others more effectively. This advantage also extends to work schedule flexibility. For example, pathologists who wish to work remotely can have the opportunity to do so. DAIP also promises greater productivity in the workplace, as algorithms are quicker to diagnose and do not experience fatigue like humans do. Due to this reduced turnaround time, AI algorithms are useful for triaging cases. Indeed, AI algorithms can quickly diagnose cases that require a fast turnaround, such as lymph node metastasis in tumors.²⁸

Equity and labor shortage

DAIP has the potential to bridge the gaps between smaller (often less-funded) and larger institutions, as well as between rural and urban locations. If implemented, remote areas with fewer staff or less specialized staff would be able to access the same quality of diagnostic services as less remote places with more staff or specialized staff.

An example of digital pathology (DP) benefiting rural areas is demonstrated through the Eastern Quebec Telepathology Network, an extensive implementation of DP starting in 2011, which encompasses 22 hospitals serving around 1.7 million patients.¹¹ A 2018 study showed a reduction in two-stage surgeries and patient transfers to more urban centers. Service breaks and diagnostic delays were also reduced.¹¹

In addition to benefiting rural and smaller institutions, DAIP may also help alleviate the overall labor shortage in the pathology workforce. Due to increased cancer incidence, an aging population, and increasingly complex cancer diagnoses, there is a growing demand for pathologists. However, the number of practicing pathologists is decreasing and is projected to shrink by 20% over the next two decades.¹⁶ Therefore, AI can reduce the burden on pathologists by accelerating the rate of certain tasks.¹⁶

Cost

A study demonstrated that when Memorial Sloan Kettering Cancer Center’s pathology department in the U.S. implemented DP in 2015, there was a 93% decrease in glass slide requests, and the use of whole-slide imaging (WSI) increased to around 23,000 slides per month by 2017.^{11,17} Their estimated savings through the use of DP was around \$267,000 per year, and taking implementation costs (WSI setup and maintenance) into account, the estimated year of breaking even would be 2021 (seven years after implementation).^{11,17} However, our literature search did not identify any follow-up publications regarding their anticipated savings. Additional cost savings are made possible through a customizable calculator developed by a working group of 30 experts.²⁹

Pitfalls of DAIP

Legal and ethical concerns

The pitfalls of DAIP pathology are tied to its reliance on AI technology. An AI-based diagnosis is only as good as the data it is trained on (“garbage in, garbage out”). Many of these AI algorithms are new and still in the early stages of development and application, meaning they will improve over time as they are used more frequently.¹⁵ In addition, AI performance may be biased or unfair to certain populations due to the inherent bias in the datasets used for training.^{30,31} This ethical concern stems from biases in the healthcare system. A study found that using a deep learning-based computational pathology system with common modeling approaches, performance gaps were observed between White and Black patients, with differences of 3.0% for breast cancer subtyping, 10.9% for lung cancer subtyping, and 16.0% for IDH1 mutation prediction in gliomas.³¹ The AI itself is not inherently biased, but if it is only fed biased data, its performance will reflect that bias. Therefore, addressing biases in AI will be difficult, if not impossible, as it may require a drastic change in practices related to data collection, modeling, and AI application. Nonetheless, addressing the AI fairness issue must be the starting point.

In addition to the risk of complacency and perpetuation of disparities, another major concern of AI in pathology is that it could reduce the pathology job market and the number of pathologists needed. Given its demonstrated accuracy in previous studies, AI may play an increasingly important role in diagnostic pathology, as it does not become fatigued, demand salary increases, or need breaks. However, pathologists also contribute significantly to clinical activities beyond diagnostics. They actively participate in multidisciplinary tumor boards and can solve complicated and rare diseases that ML models cannot reliably be trained on. Furthermore, pathologists play a vital role in consulting patients and clinicians on complicated or ethically challenging cases, which DAIP cannot do. Lastly, the interventional procedures performed by pathologists, such as fine needle aspiration, cannot be replaced by DAIP.

Because AI is still relatively new, it lacks federal or state regulation. Therefore, there are numerous ethical and legal concerns associated with its use.^{9,22,32} It is particularly concerning that we do not know who should be held primarily responsible when a DAIP algorithm makes a wrong diagnosis or influences a pathologist to make an incorrect decision. Similarly, who will be responsible when a DAIP algorithm makes a correct diagnosis, but the pathologist incorrectly overrides it? A more interesting question is whether and when pathologists will be able to trust AI algorithms more than themselves (and their peer pathologists) as AI improves.

