Review paper

Hardware and quality-control factors influencing digital pathology reproducibility

Yan Li[1], Tiantian Li[2]

[1]School of Mechanical and Electrical Engineering, Jilin Engineering Vocational College, Jilin, China
[2]Department of Pathology and Experimental Medicine, Graduate School of Medicine, Dentistry and Pharmaceutical Sciences, Okayama University, Okayama, Japan

Digital pathology (DP) converts glass slides into whole-slide images (WSI), enabling quantitative and remote diagnostics. This systematized review examines how hardware and instrumentation (HI) form the technical basis for reliable DP. Evidence from WSI validation and engineering studies shows that section-thickness variation, staining consistency, and scanner calibration are major determinants of diagnostic reproducibility. Comparative studies consistently show that even small mechanical or optical deviations – such as micron-level thickness variation or illumination shifts – produce measurable image differences that affect both inter-site diagnostic concordance and artificial intelligence model stability. We propose a unified framework linking HI parameters to quality-control (QC) indicators and regulatory requirements, including device documentation, software validation, and metadata traceability. Integrating structured metadata from tissue processing through WSI acquisition supports reproducible, auditable, and clinically compliant DP workflows. By connecting validated WSI practice, QC mechanisms, and HI-centered standardization, this review emphasizes that reliable DP depends not only on computational methods but also on precise, well-regulated, and interoperable engineering systems.

Key words: digital pathology, artificial intelligence, hardware and instrumentation, whole-slide imaging, quality control, workflow standardization.

Introduction

Digital pathology (DP) has transformed diagnostic histopathology by enabling whole-slide imaging (WSI), remote consultation, and artificial intelligence (AI)-assisted interpretation. Over the past decade, advances in scanners, automation, and data infrastructure have shifted practice from manual microscopy to integrated digital workflows capable of high-throughput scanning, automated quality control (QC), and large-scale analytics [1]. These developments were driven by telepathology needs, workforce shortages, and the rise of precision oncology.

However, despite rapid progress in algorithms, the reproducibility and generalizability of AI in pathology remain limited by upstream device- and process-level variability. Small changes in section thickness, staining uniformity, or scanner optics can change image appearance and bias downstream algorithms [2]. While computational methods are well studied, the device-level factors that determine image fidelity – microtomes, automated stainers, scanners, and laboratory automation – are less well represented in the literature.

This paper is a systematized review with an explicit positional perspective. We performed a structured search and screening (see Material and methods) to identify studies examining hardware, instrument-level QC, and workflow factors that affect WSI fidelity and AI performance. Our goals are:

- to synthesize measurable device-level parameters and QC indicators shown to influence image and algorithm reproducibility,
- to propose a practical framework linking hardware metadata, QC practices, and regulatory considerations for clinical deployment.

For clarity, we use the term hardware and instrumentation (HI) to refer specifically to operational characteristics of microtomes, stainers, scanners, and laboratory automation systems. Our focus is on parameters that are routinely encountered in practice and can be objectively monitored – such as section-thickness variability, staining consistency, optical calibration, and metadata capture – rather than on theoretical engineering constructs. By centering the review on these tangible factors, we aim to clarify how device-level variability propagates into diagnostic variability, and to highlight where standardized QC can most effectively mitigate these risks.

Material and methods

Study design and protocol

This work was conducted as a systematized review with a defined position perspective, following PRISMA 2020 principles. Although the scope did not qualify for PROSPERO registration, a predefined protocol was used to determine search strategy, screening procedures, and data-extraction criteria. The full protocol and search strings are available in the supplementary material.

Information sources and search strategy

A structured search was performed in PubMed, Web of Science, and IEEE Xplore for studies published between January 2018 and October 2025. Search terms combined concepts related to digital pathology, WSI, artificial intelligence, and hardware/device-level QC. Only peer-reviewed English-language studies were included. Complete search strings for each database are provided in the Supplementary Material.

