




REVIEW

Pros and cons of artificial intelligence implementation in diagnostic pathology

Paul J van Diest,^{1,*}  Rachel N Flach,^{1,2,*}  Carmen van Doijeweert,¹ Seher Makineli,^{1,3} Gerben E Breimer,¹ Nikolas Stathonikos,¹  Paul Pham,¹ Tri Q Nguyen¹ & Mitko Veta^{2,4}
¹Department of Pathology, ²Department of Oncological Urology, ³Department of Surgical Oncology, University Medical Center Utrecht, Utrecht and ⁴Department of Biomedical Engineering, Eindhoven University of Technology, Eindhoven, the Netherlands

van Diest P J, Flach R N, van Doijeweert C, Makineli S, Breimer G E, Stathonikos N, Pham P, Nguyen T Q & Veta M

(2024) *Histopathology* 84, 924–934. <https://doi.org/10.1111/his.15153>

Pros and cons of artificial intelligence implementation in diagnostic pathology

The rapid introduction of digital pathology has greatly facilitated development of artificial intelligence (AI) models in pathology that have shown great promise in assisting morphological diagnostics and quantitation of therapeutic targets. We are now at a tipping point where companies have started to bring

algorithms to the market, and questions arise whether the pathology community is ready to implement AI in routine workflow. However, concerns also arise about the use of AI in pathology. This article reviews the pros and cons of introducing AI in diagnostic pathology.

Keywords: algorithms, diagnostics, digital pathology, implementation, keyword artificial intelligence, machine learning

Introduction

Diagnostic surgical- and cytopathology offer excellent value for the money. For an estimated 0.2% of global healthcare costs in developed countries, patients get a cytological or histological diagnosis of their disease, starting with the benign or (pre)malignant nature. For benign processes, pathological evaluation can distinguish between congenital, reactive, degenerative, or inflammatory diseases with leads for treatment. For (pre)malignant diseases, the cancer type, grade, and other prognostic factors are established, offering information about the cancer stage whether the tumour has been radically removed, therapeutic

targets, and whether cancer might be hereditary. Based on all this information, treating clinicians can identify an optimal personalized treatment, including surgery, radiotherapy, systemic therapies like chemotherapy and endocrine therapy, and targeted and immunotherapy.

While pathology diagnoses are generally valuable and reliable, challenges still arise from the visual interpretation of morphological patterns, with the potential of misdiagnosis or reproducibility issues.¹ For example, the grading of breast and prostate cancer is subject to significant inter- and intralaboratory variation, with consequences for treatment.^{2–4} Furthermore, some tasks can be tedious and time-consuming, potentially leading to missed or misinterpreted diagnostic areas. An example of such a task is the assessment of sentinel lymph nodes of breast cancer patients, which in most cases do not contain metastases. Strikingly, in the MIRROR-study, central

Address for correspondence: P J van Diest, Department of Pathology, University Medical Center Utrecht, Utrecht, the Netherlands.
e-mail: p.j.vandiest@umcutrecht.nl

*These authors contributed equally.

pathology review of sentinel lymph nodes in over 1000 breast cancer patients resulted in a change of the N-classification in 24% of patients, mainly upstaging.⁵ Because of the therapeutic consequences of these failures,⁶ we need help to do a better job. Over the last decades, we have fortunately seen the introduction of different technologies to support the diagnostic process, such as electron microscopy, (enzyme) histochemistry, immunohistochemistry, and molecular pathology. In addition, digital pathology has entered many laboratories in the Netherlands, Sweden, the USA, the UK, and many other countries worldwide. More recently, we have seen the introduction of artificial intelligence (AI) techniques in pathology that have shown great promise in assisting morphological diagnostics and quantitation of therapeutic targets.^{7–9}

We are now at a tipping point where companies have started to bring algorithms to the market, and questions arise whether we, the pathology community, are ready to implement AI in our routine workflow. However, concerns also arise about the use of AI in pathology.¹⁰ This article briefly explains how AI works and reviews the current pros and cons of introducing AI in diagnostic pathology.

What is AI, and how does it work?

AI is a branch of computer science dedicated to developing computer models capable of performing tasks typically requiring human-level intelligence, such as perception, reasoning, decision-making, and natural language processing. A critical aspect of AI is its ability to learn from data using mathematical algorithms and statistical models, distinguishing it from traditional software, which relies on explicitly programmed rules and instructions. AI algorithms instead derive rules suited for the tasks they aim to solve directly from available training data. Unlike humans, AI systems are not subject to limitations like fatigue, enabling them to operate objectively and consistently. This makes AI especially valuable in healthcare, where precise and timely decision-making is crucial for patient outcomes and therapeutic decision-making.

AI is an umbrella term for different training modalities for algorithms. Most currently developed algorithms are trained via machine learning, in which computer software learns to identify data patterns in large numbers of example cases, and subsequently match these patterns to new cases. A subset of

machine learning is deep learning, in which software is trained by exposing a multilayered artificial neural network to large datasets, loosely mimicking the hierarchical organization of the human brain. A specific subset of these deep neural networks are convolutional neural networks (CNN), which are trained to recognize data patterns in images.¹¹

In computational pathology, the branch of pathology that involves computational analysis of any kind to analyse patient specimens, AI is at present primarily employed to analyse full digital images of tissue sections and cytological specimens [“whole slide images” (WSI)] to assist pathologists in making more accurate and efficient diagnoses. Additionally, AI can serve as quality control for stains and process nonimage-based modalities like molecular data or multimodal datasets to provide deeper insights into disease processes. By training AI algorithms on extensive datasets of labelled images and other relevant data, patterns, and features indicative of various diseases or conditions can be discerned. This enables algorithms to automatically classify new images and offer suggestions to pathologists, leading to enhanced accuracy and efficiency in diagnosis and prognosis, as well as providing novel insights into the biology of diseases.