Unintended and hidden costs

Theoretically, DAIP can save money by increasing productivity and containing labor costs. However, several costs were often overlooked. For example, the guidelines of the College of American Pathologists require keeping the glass slides for at least 10 years. Therefore, digital storage costs should be considered in addition to physical storage.¹¹ Moreover, digital storage is nowadays billed at a preset interval (commonly per year) and will incur costs perpetually unless the institution decides otherwise. Furthermore, the capital costs of implementing DAIP may be prohibitive or burdensome for a small practice or hospital with a limited capital budget. Should a small practice or hospital use bonds or other long-term financial instruments to afford the DAIP equipment, it may take decades to pay off the costs, and by then, the equipment may have lost maintenance support.

Linked to the potential medicolegal concerns, pathologists using DAIP may face higher malpractice insurance premiums than those who do not. The costs of cybersecurity and computing hardware should also be considered, along with the labor costs of additional quality assurance programs. Additionally, the labor costs involved in conducting cost-analysis for implementing DAIP should be taken into account. Therefore, many unintended and hidden costs are associated with implementing DAIP.

Challenges in digital implementation

While equity issues can be addressed through the implementation of DAIP, some places do not even have the funds to fully implement it. There are many steps in the implementation process, including institutional approval, cost analysis, procurement of AI tools (which requires time and money for training and research), configuration, and adoption.

In addition to the numerous implementation steps, DP has many requirements for its maintenance and use. It needs sufficient data storage capacity, reliable networks (connectivity), and high-resolution scanners.²⁸ WSI requires dedicated infrastructure and information technology support to minimize downtime. Specifically, because WSI allows for remote viewing, certain latency and bandwidth requirements must be met (usually more than 10 megabits per second). Additionally, depending on the institution, color calibration may be necessary for optimal image analysis.¹¹ The use of DAIP also requires lab technicians to modify standard processes, such as performing thinner sections, placing sections closer together for optimal scanning ($n = 15$, 20.8%), and dividing large specimens ($n = 14$, 58.1%).³³ Moreover, if the manufacturer of the WSI equipment has not validated the slides, institutions must validate them themselves with a certain number of cases (CAP guidelines require a minimum of 60 slides for hematoxylin and eosin staining and at least 20 slides for each auxiliary technique).¹¹ In addition to the extra labor costs and time, images are slow to load and difficult to view on digital slides. Because of these issues, an Asian study group had to hire technicians to help ameliorate the technological difficulties.²⁸

Specifically, the pure cost of DAIP combined with its tricky implementation comes at a significant price. A study focusing on the implementation of DAIP in labs in Europe and Asia found that the majority (63%) of the surveyed labs spent between \$100,000 and \$1,000,000 on implementation. However, some institutions spent between \$1,000,000 and \$5,000,000, and one institution (2.2%) spent more than \$5,000,000.³³ It is also important to note that these numbers are likely to be higher in the United States due to inflation and the higher cost of labor.

Challenges in AI implementation

There are two main challenges in the implementation and application of AI algorithms in the DP workflow: the generalizability of the model and the explainability of the model.²⁰

Generalizability refers to how well the model's complexity matches the complexity of the data. Overfitting is one of the generalizability problems. For example, a model trained on many specific cases and slides may struggle with another set of specific cases. The biggest challenge for algorithms is applying what they were trained on to new datasets.²⁰

Explainability, or interpretability, involves an algorithm's ability to allow users to understand the factors that lead to a specific decision.²⁰ DAIP faces a common issue in AI known as the "black box" problem, where the system's internal decision-making process cannot be traced. This issue makes it difficult for DAIP users to analyze each decision made by the AI, and therefore they may not be able to provide the best feedback to improve the AI algorithm.³⁴ Transparency is imperative in larger institutional settings, where failure to detect bias could have enormous ethical and legal consequences.

Technological limitations of LLMs

While DAIP has the potential to revolutionize the pathology field, it still has many technological limitations. So far, DAIP programs have mostly focused on image analysis and have not taken advantage of the rise of LLMs. For example, a pathologist can use an LLM such as ChatGPT to ask general questions about a case that image analysis cannot address. However, LLMs still lack a detailed understanding of specific cases, making them less flexible than image analysis technologies.²² A recent study also showed that ChatGPT was helpful in identifying differential diagnoses in 62.2% of simulated cases, but it provided at least one erratic differential diagnosis in 3.7% of the simulated cases.³⁵ Strikingly, among the 214 references ChatGPT provided, 12.1% were irrelevant or inaccurate, and 17.8% were non-existent.³⁵ Therefore, careful review and thorough validation are required before fully adopting LLMs in diagnostic pathology.