Eligibility criteria

Studies were included if they met the following criteria:

- addressed DP or WSI,
- examined hardware, device behavior, QC parameters, or workflow interoperability,
- reported experimental evaluation or quantitative metrics,
- provided full-text original data.

Exclusion criteria: reviews, commentaries, or editorials without primary data; purely algorithmic studies with no laboratory, imaging, or device context; duplicate records or studies with insufficient methodological detail.

To reduce ambiguity, "hardware/QC studies" were defined as research examining sectioning, staining, scanning, imaging calibration, color variability, mechanical tolerances, or workflow-level QC instrumentation.

Study selection and data extraction

Two reviewers independently screened titles, abstracts, and full texts using Rayyan. Duplicates were removed automatically. Conflicts were resolved by discussion. The study selection process is illustrated in the PRISMA flow diagram (Figure 1).

Quality and bias assessment

Depending on the study type, the following tools were applied: QUADAS-2 for diagnostic accuracy studies, TRIPOD-AI for AI validation studies, AMSTAR-2 (adapted) for technical/mechanical validation studies. Inter-reviewer agreement was quantified using Cohen's κ, with $\kappa = 0.612$ considered acceptable. Summary ratings are provided in Supplementary Table S1–S3.

Data extraction and synthesis

Data extraction followed a standardized template covering study design, device category, QC parameters, performance metrics, and regulatory notes. Given the heterogeneity of the included studies, the findings were synthesized narratively and structured into categories related to DP validation. (Table I), clinical-grade AI studies (Supplementary Table S4), mechatronic engineering frameworks (Tables II, III, Supplementary Table S5).

Methodological note on synthesis

This systematized review included studies on scanner performance, staining reproducibility, workflow-related QC, and AI validation. Literature was identified through targeted searches using predefined keywords, with inclusion criteria focused on device-level parameters, quality-control practices, and external validation outcomes. Extracted data were categorized and integrated into the three frameworks described above to systematically illustrate the foundational role of HI within the DP ecosystem.

Digital pathology and artificial intelligence: current progress

Validation of whole slide imaging

Since the first Food and Drug Administration (FDA) clearance of WSI systems for primary diagnosis, multiple multicenter studies have evaluated whether WSI can reliably replace light microscopy. Across tissue types and clinical settings,