Pros of introducing AI in diagnostic pathology

Many studies published over the last 5 years have shown the potential power of AI in diagnostic pathology, algorithms often outperforming pathologists. These were often derived from so-called “challenges” (grand-challenge.org), where international groups compete for making the best algorithm on the same set of online (usually annotated) images. Most AI algorithms are cancer-focused, including cancer detection and grading, predicting prognosis, therapeutic targets and molecular changes, but also for stain quality control and artefact recognition.¹² Besides, some good noncancer algorithms have been developed, like for classifying kidney biopsies.^{13,14} In the following, examples of AI algorithms will be presented that are more sensitive than eyeballing, thereby saving time and costs, or make diagnosis more reproducible for better treatment selection. We both regard histopathology, as well as cytopathology algorithms, which are not as widely developed or tested, yet as the specimens originating in surgical pathology, but yield promising results.

ARTIFICIAL INTELLIGENCE IS MORE SENSITIVE

In some instances, AI on hematoxylin eosin (HE) can be more sensitive than eyeballing by the pathologist. Examples are algorithms for detecting breast cancer metastases in lymph nodes, which had in the first study an accuracy of 99% compared to 81% for pathologists in a diagnostic setting.¹⁵ This was confirmed in follow up studies,^{16,17} and widened to e.g. colon cancer,^{18,19} head and neck cancer,¹⁹ and melanoma metastases.²⁰ This is a critical advantage in reducing errors in diagnosis and classification. Also, for prostate cancer algorithms, sensitivity increases most for discriminating between small foci of atypical glands and tumour, improving sensitivity from 74% without AI to 90% with AI-assistance in a study that focused on the value in small tumour foci in prostate needle biopsies.²¹ In another study, sensitivity in prostate needle biopsies slightly increased from 92.6% in unassisted reviews to 93.9% in assisted reviews, to 96.5 in the AI-standalone performance. Although sensitivity in the AI-standalone performance outperformed the (assisted) pathologists, specificity was highest in an AI-assisted workflow (96.1% in assisted pathologists versus 93.5 in unassisted pathologists versus 92.5% AI standalone).²²

Examples of (relatively) high algorithm sensitivity can also be found in cytopathology, within the screening of cervical intraepithelial lesion (sensitivities ranging between 85.7% and 94.7),²³ lung cancer (sensitivity and specificity 95.9% and 98.2% respectively),²⁴ or urine cytology (sensitivity and specificity 80.9% and 61.8%, compared to histopathological samples, respectively).²⁵

ARTIFICIAL INTELLIGENCE SAVES TIME

Support from algorithms may save the pathologist time, especially when algorithms run automatically in the background, and cases are preprocessed before the pathologist opens up the case, hereby improving the turnaround time. Examples are algorithms for finding mitoses (Figure 1) and lymph node metastases detection,^{15,26,27} and grading of tumours in prostate needle biopsies, and algorithms for stratifying patients into relevant routines for pathologists, thereby optimizing workflow. Studies suggest that diagnostic time can be reduced by 13–65% during the assessment of prostate needle biopsies if a pathologist uses AI during evaluation.^{21,22,28} Another study showed the potential of AI to automatically discriminate between slides with prostate cancer and between WSI of lymph nodes that contained micro- and

macrometastases or benign and normal tissue and thereby correctly signing out 30–40% of the cases as benign without further need of additional immunohistochemistry or human intervention.²⁹ In dermatopathology, stratifying cases into routine or more complex cases can help optimize workflow by reducing double reading and deeper cuts and stains, resulting in a lower turnaround time.^{30,31} In cytopathology, computer-assisted Papanicolaou (Pap) test screening already facilitated faster screening and improved accuracy, by limiting the number of fields of view to cytopathologists.³² The cytopathologists still need to assess these specific cells, but do not have to assess the entire slide.^{23,33}

ARTIFICIAL INTELLIGENCE MAY SAVE COSTS BY REDUCING ANCILLARY TESTING

Algorithms can save costs by decreasing the use of costly ancillary tests, such as immunohistochemistry, but also molecular tests or next-generation sequencing (NGS).

Multiple algorithms are being developed that help predict the presence of a mutation, thereby potentially narrowing the indication for molecular testing, as pretest probability of a molecular feature increases. Examples are detection of Epstein–Barr virus in gastric cancer,³⁴ detection of *BRCA1/2* mutation on HE slides in breast cancer,³⁵ *TP53* in prostate cancer,³⁶ microsatellite instability in colorectal carcinoma,^{37,38} or a pancancer detection algorithm of multiple genetic alterations like *BRAF*, *PIK3CA*, *KRAS*, *TP53*, *FOXA1*, and more.³⁹

Estimating the financial impact of their AI algorithm³⁷ in a clinical context, Kacew *et al.*⁴⁰ modelled an AI-based determination of mismatch repair/microsatellite instability status into a first-line metastatic colorectal carcinoma setting in the US. They developed a model to compare cost savings in eight testing strategies in a hypothetical nationally representative population in the US. These included different strategies containing NGS, a high-sensitivity AI, a high-sensitivity polymerase chain reaction or immunohistochemistry panel, and a high-sensitivity panel. In their analysis, most cost-effectiveness was achieved when using a high-sensitivity AI, followed by an immunohistochemistry panel for those that tested negative using AI (12.9% of costs, \$400 million), when compared to performing NGS alone.

If algorithms can reliably identify cancer cells on HE-stained sections, this may save the high costs of immunohistochemistry, such as keratin stains on sentinel lymph nodes for breast or cervix carcinoma,⁴¹