Mitigation approach and future directions

While using DAIP undoubtedly signifies advancements in technology, it is essential to critically assess whether it truly meets the unique needs and goals of individual laboratories or institutions. Given the costs and legal/ethical concerns associated with implementing DAIP, manufacturers of DAIP software and equipment should work to minimize implementation costs. Additionally, we must recognize the need to optimize DAIP implementation, including reducing storage costs, improving access to experts and expert opinions, encouraging institutional administration support for installation, and replacing microscopes. In 2020, the primary adopters of DAIP were still large institutions, while smaller and more rural laboratories could also benefit significantly from DAIP implementation.³³

AI holds great promise, especially in anatomic pathology. However, storage remains a significant concern.^{29,33} For example, one hematoxylin and eosin slide can take up to 10 gigapixels of information at 40x magnification, which presents challenges for pathologists to notice and account for all the information—not just on one slide but across many. AI algorithms are able to produce consistent results when provided with the same slides and can identify anomalies even within large volumes of data.

Despite AI's promising effects on the pathology workflow,

one gap that remains is the “translation gap”, where challenges arise in implementing these algorithms in practice. For instance, variations in tissue acquisition and slide preparation processes may influence the performance of image analyses.¹⁶ Another challenge is selecting the most reliable performance metric, as accuracy, balanced accuracy, area under the receiver-operator curve, sensitivity, and specificity are all important. The College of American Pathologists has acknowledged this issue and has led the development of recommendations for ML performance,³⁶ although further validation and research are still needed.

Several approaches can mitigate ML biases or fairness issues, and future work on this subject is warranted. First, radiology colleagues have addressed three key components for reducing biases in ML: rigorous dataset curation and data handling, development of robust ML models, and a better understanding and utilization of various performance metrics.^{37–39} Their experiences may be helpful in mitigating biases and fairness issues in DAIP. Second, regrouping study subjects by the variable of interest could help reduce ML biases and improve AI fairness when classifying cancer outcomes.⁴⁰ Third, newer ML algorithms, such as transfer learning, fairness-aware classifiers, and unbiased prompts for LLMs,^{30,41,42} may help reduce ML biases in DAIP implementation. Finally, pathologists and software engineers can only mitigate ML biases if they are aware of them, highlighting the need for related education. Caution should also be exercised when DAIP biases are more prevalent or subtle.

AI and the pathology profession will grow together, influencing each other as AI continues to advance and become increasingly embedded in pathology and other medical practices.⁹ AI has already been found useful in pathology and medical education and can help identify reporting errors and discrepancies.^{43,44} Clearly, pathologists need more education on DAIP technology and its applications.²⁸ Residency programs may formally incorporate DAIP into their curriculum as a required or optional rotation in the future, focusing on general and practical subjects. Pathologists who specialize in DAIP will be needed in the near future to bridge AI/data science and pathology practice. These specialists may be trained through designated fellowship programs alongside pathology informaticians but will have significant differences in their deeper knowledge of AI/data science. However, it remains unclear whether their fellowship should be separate from or integrated with current pathology informatics programs.

We anticipate that pathologists will continue to play an important and integrated role in patient care. The non-diagnostic roles of pathologists are unlikely to be replaced by DAIP, as described earlier. However, DAIP will reshape the pathology profession in several ways. First, pathology reports may become more patient-centered and more understandable to laypersons.²⁶ Pathologists may then have direct interactions with patients, aided by LLMs (e.g., copilot).⁴⁵ Second, DAIP’s assistance will enable the efficient handling of many simpler cases,^{9,32,46} while pathologists will be able to devote more time to complex cases or direct patient care. This may challenge the value of general surgical pathologists but could be justified in small institutions. Third, seeking second opinions will become much easier, faster, and cheaper on the DAIP platform compared to using glass slides. Many pathologists could work remotely as consultants. Finally, pathologists will need to work hard to “outsmart” DAIP and AI and demonstrate their undeniable clinical value, especially since DAIP may appear more accurate than pathologists in diagnosing certain diseases.^{5,14,15} One example is melanoma diagnosis. Although dermoscopy and digital dermoscopy were first pro-

posed in 2009,⁴⁷ recent studies have shown the possibility of dermatologist-like AI tools for diagnosing melanoma.⁴⁸