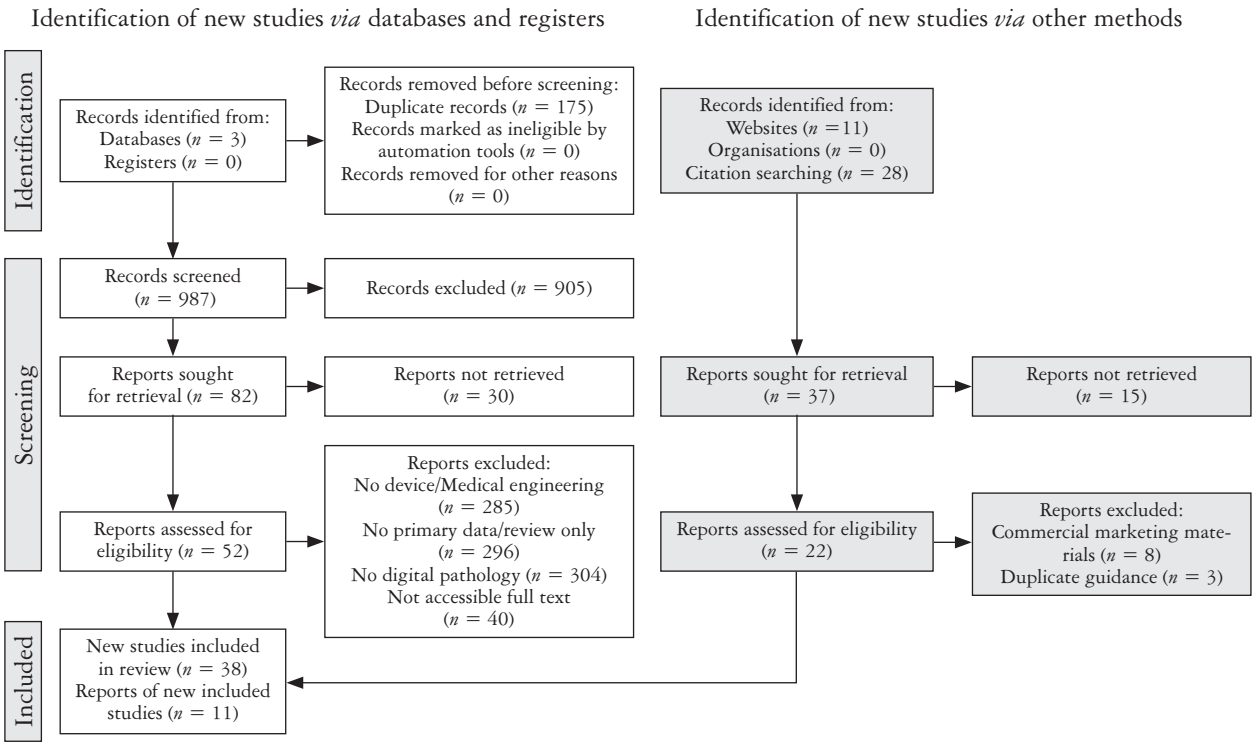

Figure 1. The study selection process in the PRISMA flow diagram

concordance rates remain consistently high – often > 95% for primary diagnoses when standardized scanning protocols and QC procedures are applied [3]. These results demonstrate that WSI can safely support diagnostic workflows when operated under controlled technical conditions.

Validation studies generally use retrospective or prospective designs across 5–20 participating sites and include tens of thousands of slides. Key performance indicators include diagnostic accuracy, turnaround time, and interobserver agreement, commonly assessed with non-inferiority margins of ±4%. Table I summarizes representative studies, highlighting design parameters, case numbers, diagnostic endpoints, and non-inferiority outcomes.

Taken together, these studies provide strong evidence supporting the clinical safety of WSI under appropriate device-level QC. However, variation in scanner models, staining techniques, and slide-preparation processes introduces heterogeneity that can influence external validity. Therefore, cross-device and cross-laboratory validation remains essential, particularly in settings where hardware configurations and workflow practices differ.

Clinical-grade AI in pathology

Recent AI advances have enabled automated tumor detection, grading, and biomarker quantification across multiple tissues. Large-scale studies employing weakly supervised, self-supervised, and multi-modal learning strategies have reported diagnostic performance with area under the curve values exceeding 0.90 in lung [11], breast [12], and colon cancer cohorts [13].

Several tools developed under the FDA software as a medical device (SaMD) framework have also demonstrated consistent external validation. However, direct comparison across systems remains challenging due to differences in annotation granularity, dataset size, and validation methodology. For example, some studies include more than 10,000 slides with prospective multicenter evaluation, whereas others rely on retrospective single-institution cohorts, resulting in substantial variability in generalizability [14–16].

Across the reviewed evidence, a consistent finding is that many performance limitations arise not from model architecture itself but from upstream technical variability, specifically staining inconsistency, section-thickness variation, tissue artifacts, and scanner-dependent optical differences. As summarized in Supplementary Table S4, issues such as color drift, staining variation, tissue folds, and device-specific imaging characteristics frequently contribute to misclassification or reduced model confidence even when the algorithm remains unchanged.

Taken together, three recurrent determinants emerged as essential for robust clinical-grade AI:

- dataset diversity – inclusion of heterogeneous staining protocols, scanner models, and tissue sources,
- staining and imaging consistency – maintained through standardized workflows and hardware-level QC,
- transparent validation – external, multi-center, and prospective testing aligned with SaMD principles.