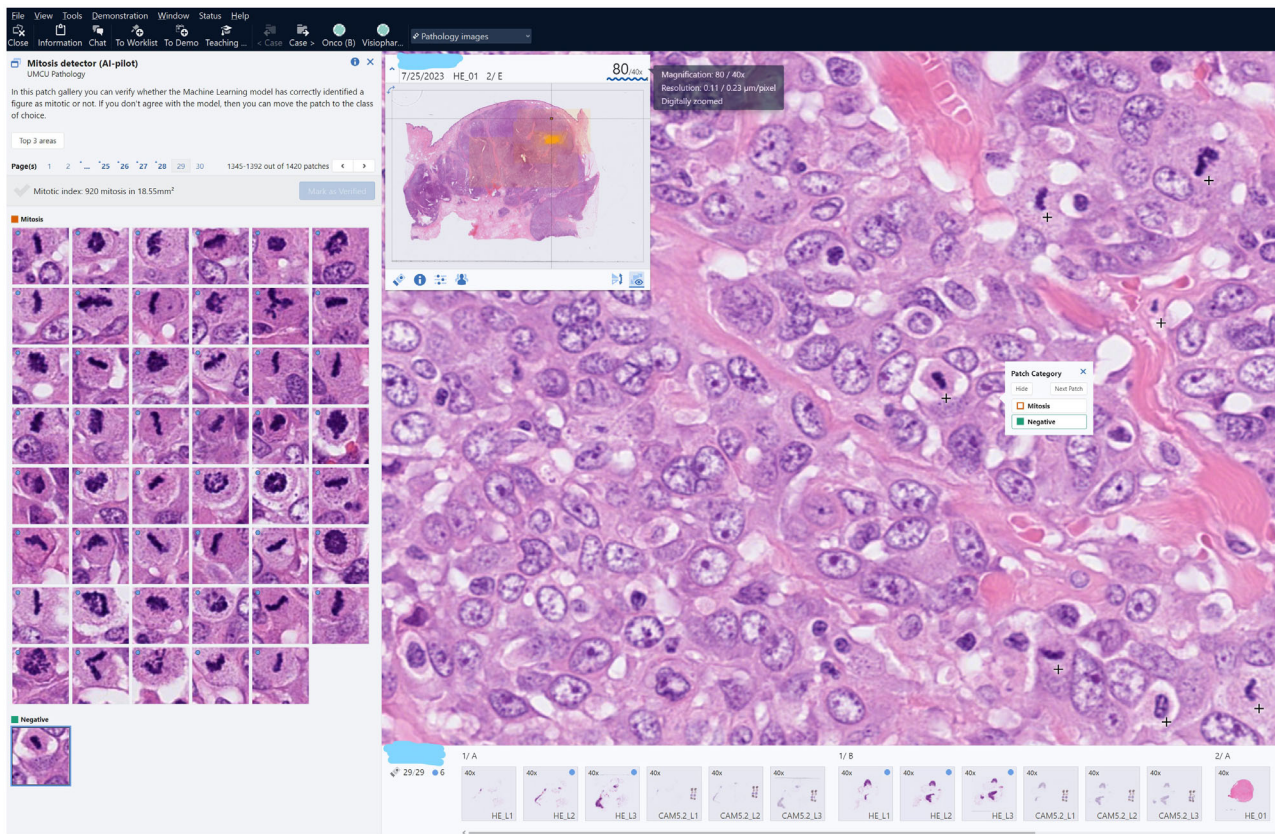


Figure 1. Screenshot of the UMC Utrecht home-brew mitoses detector integrated within the Sectra PACS. On the left, thumbnail galleries with candidate mitoses are displayed that can be clicked and reviewed at high magnification in the viewer on the right and confirmed as “mitoses” or “negative”, after which they will stay in the mitoses gallery or be moved to the negative gallery on the bottom left. Alternatively, thumbnails can be dragged from one gallery to the other.

prostate biopsies,⁴¹ or in identifying intrahepatic cholangiocarcinoma from colorectal liver metastases.⁴² For example, in the UMC Utrecht, where HE slides of sentinel lymph nodes are assessed first, and stains are requested when these slides are morphologically negative, we spent over €30,000 on keratin stains for sentinel lymph node assessment in approximately 180 breast cancer patients. Again, in most cases, these stains do not show any (clinically relevant) metastases. If a metastases detection algorithm would find all micro-metastases (Figure 2), less immunohistochemistry (IHC) would be necessary, since in breast cancer isolated tumour cells in a lymph node have mostly no clinical consequences. Savings on immunohistochemistry may be even more significant in laboratories that routinely stain all sections without assessing the HE slides first. This is currently prospectively evaluated in the CONFIDENT-B trial.⁴³

In the example of the assessment of prostate needle biopsies, all slides in the UMC Utrecht are routinely

stained using a cocktail of 34BE12, p63, and AMACR due to workflow optimization, which is not an uncommon practice in the Netherlands. Nevertheless, high variability between laboratories (and pathologists) exists in the use of IHC,^{44,45} and using an algorithm for tumour detection might decrease the need for expensive stains across the board, thereby decreasing overall costs. The CONFIDENT-P trial will provide important data on the role of AI in reducing the need for costly IHC stains in the diagnosis of prostate cancer, while maintaining diagnostic speed and accuracy.⁴³

ARTIFICIAL INTELLIGENCE MAKES DIAGNOSTICS MORE REPRODUCIBLE

Artificial intelligence has shown great promise in making pathology diagnostics more reproducible among pathologists. Examples include AI-supported mitoses counting,^{26,27,46} Gleason grading of prostate