LLMs

In addition to image analysis, a future direction that institutions should consider is incorporating LLMs alongside image analysis in pathology. Paired with image analysis, LLMs could prove to be extremely helpful to pathologists, but they also come with their limitations. LLMs are not yet fine-tuned enough for pathological purposes, often providing broad and generic responses. They also present unique challenges—ChatGPT-4, for example, has been known to hallucinate by generating fabricated sources when asked about histopathology, a problem not seen in digital pathology-specific AI. However, there is progress with more domain-specific LLMs, such as BERT (from Google) derivatives (BERT (Bi-directional Encoder Representations from Transformers)) and GPT (from OpenAI) derivatives (Generative Pre-trained Transformer), demonstrating significant advancements in a very new field.⁴⁹ While LLMs are probably not ready for wide clinical laboratory application yet, they have great potential for streamlining and organizing fields such as public health and pathology.

For this direction with LLMs to be fruitful, four main improvements are essential: collaboration between AI developers and institutions, minimization of bias through training on how to recognize biases, more longitudinal studies on AI use in laboratories across the world, and clear regulations governing AI at every scale (local, regional, and global).²² Without further understanding, research, and data assessing DAIP and LLMs, the research and clinical fields may not be able to fully take advantage of these risky yet potentially valuable tools.

Conclusions

Our literature search shows that implementing DAIP is not suitable for every clinical laboratory, and its benefits and pitfalls must be thoroughly evaluated. AI demonstrates the potential to make a pathologist’s and institution’s workflow more efficient and accessible, but there are many other considerations to take into account, including ethical and legal issues as well as cost restrictions. Therefore, we are still in the early phases of implementing DAIP. It is probably too early to fully implement DAIP in every clinical laboratory. In summary, while many studies have found promising outcomes in using AI for diagnosis and analysis, there are also several noteworthy limitations in implementing DAIP. Therefore, a balance between the benefits and pitfalls of DAIP must be carefully examined and aligned with the institution’s own needs and goals before deciding whether to implement DAIP.

Acknowledgments

None.

Funding

This work was supported by the National Cancer Institute, National Institutes of Health (R37CA277812 to LZ).

Conflict of interest

Lanijing Zhang is a deputy editor-in-chief of the *Journal of Clinical and Translational Pathology*. The authors declare no other conflicts of interest.

Author contributions

Database query, literature review, and manuscript writing (IZS, LS). All authors have approved the final version and publication of the manuscript.