Table I. Representative multicenter validation studies for whole slide imaging

Study	Design	*n* slides	Body sites	Endpoints (primary)	NI outcome	External validation	Limitations
Botros *et al.*, 2024 [4]	Retrospective cohort	290	Esophageal biopsies	Diagnostic concordance/ pathologist assessment	Yes	Yes	Single scanner type used/limited generalizability/small sample size for rare categories
Wu *et al.*, 2024 [5]	Multicenter validation retrospective cohort	8225	Prostate cancer	Remote reporting feasibility/ diagnostic concordance	Yes	Yes	Potential selection bias/small sample size in some subgroups/need for prospective validation/limited to high-resolution WSIs
Geread *et al.*, 2020 [6]	Multicenter validation	400	Breast cancer	Diagnostic concordance/ reproducibility (inferred)	Yes	Yes	Potential impact of outliers/need for integration with slide viewers
Mohanty *et al.*, 2023 [7]	Multicenter validation retrospective cohort	113	Dental epithelium	Turnaround/remote feasibility/ diagnostic concordance	N/A	No	Single-center/small sample size/exclusion of imperfect slides/time-consuming manual labeling
Study	Design	*n* slides	Body sites	Endpoints (primary)	NI outcome	External validation	Limitations
Yuenyong *et al.*, 2023 [8]	Retrospective cohort	88	Lymphoma	Diagnostic accuracy/validation metrics	N/A	No	Dataset representativeness/manual ROI selection required/color fading in older slides affects performance
Yu *et al.*, 2025 [9]	Retrospective cohort	633	Head/neck	Diagnostic/biomarker prediction (AUC – 0.998 in snippet)	N/A	No	Single-institution scanner data/lack of multi-institutional validation/ potential morphological changes in FFPE specimens
Parra-Medina *et al.*, 2025 [10]	Retrospective cohort	4187	Gastric	Diagnostic metrics (AUC – 0.998 in snippet)	N/A	Yes	Small training set size/single-center data/lack of generalizability/technical variations/inconsistent ground truth labeling at annotation boundaries
Tsuneki *et al.*, 2022 [11]	Multicenter validation retrospective cohort	5103	Poorly differentiated adenocarcinoma (gastric)	Classification accuracy for poorly differentiated adenocarcinoma	N/A	Yes	Potential bias toward specific specimen characteristics/need for further validation *via* randomized trials/limited number of hospitals

AUC – area under the curve, FFPE – formalin-fixed paraffin-embedded, N/A – not applicable, NI outcome – non-inferiority outcome, ROI – region of interest, WSI – whole-slide images

These findings underscore that dependable AI performance requires tight coupling between algorithm development and device-level quality assurance. AI models achieve reproducible results only when deployed within stable, well-controlled imaging environments.

Hardware and instrumentation and quality control in digital pathology systems

Engineering parameters and quality-control framework

Although computational pathology has advanced rapidly [17], the reproducibility of digital diagnostics continues to rely heavily on the instruments and processes that generate WSI. Small but measurable sources of variability – such as section-thickness deviations, minor focus errors, or reagent batch differences – can alter image appearance and subsequently influence AI performance.

To control this variability, device performance should be defined through specific and testable parameters, including section-thickness tolerance, autofocus repeatability, and illumination stability. Table II summarizes these key parameters, routine QC checks, and their relevance to AI-driven analyses. Capturing these measurements as structured metadata linked directly to image files enables downstream systems to detect, interpret, or correct for hardware-driven variation.

Reference architecture for interoperable metadata and quality-control flow

Ensuring reproducibility across laboratories requires more than standardized procedures – it demands a unified data and metadata architecture that connects mechatronic devices, QC logs, and AI analytics in real time [24, 25].

To address this need, we propose an interoperable reference framework integrating both physical and digital QC (Figure 2).

In this framework, each device contributes structured metadata to the laboratory information system and electronic health record:

- the microtome records section thickness, blade angle, and cutting speed,
- the water bath captures temperature and immersion time,
- the stainer/immunohistochemistry module logs reagent batch numbers, pH, temperature, and protocol version,
- the scanner provides optical calibration data, z-step size, illumination uniformity, and focus performance.