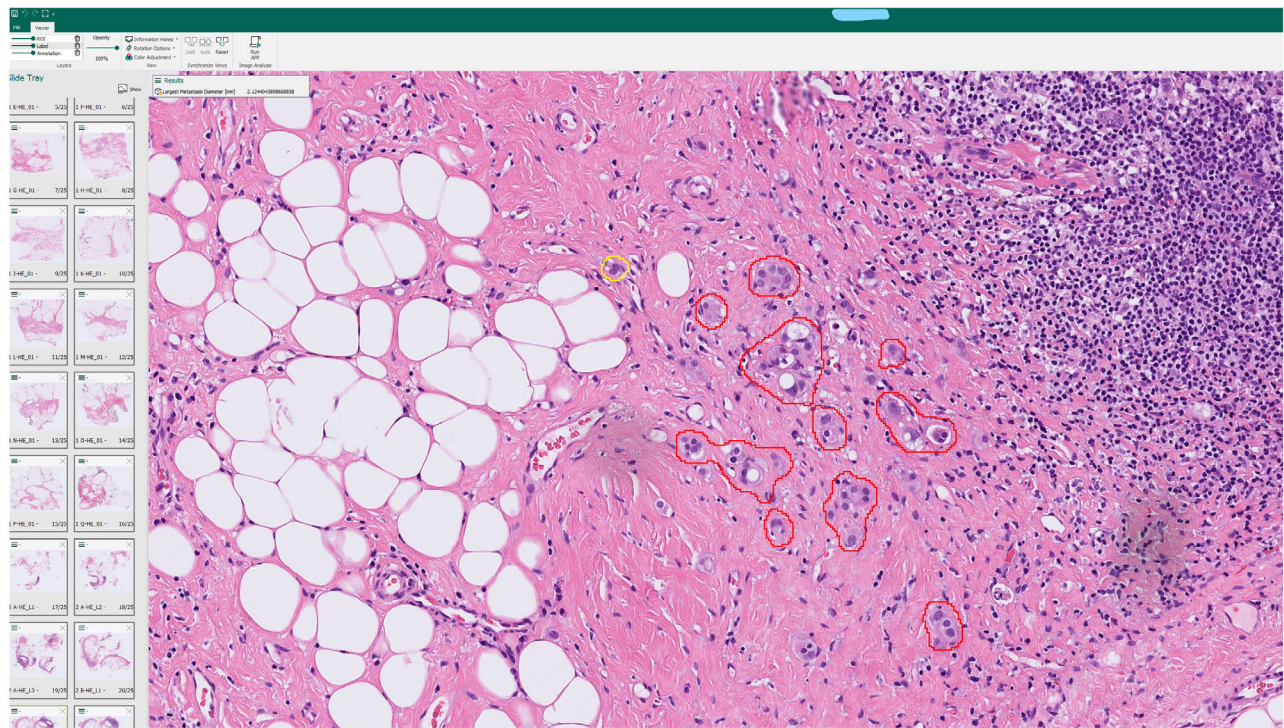


Figure 2. Screenshot of the Visiopharm lymph node metastases detector that has found isolated tumour cells after chemotherapy, delineated in red (high suspicion) and yellow (low suspicion).

cancer,^{22,47–49} and breast cancer grading,^{50,51} counting Ki67-positive cells,⁵² and therapeutic targets such as ER, PR, HER2, and PD-L1.⁵³

Scoring of HER2 IHC staining intensity, crucial for the decision on HER2 inhibition in breast cancer, is done more accurately by AI-supported pathologists compared to non-AI-assisted pathologists.⁵⁴ For prostate needle biopsies, variability in Gleason grading is higher in general pathologists than in expert urogenital pathologists.^{47,49} Grading prostate cancer more consistently improves the prognostic utility significantly. Some algorithms have even reached a kappa comparable to that of expert genitourinary pathologists.^{22,47} Multiple studies with different algorithms showed that general pathologists could reach more reproducible grading results, and have higher concordance with expert urogenital pathologists, compared to unassisted general pathologists or a standalone AI system.^{21,22,47}

In cytopathology, an example of improved reproducibility can be found in lung cancer cytology.²⁴ Using a deep learning-based image analysis algorithm, interrater agreement of three experienced pathologists improved significantly when comparing unassisted pathologists to AI-assisted pathologists

(Fleiss kappa 0.553–0.908).²⁴ This highlights the benefits of an AI-assisted approach in improving reproducibility, both in histopathology and cytopathology.

ARTIFICIAL INTELLIGENCE MAY IMPROVE PATIENT OUTCOMES AND GLOBAL HEALTHCARE COSTS

In a time where global healthcare costs are rising, and an almost 50% increase in cancer incidence is projected by GLOBOCAN⁵⁵; looking for cost and time savings is pivotal for the next few decades. Although diagnostic surgical and cytopathology only account for roughly 0.2–0.5% of global healthcare costs in developed countries, their role is pivotal in treatment decisions for a large number of patients. Patients' treatments are currently, however, influenced by considerable interobserver variation and institutional variation.^{6,56} Thereby, investment in good pathology, in which digital pathology, AI, and molecular pathology all play a big role, is inherently cost-effective.⁵⁷ While AI could lead to more cost-effective pathology workflows, leading to cost savings for pathology laboratories, even larger savings are probably found

outside the pathology laboratory. AI assistance could make pathology assessment more reproducible, leading to better treatments for patients, by reducing over- and undertreatment, saving side effects, and reducing healthcare costs.

In some cases, it can also be implemented as a screening tool, like the algorithm that Gao *et al.* developed and combined cytology with epidemiological factors to screen patients for oesophageal cancer, which could save endoscopy screenings and subsequent costs.⁵⁸ Some of this money should be redirected to pathology labs to facilitate implementation of digital pathology and AI.

Cons of introducing AI in diagnostic pathology

The cons of AI in diagnostic pathology are various in nature. They encompass safety issues, practical issues, such as implementation problems and ethical and regulatory concerns. In the following section we will first discuss our view on different safety issues, followed by the practical issues we envision, and encountered ourselves in our journey of digital pathology and AI implementation, followed by ethical and legal questions.

SAFETY ISSUES

Current algorithms are first generation and imperfect

Indeed, some of the algorithms on the market are first generation. However, much effort has been put into training and validating these algorithms, sometimes on tens of thousands of WSI.⁴⁸ No algorithm is perfect, but that does not preclude using them as long as this is done in a critically supervised way. There is room for improvement, considering that algorithms are only valuable under specific circumstances. For example, algorithms for prostate needle biopsies do not (yet) work on specimens obtained by transurethral resection of the prostate or by radical prostatectomy. They can only diagnose adenocarcinoma, but not other types of prostate cancer (e.g. small-cell carcinoma). Over time, better algorithms will become available based on feedback from using them in daily practice. Lastly, algorithms may eventually become self-learning based on immediate feedback during the diagnostic process, but presently, there is little proof of principle for this feature. Algorithms will get worse when fed with the wrong feedback. This is the reason that Conformité Européenne - In Vitro Diagnostics

(CE-IVD) approved algorithms cannot currently be trained and modified based on immediate feedback systems.

Trusting AI is dangerous

Some argue that using AI is dangerous, since trusting these algorithms too much could make us uncritical and lazy, and, consequently, prone to diagnostic errors. At this point, we would like to draw a parallel to stains since, in many ways, AI algorithms are just “digital stains”. At the beginning of IHC, we may have had the same problems. Initial overestimations of sensitivity and specificity claimed in the first publications have perhaps also made us uncritical and lazy. However, we learned that IHC stains might show artefacts and be false-positive or -negative, and thus should be disregarded, and that sensitivity and specificity are never perfect.