References

- [1] Zhang DY, Venkat A, Khasawneh H, Sali R, Zhang V, Pei Z. Implementation of Digital Pathology and Artificial Intelligence in Routine Pathology Practice. *Lab Invest* 2024;104(9):102111. doi:10.1016/j.labinv.2024.102111, PMID:39053633.
- [2] Liu DD, Zhang L. Trends in the characteristics of human functional genomic data on the gene expression omnibus, 2001-2017. *Lab Invest* 2019;99(1):118-127. doi:10.1038/s41374-018-0125-5, PMID:30206311.
- [3] Cui M, Deng F, Disis ML, Cheng C, Zhang L. Advances in the Clinical Application of High-throughput Proteomics. *Explor Res Hypothesis Med* 2024;9(3):209-220. doi:10.14218/erhm.2024.00006, PMID:39148720.
- [4] Cui M, Cheng C, Zhang L. High-throughput proteomics: a methodological mini-review. *Lab Invest* 2022;102(11):1170-1181. doi:10.1038/s41374-022-00830-7, PMID:35922478.
- [5] van Diest PJ, Flach RN, van Dooijeweert C, Makineli S, Breimer GE, Stathonikos N, et al. Pros and cons of artificial intelligence implementation in diagnostic pathology. *Histopathology* 2024;84(6):924-934. doi:10.1111/his.15153, PMID:38433288.
- [6] Satturwar S, Parwani AV. Artificial Intelligence-Enabled Prostate Cancer Diagnosis and Prognosis: Current State and Future Implications. *Adv Anat Pathol* 2024;31(2):136-144. doi:10.1097/PAP.0000000000000425, PMID:38179884.
- [7] Kim D, Sundling KE, Virk R, Thrall MJ, Alperstein S, Bui MM, et al. Digital cytology part 1: digital cytology implementation for practice: a concept paper with review and recommendations from the American Society of Cytopathology Digital Cytology Task Force. *J Am Soc Cytopathol* 2024;13(2):86-96. doi:10.1016/j.jasc.2023.11.006, PMID:38158316.
- [8] Kalra N, Verma P, Verma S. Advancements in AI based healthcare techniques with FOCUS ON diagnostic techniques. *Comput Biol Med* 2024;179:108917. doi:10.1016/j.combiomed.2024.108917, PMID:39059212.
- [9] Berbis MA, McClintock DS, Bychkov A, Van der Laak J, Pantanowitz L, Lennerz JK, et al. Computational pathology in 2030: a Delphi study forecasting the role of AI in pathology within the next decade. *EBioMedicine* 2023;88:104427. doi:10.1016/j.ebiom.2022.104427, PMID:36603288.
- [10] Janowczyk A, Baumhoer D, Dirnhofer S, Grobholz R, Kipar A, de Leval L, et al. Towards a national strategy for digital pathology in Switzerland. *Virchows Arch* 2022;481(4):647-652. doi:10.1007/s00428-022-03345-0, PMID:35622144.
- [11] Jahn SW, Plass M, Moiraf F. Digital Pathology: Advantages, Limitations and Emerging Perspectives. *J Clin Med* 2020;9(11):3697. doi:10.3390/jcm9113697, PMID:33217963.
- [12] Chong Y, Kim DC, Jung CK, Kim DC, Song SY, Joo HJ, et al. Recommendations for pathologic practice using digital pathology: consensus report of the Korean Society of Pathologists. *J Pathol Transl Med* 2020;54(6):437-452. doi:10.4132/jptm.2020.08.27, PMID:33027850.
- [13] Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* 2021;372:n71. doi:10.1136/bmj.n71, PMID:33782057.
- [14] Försch S, Klauschen F, Hufnagl P, Roth W. Artificial Intelligence in Pathology. *Dtsch Arztebl Int* 2021;118(12):194-204. doi:10.3238/arztebl.m2021.0011, PMID:34024323.
- [15] Hanna MG, Reuter VE, Samboy J, England C, Corsale L, Fine SW, et al. Implementation of Digital Pathology Offers Clinical and Operational Increase in Efficiency and Cost Savings. *Arch Pathol Lab Med* 2019;143(12):1545-1555. doi:10.5858/arpa.2018-0514-OA, PMID:31173528.
- [16] Bruce C, Prassas I, Mokhtar M, Clarke B, Youssef E, Wang C, et al. Transforming diagnostics: The implementation of digital pathology in clinical laboratories. *Histopathology* 2024;85(2):207-214. doi:10.1111/his.15178, PMID:38516992.
- [17] Liu Y, Kohlberger T, Norouzi M, Dahl GE, Smith JL, Mohtashamian A, et al. Artificial Intelligence-Based Breast Cancer Nodal Metastasis Detection: Insights Into the Black Box for Pathologists. *Arch Pathol Lab Med* 2019;143(7):859-868. doi:10.5858/arpa.2018-0147-OA, PMID:30295070.
- [18] Steiner DF, Chen PC, Mermel CH. Closing the translation gap: AI applications in digital pathology. *Biochim Biophys Acta Rev Cancer* 2021;1875(1):188452. doi:10.1016/j.bbcan.2020.188452, PMID:33065195.
- [19] Steiner DF, MacDonald R, Liu Y, Truszowski P, Hipp JD, Gammage C, et al. Impact of Deep Learning Assistance on the Histopathologic Review of Lymph Nodes for Metastatic Breast Cancer. *Am J Surg Pathol* 2018;42(12):1636-1646. doi:10.1097/PAS.0000000000001151, PMID:30312179.
- [20] Thagaard J, Broeckx G, Page DB, Jahangir CA, Verbandt S, Kos Z, et al. Pitfalls in machine learning-based assessment of tumor-infiltrating lymphocytes in breast cancer: A report of the International Immuno-Oncology Biomarker Working Group on Breast Cancer. *J Pathol* 2023;260(5):498-513. doi:10.1002/path.6155, PMID:37608772.
- [21] Sobhani F, Robinson R, Hamidinekoo A, Roxanis I, Somaiah N, Yuan Y. Artificial intelligence and digital pathology: Opportunities and implications for immuno-oncology. *Biochim Biophys Acta Rev Cancer* 2021;1875(2):188520. doi:10.1016/j.bbcan.2021.188520, PMID:33561505.
- [22] Lu MY, Chen B, Williamson DFK, Chen RJ, Zhao M, Chow AK, et al. A multimodal generative AI copilot for human pathology. *Nature* 2024;634(8033):466-473. doi:10.1038/s41586-024-07618-3, PMID:38866050.