All metadata are encoded through digital imaging and communications in medicine (DICOM)-WSI

Table II. Engineering parameters, quality-control indicators, and corresponding AI failure modes

Device class	Key engineering parameters	QC/monitoring indicators	Why it matters for AI/DP	Expected AI failure modes
Microtome (tissue sectioning) [18]	Section thickness tolerance: ±0.5–1 μm Stage straightness < 2 μm Blade angle 5–10° Backlash < 1 μm	Micrometer gauges/optical profilometry/routine calibration	Section thickness variation alters nuclear size and chromatin texture features, leading to non-reproducible morphology	Apparent nuclear atypia or loss of nuclear detail → misclassification of tumor grade
Water-bath/slide transfer [19]	Temperature 37–40°C Automated flotation Transfer repeatability ±1 mm	Temperature sensors/automated slide tracking	Prevents folds, chatter, or compression artifacts	Folded/overstretched nuclei → segmentation errors
Automated stainers/ IHC platforms [20, 21]	Bath temperature ±1°C pH 7.2 ±0.2 Reagent turnover every 200 slides Reagent lot/barcode tracking	On-board sensors/reagent barcodes/external QA (NordiQC, UK NEQAS, CAP pIQAP)	Dye concentration and pH stability determine stain color consistency across labs	Cross-site domain shift; color normalization errors; false negatives in IHC positivity
Slide scanners (WSI) [22]	Stage repeatability < 1 μm Autofocus step 0.5–1 μm Pixel size 0.25 μm/px at 40× (Nyquist-compliant) Illumination drift ΔE < 2 over 1000 scans	Built-in blur/fold detection/ coverage heatmaps/ Periodic MTF checks	Resolution and focus stability directly affect tissue detail; illumination drift alters color distribution	Out-of-focus tiles → false negatives at invasive fronts; color drift → biomarker misclassification
File formats and interoperability [23]	Proprietary (SVS, NDPI) vs. DICOM-WSI; metadata completeness (scanner model, firmware, objective)	Metadata QA/LIS/FHIR integration/audit trail	Metadata gaps reduce reproducibility and hinder cross-center AI benchmarking	Missing scanner parameters → AI performance not generalizable

CAP – College of American Pathologist, DICOM – digital imaging and communications in medicine, DP – digital pathology, FHIR – Fast Healthcare Interoperability Resources, IHC – immunohistochemistry, LIS – laboratory information system, MTF – modulation transfer function, NEQAS – National External Quality Assessment Service, NDPI – Hamamatsu digital pathological image, NEQAS – National External Quality Assessment Service, pIQAP – pathology Imaging Quality Assurance Program, QA – quality assurance, QC – quality -control, SVS – aperio scan scope virtual slide, WSI – whole-slide images

Table III. Regulatory and standards matrix for mechatronics-artificial intelligence integration

Device class	Regulatory category	Key applicable standards	Compliance requirements	Post-market monitoring
Microtome/ stainer [34]/ scanner	Medical device (FDA class II/EU IVDR)	ISO 13485 (QMS)/IEC 62304 (software)/ IEC 60601-1 (electrical safety)	Design control/risk management/ process validation	Periodic revalidation/ QC logs/CAP/CLIA inspections
Image management/ WSI server [35]	SaMD	DICOM-WSI/HL7/FHIR/ ISO 14971 (risk) [36]/IEC 62366 (usability)	Metadata interoperability/ data integrity/ cybersecurity	Audit trail/ Access control
AI algorithm (diagnostic/ quantitative)	SaMD (FDA/CE-IVD)	GMLP/TRIPOD-AI/ CONSORT-AI/DECIDE-AI/ PCCP guidance [37]	Model validation/ bias assessment/ change control plan	Continuous learning documentation/ drift monitoring
Integrated DP system	Networked medical system	ISO/IEC 27001 (information security), IEC 80001 (risk for IT networks) [38, 39]	Cybersecurity/ encryption/ traceability	Cyber risk audit/ performance surveillance