For AI algorithms, it is no different. We will have to be critical of the output of any AI algorithm, carefully supervise everything, and use what we like and ignore, overrule, or improve what we do not like, just as with stains. We must realize that AI algorithms can only recognize what they have been trained for. Further, a few algorithms give a level of certainty for their output, which would nevertheless be a highly recommendable feature for most algorithms. Hence, we advocate AI-assisted diagnostics, AI-supported diagnostics, or augmented diagnostics. Ultimately, we remain fully responsible and accountable for the contents of the pathology report, AI-supported or not. We do not anticipate AI algorithms ever to work entirely independently. Therefore, pathologists must be educated in understanding the abilities and limitations of AI assistance, before deploying them in daily practice.

Artificial intelligence will replace pathologists

An international survey among 718 dermatopathologists showed that, although the vast majority agreed that AI would improve dermatopathology, 6% of the pathologists feared that AI would replace the human pathologist in the foreseeable future.⁵⁹ However, instead of being mutually exclusive, artificial and human intelligence complement each other, a concept known as *augmented intelligence*.^{48,49} Rather than replacing humans, AI enhances our human intelligence, leading to a valuable AI-pathologist synergy. Furthermore, independently operating algorithms (e.g., without human supervision) are highly unlikely to be implemented and accepted due to the medical, ethical, and legal consequences when the AI-generated diagnosis is incorrect.

Pathologists will always have to supervise and check AI output and be aware of potentially false-negative or false-positive results, as they do in daily practice with nondigital stains and molecular test results. Yet, in the near future, AI will help reduce costs and make our jobs faster, more reproducible, more satisfying, and more appealing for the younger generation. This way, pathologists can focus on their most important crucial task: interpreting findings instead of tediously looking for the right correct diagnostic clues and, unfortunately, sometimes missing them anyway.

Next to this, AI tools that can increase the diagnostic volume that pathologists can handle are desperately needed, considering the existing global shortage of pathologists, which is only expected to become worse in a world where the need for pathologists' expertise is dire, due to an increasing cancer incidence and increasingly complex diagnostic options and reporting guidelines.^{55,60} Therefore, the only pathologists to be replaced by AI may be the inflexible ones unwilling to go digital and use AI to assist them in their jobs.

ETHICAL AND REGULATORY ISSUES

Artificial intelligence contains bias

Models with relatively homogenous data are at risk for data bias. A limited range of patient populations and clinical settings can lead to a lack of diversity in data, resulting in biased decisions made by AI and a decreased ability to apply results to a broader population. An algorithm can function correctly if the data are realistic and diverse, encompassing the full range of features present in future data to limit bias. The data used to teach AI systems can be biased due to past human actions related to factors like race, religion, ethnicity, or gender.⁶¹ While common cancers have a consistent appearance, certain groups may be affected differently due to prevalence or aggressiveness. Obermeyer and Topol⁶² argue that AI can scale up (racial) bias and redress it if trained properly by aligning algorithms with patients' outcomes and experiences.

Training sets for AI algorithms may have been biased, and hardly ever a detailed description is given of the training set for commercial algorithms on company websites, which is something that the companies marketing algorithms must certainly address better. Validation studies should be published, or at least websites should provide detailed information, for example, on numbers of cases with a number breakdown of the various categories, histological diagnoses,

length of follow-up, years of diagnoses, and type(s) of scanners, etc.^{10,63}

FDA- or CE-IVD approval

Integrating AI into pathology practice does not present any inherent ethical obstacles, similar to the absence of such hurdles for stains, quantitation, and molecular techniques. However, it is essential to ensure that AI algorithms meet rigorous quality standards and undergo thorough validation, acknowledging that they may not achieve perfection, but can still be highly valuable and effective tools.⁶⁴ Companies should therefore disclose on their websites how algorithms were trained, publish the results of validation studies in peer-reviewed journals through collaborations with pathology labs and go for international quality certifications such as CE-IVD or FDA approval.

However, if not available, it is still possible to use (even home-brew) algorithms after proper in-house validation. Guidelines for this are eagerly awaited.

PRACTICAL ISSUES

Algorithms will not work across labs with different staining procedures and scanners

For some older algorithms, this may have been the case. However, AI algorithm training procedures nowadays include slides stained in different labs and scanned with different types of scanners. Further, stain normalization procedures have become available,⁶⁵ and existing algorithms can be retrained to work on the staining spectrum in a specific lab.

Algorithms are expensive and not reimbursed

For many algorithms brought to the market by companies, it is unclear what the market price will be for a licence or pay-per-view, since the pathology market still needs to assess what prices are fair. A major problem here is that in most countries around the globe, there will be no extra reimbursement for using AI. Although money will be saved elsewhere when we go digital and start using AI (e.g., savings on IHC stains and time) because we will do a better (and more efficient) job as pathologists, it is improbable that those budgets will be redirected to pathology departments. Nevertheless, it is something for which we should keep advocating by showing the benefits of digital pathology and AI.

Since pathology is in general a poor specialty, and going digital, if at all financially possible, will already deplete most laboratory resources, there must be a business case for implementing AI based on tangible

savings in personnel or consumable costs. Because of limited daily AI case volumes in most pathology laboratories, tangible savings in personnel will not be easy to achieve. Lower turn-around times (TAT) will likely result in pathologists going home a little earlier every day instead of lowering the number of full-time equivalent pathologists, except for mono-specialty or very large laboratories. So, we must look for a business case for saving costs on consumables, such as IHC. To this end, we are currently running two clinical trials at the UMC Utrecht on breast cancer sentinel lymph nodes (CONFIDENT-B) and prostate biopsies (CONFIDENT-P) to assess in daily pathology practice how often pathologists forgo IHC when they are AI-supported.⁴³

AI requires a fully digital workflow

Undoubtedly, AI will be most efficiently deployed when integrated into a fully digital workflow, where it can run automatically in the background, or inter-actively kicked in on the spot when needed during diagnostics. Going digital means that pathologists need to be trained, and this requires a significant investment of time and money. For labs that are not yet fully digital, the possibility of running AI could be a good incentive and selling point to the hospital management to decide to go fully digital. Still, it is possible to run AI on cloud servers in selected cases scanned on demand.