- [23] Ullah E, Parwani A, Baig MM, Singh R. Challenges and barriers of using large language models (LLM) such as ChatGPT for diagnostic medicine with a focus on digital pathology - a recent scoping review. *Diagn Pathol* 2024;19(1):43. doi:10.1186/s13000-024-01464-7, PMID:38414074.
- [24] Shahid F, Hsu MH, Chang YC, Jian WS. Using Generative AI to Extract Structured Information from Free Text Pathology Reports. *J Med Syst* 2024;49(1):36. doi:10.1007/s10916-025-02167-2, PMID:40080229.
- [25] Fung D, Arbour G, Malik K, Muzio K, Ng R. Using a Longformer Large Language Model for Segmenting Unstructured Cancer Pathology Reports. *JCO Clin Cancer Inform* 2025;9:e2400143. doi:10.1200/CCI-24-00143, PMID:40036729.
- [26] Park S, Choi W, Choi I. Enhancing Clinical Data Extraction from Pathology Reports: A Comparative Analysis of Large Language Models. *Stud Health Technol Inform* 2024;316:756-760. doi:10.3233/SHTI240523, PMID:39176904.
- [27] Proctor ES, Nusbaum DJ, Lee JM, Benirschke RC, Freedman A, Raster G, et al. Bridging the gap: Evaluating ChatGPT-generated, personalized, patient-centered prostate biopsy reports. *Am J Clin Pathol* 2025. doi:10.1093/ajcp/aae185, PMID:39838829.
- [28] Kim S, Jang S, Kim B, Sunwoo L, Kim S, Chung JH, et al. Automated Pathologic TN Classification Prediction and Rationale Generation From Lung Cancer Surgical Pathology Reports Using a Large Language Model Fine-Tuned With Chain-of-Thought: Algorithm Development and Validation Study. *JMIR Med Inform* 2024;12:e67056. doi:10.2196/67056, PMID:39705675.
- [29] Ardon O, Asa SL, Lloyd MC, Lujan G, Parwani A, Santa-Rosario JC, et al. Understanding the financial aspects of digital pathology: A dynamic customizable return on investment calculator for informed decision-making. *J Pathol Inform* 2024;15:100376. doi:10.1016/j.jpi.2024.100376, PMID:38736870.
- [30] Li Y, Zhang L, Zhang Y. Fairness of chatgpt. *arXiv* 2023. doi:10.48550/arXiv.2305.18569.
- [31] Vaidya A, Chen RJ, Williamson DFK, Song AH, Jaume G, Yang Y, et al. Demographic bias in misdiagnosis by computational pathology models. *Nat Med* 2024;30(4):1174-1190. doi:10.1038/s41591-024-02885-z, PMID:38641744.
- [32] Verghese G, Lennerz JK, Ruta D, Ng W, Thavaraj S, Siziopikou KP, et al. Computational pathology in cancer diagnosis, prognosis, and prediction - present day and prospects. *J Pathol* 2023;260(5):551-563. doi:10.1002/path.6163, PMID:37580849.
- [33] Pinto DG, Bychkov A, Tsuyama N, Fukuoka J, Eloy C. Real-World Implementation of Digital Pathology: Results From an Intercontinental Survey. *Lab Invest* 2023;103(12):100261. doi:10.1016/j.labinv.2023.100261, PMID:37839634.
- [34] Evans T. The explainability paradox: Challenges for xAI in digital pathology. *Future Gener Comput Syst* 2022;133:281-296. doi:10.1016/j.future.2022.03.009.
- [35] Guastafierro V, Corbitt DN, Bressan A, Fernandes B, Mintemur Ö, Magnoli F, et al. Unveiling the risks of ChatGPT in diagnostic surgical pathology. *Virchows Arch* 2024. doi:10.1007/s00428-024-03918-1, PMID:39269615.
- [36] Hanna MG, Olson NH, Zarella M, Dash RC, Herrmann MD, Furtado LV, et al. Recommendations for Performance Evaluation of Machine Learning in Pathology: A Concept Paper From the College of American Pathologists. *Arch Pathol Lab Med* 2024;148(10):e335-e361. doi:10.5858/arpa.2023-0042-CP, PMID:38041522.
- [37] Zhang K, Khosravi B, Vahdati S, Faghani S, Nugen F, Rassoulnejad-Mousavi SM, et al. Mitigating Bias in Radiology Machine Learning: 2. Model Development. *Radiol Artif Intell* 2022;4(5):e220010. doi:10.1148/ryai.220010, PMID:36204532.
- [38] Rouzrokh P, Khosravi B, Faghani S, Moassefi M, Vera Garcia DV, Singh Y, et al. Mitigating Bias in Radiology Machine Learning: 1. Data Handling. *Radiol Artif Intell* 2022;4(5):e210290. doi:10.1148/ryai.210290, PMID:36204544.
- [39] Faghani S, Khosravi B, Zhang K, Moassefi M, Jagtap JM, Nugen F, et al. Mitigating Bias in Radiology Machine Learning: 3. Performance Metrics. *Radiol Artif Intell* 2022;4(5):e220061. doi:10.1148/ryai.220061, PMID:36204539.
- [40] Feng CH, Deng F, Disis ML, Gao N, Zhang L. Towards machine learning fairness in classifying multicategory causes of deaths in colorectal or lung cancer patients. *bioRxiv* 2025. doi:10.1101/2025.02.14.638368, PMID:40027644.
- [41] Gu T, Pan W, Yu J, Ji G, Meng X, Wang Y, et al. Mitigating bias in AI mortality predictions for minority populations: a transfer learning approach. *BMC Med Inform Decis Mak* 2025;25(1):30. doi:10.1186/s12911-025-02862-7, PMID:39825353.
- [42] Yaseliiani M, Noor-E-Alam M, Hasan MM. Mitigating Sociodemographic Bias in Opioid Use Disorder Prediction: Fairness-Aware Machine Learning Framework. *JMIR AI* 2024;3:e55820. doi:10.2196/55820, PMID:39163597.
- [43] Burisch C, Bellary A, Breuckmann F, Ehlers J, Thal SC, Sellmann T, et al. ChatGPT-4 Performance on German Continuing Medical Education-Friend or Foe (Trick or Treat)? Protocol for a Randomized Controlled Trial. *JMIR Res Protoc* 2025;14:e63887. doi:10.2196/63887, PMID:39913914.
- [44] Bartley M, Huemann Z, Hu J, Tie X, Ross AB, Kennedy T, et al. Artificial Intelligence for Teaching Case Curation: Evaluating Model Performance on Imaging Report Discrepancies. *Acad Radiol* 2025. doi:10.1016/j.acra.2025.02.011, PMID:40058990.
- [45] Wang A, Zhou J, Zhang P, Cao H, Xin H, Xu X, et al. Large language model answers medical questions about standard pathology reports. *Front Med (Lausanne)* 2024;11:1402457. doi:10.3389/fmed.2024.1402457, PMID:39359921.