AI – artificial intelligence, CAP – College of American Pathologist, CLIA – clinical laboratory improvement amendments, DICOM – digital imaging and communications in medicine, DP– digital pathology, EU IVDR – European Union's In Vitro Diagnostic Regulation, FDA – Food and Drug Administration, GMLP – Good Machine Learning Practice, HL7/FHIR – Health Level Seven – Fast Healthcare Interoperability Resources, PCCP – Predetermined Change Control Plan, QC – quality control, SaMD – software as a medical device, WSI – whole-slide images

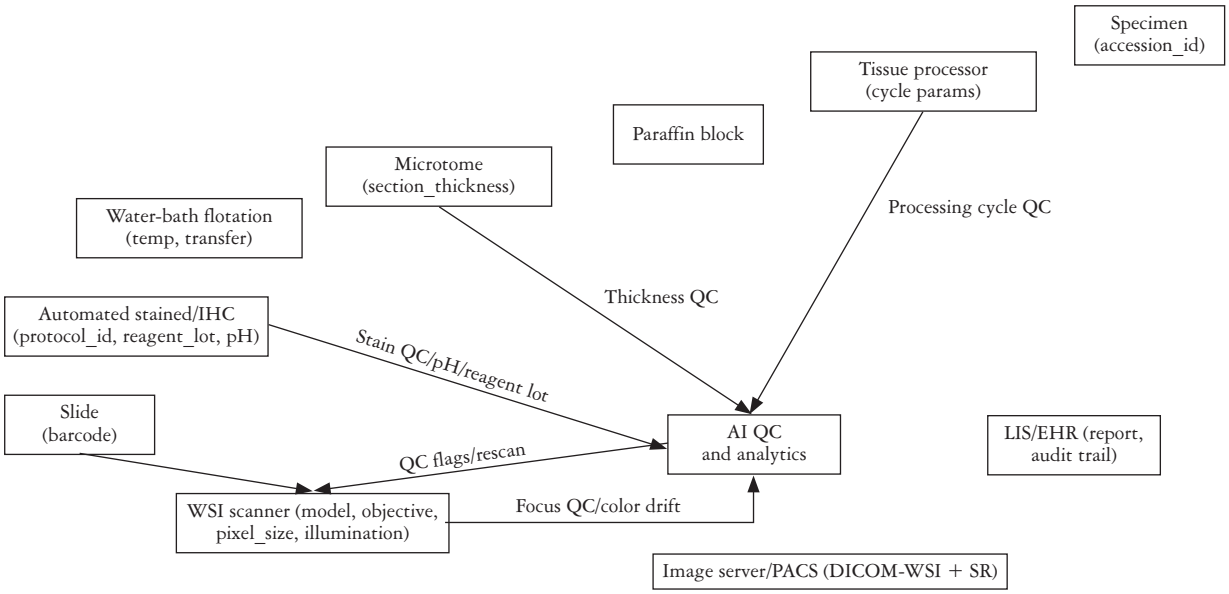

Figure 2. Reference architecture of metadata and quality-control data flow across the digital pathology pipeline

AI – artificial intelligence, EHR – electronic health record, IHC – immunohistochemistry, LIS – laboratory information system, PACS – picture archiving and communication system, QC – quality control, WSI – whole-slide images

extensions and transferred using Health Level Seven – Fast Healthcare Interoperability Resources (HL7/ FHIR) interfaces, enabling longitudinal traceability across the tissue lifecycle. AI modules can then perform continuous QC – detecting focus deviations, color drift, tissue folds, or tissue loss – and generate automated alerts for rescanning or recalibration.

This architecture integrates hardware-derived metadata, digital traceability, and AI-based QC into a single ecosystem, supporting more reliable clinical deployment by ensuring that device performance and AI outputs are evaluated within the same feedback loop [26].