For cytopathology specifically, the digital workflow renders a new set of challenges, compared to histopathology, such as the 3D distribution of cytology material on a glass slide, which warrants the need to capture images on multiple planes (z-stacking).⁶⁶ Cytopathology files are therefore much larger than histopathology slides. Moreover, the amount of colour information due to multiple staining techniques is higher. However, some imaging systems that work on glass slides have received FDA approval, such as the ThinPrep Imaging System (Hologic, Marlborough, MA, US), or the FocalPoint GS Imaging System (Becton Dickinson, Franklin Lakes, NJ, US).⁶⁶ Algorithms for such systems are being developed, enabling implementation of AI assistance even in not fully digital workflows,^{10,15} and the first CE-marked algorithm for cervical cytology on Thinprep glass slides has also been released.⁶⁷

Integration in PACS systems needs to be improved

Most companies have yet to realize full integration with the major Picture Archiving and Communication System (PACS) systems. However, this is of the utmost importance, since we do not want to work

with different viewers for the various AI algorithms we use. Most PACS systems have the option to integrate AI algorithms through an Application Programming Interface (API). However, this is a challenge. A lab may need in-house AI specialists to crack the API, together with the PACS vendor, to achieve this. Therefore, companies bringing AI algorithms to the market should realize that customers want full integration with their PACS. Therefore, realizing full integration with the major PACS systems and optimizing PACS display should be an early step in the development roadmap of AI companies.

Implementation is difficult

In view of the above, implementation may be challenging, especially for algorithms that have yet to be integrated with the PACS system that is used. It will be difficult for many labs to fund AI specialists, and central Information Technology (IT) services usually cannot handle this. Therefore, some extra incidental budget to achieve AI implementation may be required.

Further, sufficient graphics processing unit (GPU) computing needs to be realized to process the large whole-slide images we produce in pathology. This will be even more so in cytopathology slides, where whole-slide data analysis is extremely challenging, due to giga-scale slides.⁶⁸ This can be done locally within the firewall of institutions that precludes security problems, often needing the help of central IT if one wants to avoid buying and maintaining their servers, or through a cloud GPU service that is easier scalable but may bring on security problems. AI companies should therefore offer as much as possible turn-key compute solutions.

AI startup companies will disappear

Numerous AI startup ventures, driven by ambitious marketing strategies, may find themselves excessively optimistic, considering the formidable challenges of implementing AI in routine pathology practice. Some of these startups may not endure, particularly if the pathology community is reluctant to bear the costs associated with AI algorithms. Nevertheless, larger and more established companies will likely absorb these struggling startups, hopefully ensuring that valuable algorithms are preserved and not lost.

Conclusion

Many good AI algorithms exist, which can be utilized to gain experience in AI-assisted pathology

assessments and undergo further prospective validation in clinical practice. Continuous feedback will contribute to the ongoing development of these algorithms. We believe that AI assistance has the potential to enhance the pathologists' workflow and the overall quality of their work, leading to better diagnoses for patients, which is likely to result in better outcomes. While there are notable advantages, it is vital to acknowledge and address the associated challenges, which seem largely manageable. In summary, AI can enhance the diagnostic process by providing greater accuracy, improved reproducibility, faster TATs, increased work satisfaction for pathologists, and ultimately better patient treatment. Yes, we can!

Author contributions

Conceptualization: P.D. and R.F.; investigation: all authors; writing, original draft preparation, P.D.; writing, review and editing, all authors. All authors have read and agreed to the published version of the article.

Funding information

This research received no external funding.

Conflict of interest

P.D. is a member of the advisory board of Visiopharm and Sectra. All other authors declare no conflicts of interest.

Data availability statement

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

References

- van Doijeweert C, van Diest PJ, Ellis IO. Grading of invasive breast carcinoma: The way forward. *Virchows Archiv* 2022; **480**: 33–43.
- van Doijeweert C, van Diest PJ, Willems SM, Kuijpers CCHJ, Overbeek LIH, Deckers IAG. Significant inter- and intra-laboratory variation in grading of ductal carcinoma in situ of the breast: a nationwide study of 4901 patients in The Netherlands. *Breast Cancer Res. Treat.* 2019; **174**: 479–488.
- van Doijeweert C, van Diest PJ, Willems SM et al. Significant inter- and intra-laboratory variation in grading of invasive breast cancer: a nationwide study of 33,043 patients in The Netherlands. *Int. J. Cancer* 2020; **146**: 769–780.
- Flach RN, Willemse PPM, Suelmann BBM et al. Significant inter-and intralaboratory variation in gleason grading of prostate cancer: a nationwide study of 35,258 patients in The Netherlands. *Cancers (Basel)* 2021; **13**: 5378.
- Vestjens JHMJ, Pepels MJ, De Boer M et al. Relevant impact of central pathology review on nodal classification in individual breast cancer patients. *Ann. Oncol.* 2012; **23**: 2561–2566.
- van Doijeweert C, Baas IO, Deckers IAG, Siesling S, van Diest PJ, van der Wall E. The increasing importance of histologic grading in tailoring adjuvant systemic therapy in 30,843 breast cancer patients. *Breast Cancer Res. Treat.* 2021; **187**: 577–586.
- Yousif M, van Diest PJ, Laurinavicius A et al. Artificial intelligence applied to breast pathology. *Virchows Archiv.* 2022; **480**: 191–209.
- van der Laak J, Litjens G, Ciompi F. Deep learning in histopathology: the path to the clinic. *Nat. Med.* 2021; **27**: 775–784.
- Flach RN, Fransen NL, Sonnen AFP et al. Implementation of artificial intelligence in diagnostic practice as a next step after going digital: the UMC utrecht perspective. *Diagnostics* 2022; **12**: 1042.
- Nakagawa K, Moukheiber L, Celi LA et al. AI in pathology: what could possibly go wrong? *Semin. Diagn. Pathol.* 2023; **40**: 100–108.
- Abels E, Pantowitz 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**: 286–294.
- Senaras C, Khalid Khan Niazi M, Lozanski G, Gurcan MN. DeepFocus: detection of out-of-focus regions in whole slide digital images using deep learning. *PloS One* 2018; **13**: e0205387.
- Jacq A, Tarris G, Jaugey A et al. Automated evaluation with deep learning of total interstitial inflammation and peritubular capillaritis on kidney biopsies. *Nephrol. Dial. Transplant.* 2023; **38**: 2786–2798.
- Kers J, Bülow RD, Klinkhammer BM et al. Deep learning-based classification of kidney transplant pathology: a retrospective, multicentre, proof-of-concept study. *Lancet Digit Health* 2022; **4**: e18–e26.
- Bejnordi BE, Veta M, Van Diest PJ et al. Diagnostic assessment of deep learning algorithms for detection of lymph node metastases in women with breast cancer. *JAMA* 2017; **318**: 2199–2210.
- Jarkman S, Karlberg M, Pocevičiūtė M, Bodén A, Bándi P et al. Generalization of deep learning in digital pathology: experience in breast cancer metastasis detection. *Cancers* 2022; **14**: 5424.
- Steiner DF, Macdonald R, Liu Y et al. Impact of deep learning assistance on the histopathologic review of lymph nodes for metastatic breast cancer. 2018; **42**(12): 1636–1646. www.ajsp.com.
- Chuang WY, Chen CC, Yu WH et al. Identification of nodal micrometastasis in colorectal cancer using deep learning on annotation-free whole-slide images. *Mod. Pathol.* 2021; **34**: 1901–1911.
- Báñdi P, Balkenhol M, van Dijk M et al. Continual learning strategies for cancer-independent detection of lymph node metastases. *Med. Image Anal.* 2023; **85**: 102755.
- Jansen P, Baguer DO, Duschner N et al. Deep learning detection of melanoma metastases in lymph nodes. *Eur. J. Cancer* 2023; **188**: 161–170.
- Raciti P, Sue J, Ceballos R et al. Novel artificial intelligence system increases the detection of prostate cancer in whole slide