- [46] Jarkman S, Karlberg M, Pocevičiūtė M, Bodén A, Bándi P, Litjens G, *et al*. Generalization of Deep Learning in Digital Pathology: Experience in Breast Cancer Metastasis Detection. *Cancers (Basel)* 2022;14(21):5424. doi:10.3390/cancers14215424, PMID:36358842.
- [47] Rajpara SM, Botello AP, Townend J, Ormerod AD. Systematic review of dermoscopy and digital dermoscopy/ artificial intelligence for the diagnosis of melanoma. *Br J Dermatol* 2009;161(3):591–604. doi:10.1111/j.1365-2133.2009.09093.x, PMID:19302072.
- [48] Chanda T, Hauser K, Hobelsberger S, Bucher TC, Garcia CN, Wies C, *et al*. Dermatologist-like explainable AI enhances trust and confidence in diagnosing melanoma. *Nat Commun* 2024;15(1):524. doi:10.1038/s41467-023-43095-4, PMID:38225244.
- [49] Lujan G, Quigley JC, Hartman D, Parwani A, Roehmholdt B, Meter BV, *et al*. Dissecting the Business Case for Adoption and Implementation of Digital Pathology: A White Paper from the Digital Pathology Association. *J Pathol Inform* 2021;12:17. doi:10.4103/jpi.jpi_67_20, PMID:34221633.