- images of core needle biopsies. *Mod. Pathol.* 2020; **33**: 2058–2066.
22. Steiner DF, Nagpal K, Sayres R *et al.* Evaluation of the use of combined artificial intelligence and pathologist assessment to review and grade prostate biopsies. *JAMA Netw. Open* 2020; **3**: e2023267. <https://jamanetwork.com/>.
 23. Hou X, Shen G, Zhou L, Li Y, Wang T, Ma X. Artificial intelligence in cervical cancer screening and diagnosis. *Front. Oncol.* 2022; **12**: 851367. Available from: <https://www.frontiersin.org/articles>
 24. Kim T, Chang H, Kim B *et al.* Deep learning-based diagnosis of lung cancer using a nationwide respiratory cytology image set: improving accuracy and inter-observer variability. *Am. J. Cancer Res.* 2023; **13**: 5493–5503. www.ajcr.us/.
 25. Lebet T, Paoletti X, Pignot G *et al.* Artificial intelligence to improve cytology performance in urothelial carcinoma diagnosis: results from validation phase of the French, multicenter, prospective VISIOCYT1 trial. *World J. Urol.* 2023; **41**: 2381–2388.
 26. Veta M, Heng YJ, Stathonikos N *et al.* Predicting breast tumor proliferation from whole-slide images: the TUPAC16 challenge. *Med. Image Anal.* 2019; **54**: 111–121.
 27. Balkenhol MCA, Tellez D, Vreuls W *et al.* Deep learning assisted mitotic counting for breast cancer. *Lab. Invest.* 2019; **99**: 1596–1606.
 28. Perincheri S, Levi AW, Celli R *et al.* An independent assessment of an artificial intelligence system for prostate cancer detection shows strong diagnostic accuracy. *Mod. Pathol.* 2021; **34**: 1588–1595.
 29. Litjens G, Sánchez CI, Timofeeva N *et al.* Deep learning as a tool for increased accuracy and efficiency of histopathological diagnosis. *Sci. Rep.* 2016; **6**: 26286.
 30. Esteve A, Kuprel B, Novoa RA *et al.* Dermatologist-level classification of skin cancer with deep neural networks. 2017 <https://licensing.eri.ed.ac.uk/i/>.
 31. Hekler A, Utikal JS, Enk AH *et al.* Superior skin cancer classification by the combination of human and artificial intelligence. *Eur. J. Cancer* 2019; **120**: 114–121.
 32. Lew M, Wilbur DC, Pantanowitz L. Computational cytology: lessons learned from pap test computer-assisted screening. *Acta Cytol.* 2021; **65**: 286–300.
 33. Thrall MJ. Automated screening of papanicolaou tests: a review of the literature. *Diagn. Cytopathol.* 2019; **47**: 20–27. <https://onlinelibrary.wiley.com>.
 34. Le Vuong TT, Song B, Kwak JT, Kim K. Prediction of Epstein-Barr virus status in gastric cancer biopsy specimens using a deep learning algorithm. *JAMA Netw. Open* 2022; **5**: 2236408.
 35. Wang X, Zou C, Zhang Y *et al.* Prediction of BRCA gene mutation in breast cancer based on deep learning and histopathology images. *Front. Genet.* 2021; **12**(July): 661109.
 36. Pizurica M, Larmuseau M, Van der Eecken K *et al.* Whole slide imaging-based prediction of TP53 mutations identifies an aggressive disease phenotype in prostate cancer. *Cancer Res.* 2023; **83**: 2970–2984.
 37. Echle A, Grabsch HI, Quirke P *et al.* Clinical-grade detection of microsatellite instability in colorectal tumors by deep learning. *Gastroenterology* 2020; **159**: 1406–1416.e11.
 38. Yamashita R, Long J, Longacre T *et al.* Deep learning model for the prediction of microsatellite instability in colorectal cancer: a diagnostic study. *Lancet Oncol.* 2021; **22**: 132–141.
 39. Nikolas Kather J, Heij LR, Grabsch HI *et al.* Pan-cancer image-based detection of clinically actionable genetic alterations. *Nat. Cancer* 2020; **1**: 789–799.
 40. Kacew AJ, Strohbehn GW, Saulsberry L *et al.* Artificial intelligence can cut costs while maintaining accuracy in colorectal cancer genotyping. *Front. Oncol.* 2021; **11**: 630953.
 41. Baeten IGT, Hoogendam JP, Jonges GN *et al.* Value of routine cytokeratin immunohistochemistry in detecting low volume disease in cervical cancer. *Gynecol. Oncol.* 2022; **165**: 257–263.
 42. Albrecht T, Rossberg A, Albrecht JD *et al.* Deep learning-enabled diagnosis of liver adenocarcinoma. *Gastroenterology* 2023; **165**: 1262–1275.
 43. Flach RN, Stathonikos N, Nguyen TQ, Ter Hoeve ND, van Diest PJ, van Doijeweert C. CONFIDENT-trial protocol: a pragmatic template for clinical implementation of artificial intelligence assistance in pathology. *BMJ Open* 2023; **13**: 67437. <http://bmjopen.bmj.com/>.
 44. Bonert M, El-Shinnawy I, Rahman M *et al.* Immunohistochemistry use by diagnostic category and pathologist in 4477 prostate Core biopsy sets assessed at two hospitals. *Appl. Immunohistochem. Mol. Morphol.* 2020; **28**(4): 259–266. <http://links.lww.com/AIMM/A218>.
 45. Al Diffalha S, Shaar M, Barkan GA, Wojcik EM, Picken MM, Pambuccian SE. Immunohistochemistry in the workup of prostate biopsies: frequency, variation and appropriateness of use among pathologists practicing at an academic center. *Ann. Diagn. Pathol.* 2017; **270**: 34–42.
 46. van Bergeijk SA, Stathonikos N, ter Hoeve ND *et al.* Deep learning supported mitoses counting on whole slide images: a pilot study for validating breast cancer grading in the clinical workflow. *J. Pathol. Inform.* 2023; **14**: 100316.
 47. Bulten W, Balkenhol M, Belinga JJA *et al.* Artificial intelligence assistance significantly improves Gleason grading of prostate biopsies by pathologists. *Mod. Pathol.* 2021; **34**: 660–671.
 48. Bulten W, Kartasalo K, Chen PHC *et al.* Artificial intelligence for diagnosis and Gleason grading of prostate cancer: the PANDA challenge. *Nat. Med.* 2022; **28**: 154–163.
 49. Ström P, Kartasalo K, Olsson H *et al.* Artificial intelligence for diagnosis and grading of prostate cancer in biopsies: a population-based, diagnostic study. *Lancet Oncol.* 2020; **21**: 222–232.
 50. Wetstein SC, de Jong VMT, Stathonikos N *et al.* Deep learning-based breast cancer grading and survival analysis on whole-slide histopathology images. *Sci. Rep.* 2022; **12**: 15102.
 51. Wang Y, Acs B, Robertson S *et al.* Improved breast cancer histological grading using deep learning. *Ann. Oncol.* 2022; **33**: 89–98.
 52. Zehra N, Shams M, Ahmad Z, Chundrigger Q, Ahmed A, Jaffar N. Digital imaging AI software and its concordance with manual method. *J. Coll. Physicians Surg. Pak.* 2023; **33**: 544–547.
 53. Vahadane A, Sharma S, Mandal D *et al.* Development of an automated combined positive score prediction pipeline using artificial intelligence on multiplexed immunofluorescence images. *Comput. Biol. Med.* 2023; **152**: 106337.
 54. Yue M, Zhang J, Wang X *et al.* Can AI-assisted microscope facilitate breast HER2 interpretation? A multi-institutional ring study. *Virchows Arch.* 2021; **479**: 443–449.
 55. Ferlay J, Laversanne M, Ervik M, Lam F, Colombet M, Mery L, *et al.* *Global cancer observatory: cancer tomorrow*. Lyon, France: International Agency for Research on Cancer. 2020 [cited 2023 Jun 20]. Available at: <https://gco.iarc.fr/tomorrow>
 56. Flach RN, van Doijeweert C, Aben KKH *et al.* Interlaboratory Gleason grading variation affects treatment: a Dutch historic cohort study in 30 509 patients with prostate cancer. *J. Clin. Pathol.* 2023; **76**: 690–697. <https://jcp.bmj.com/lookup>.
 57. Ho J, Ahlers SM, Stratman C *et al.* Can digital Pathology result in cost savings? A financial projection for digital pathology

- implementation At a large integrated health care organization. *J. Pathol. Inform.* 2014; **5**: 33.
58. Gao Y, Xin L, Lin H *et al.* Machine learning-based automated sponge cytology for screening of oesophageal squamous cell carcinoma and adenocarcinoma of the oesophagogastric junction: a nationwide, multicohort, prospective study. *Lancet Gastroenterol. Hepatol.* 2023; **8**: 432–445.
 59. Polesie S, McKee PH, Gardner JM *et al.* Attitudes toward artificial intelligence within dermatopathology: an international online survey. *Front. Med. (Lausanne)* 2020; **7**: 591952.
 60. Steiner DF, Chen PHC, Mermel CH. Closing the translation gap: AI applications in digital pathology. *Biochim. Biophys. Acta Rev. cancer.* 2021; Vol. **1875**: 188452.
 61. McKay F, Williams BJ, Prestwich G, Bansal D, Hallowell N, Treanor D. The ethical challenges of artificial intelligence-driven digital pathology. *J. Pathol. Clin. Res.* 2022; **8**: 209–216. <https://onlinelibrary.wiley.com>.
 62. Obermeyer Z, Topol EJ. Artificial intelligence, bias, and patients' perspectives. *Lancet* 2021; **397**: 2038.
 63. Eelbode T, Sinonquel P, Maes F, Bisschops R. Pitfalls in training and validation of deep learning systems. *Best Pract. Res. Clin. Gastroenterol.* 2021; **52**: 101712.
 64. Bera K, Schalper KA, Rimm DL, Velcheti V, Madabhushi A. Artificial intelligence in digital pathology — new tools for diagnosis and precision oncology. *Nat. Rev. Clin. Oncol.* 2019; **16** (11): 703–715. www.nature.com/nrclinonc.
 65. Bejnordi BE, Litjens G, Timofeeva N *et al.* Stain specific standardization of whole-slide histopathological images. *IEEE Trans. Med. Imaging* 2016; **35**: 404–415.
 66. McAlpine ED, Pantanowitz L, Michelow PM. Challenges developing deep learning algorithms in cytology. *Acta Cytol.* 2021; **65**: 301–309.
 67. Genius Digital Diagnostics System. [cited 2023 Dec 28]. Available at: <https://www.hologic.com/hologic-products/cytology/genius-digital-diagnostics-system>
 68. Jiang H, Zhou Y, Lin Y, Chan RCK, Liu J, Chen H. Deep learning for computational cytology: a survey. *Med. Image Anal.* 2023; **84**: 102691.