Regulatory and standardization implications

As DP enters regulated clinical use, alignment with established engineering and software standards becomes essential. Scanners, image servers, and AI algorithms are increasingly regulated as SaMD [27] and are expected to comply with multi-layer frameworks, including:

- ISO 13485 – medical-device quality management [28],
- IEC 62304 – software life cycle and validation [29],
- ISO 15189 – laboratory accreditation and traceability [30],
- FDA Good Machine Learning Practice (GMLP) [31],

- European Union's *In Vitro* Diagnostic Regulation (EU IVDR) and medical device regulation – device classification and post-market vigilance [32].

Table III summarizes these frameworks in a regulatory-technical matrix linking device classes with their associated standards, cybersecurity expectations, and surveillance obligations [33]. Integrating these requirements ensures that engineering quality and AI safety are governed under a unified structure, supporting consistent and auditable clinical-grade DP systems.

Integration and future prospects

Digital pathology is increasingly functioning as a coordinated cyber-physical ecosystem in which instruments, metadata standards, and analytics operate jointly. Practical gains in reproducibility arise when device metadata, QC logs, and automated analytics are linked so that hardware issues are identified and corrected before they reach the diagnostic stage [40]. The following subsections outline how synchronized metadata, federated QC, and integrated system design can support cross-site reliability.

Workflow synchronization through device metadata

Each component of the histology workflow – from tissue processing to scanning – generates metadata reflecting its operational state. Harmonizing these metadata streams under DICOM-WSI and HL7/FHIR enables synchronized workflow control and real-time QC feedback [41]. Standardized metadata fields, implemented either through DICOM extensions or consistent laboratory information system (LIS) configurations, provide a practical foundation for maintaining cross-laboratory consistency while imposing minimal disruption to established workflows.

Cloud-based collaboration and federated quality control

Cloud annotation environments and federated learning frameworks now support collaborative model development without centralizing patient data [42]. Hardware-related metadata – such as section thickness, reagent batch, and scanner color profile – serve as key anchors for normalization and cross-site performance auditing. In federated settings, local QC data contribute to a shared dashboard, allowing laboratories to jointly monitor mechanical and computational performance.

Convergence toward hardware-digital co-design

Future DP systems will increasingly adopt integrated hardware-software co-design, rather than treating mechanical and computational layers separately. Embedding self-calibrating sensors, predictive maintenance routines, and AI-driven QC dashboards can create self-regulating laboratory environments that adapt based on accumulated operational data. This approach aligns with emerging concepts of adaptive standardization, where digital platforms recommend or adjust operational parameters to maintain diagnostic fidelity [43].

Limitations and challenges

Despite the progress described in this review, several obstacles hinder the establishment of a fully interoperable, quality-assured DP ecosystem.

Fragmented interoperability

Manufacturers often implement proprietary metadata schemas and closed interfaces, limiting integration between scanners, stainers, and LIS systems. This fragmentation restricts unified QC workflows. An international consensus on open APIs, metadata ontologies, and plug-and-play interoperability remainse urgently needed [44].

Quality and reproducibility gaps

While external quality assessment programs such as NordiQC [45] and College of American Pathologist pathology Imaging Quality Assurance Program [46] monitor staining quality, but few standards address hardware precision metrics – including section-thickness tolerance, temperature control, or scanner calibration. Without harmonized QC metadata, even well-validated AI models may show performance degradation when deployed across heterogeneous sites.

Ethical, legal, and data-governance issues

Cloud-based DP introduces challenges related to privacy, data ownership, and algorithmic accountability [47]. Incorporating hardware metadata adds another dimension – instrument traceability – which raises questions about liability in the event of mechanical error or digital failure. Future regulatory frameworks must ensure transparent audit trails linking sensor-level data, image files, and diagnostic outputs.

Regulatory fragmentation and adaptive AI oversight

Although multiple DP and AI systems have obtained regulatory clearance, global oversight remains heterogeneous [48]. The Food and Drug Administration, EU IVDR, and United Kingdom Medicines and Healthcare Products Regulatory Agency (MHRA) share the goals of safety and transparency but differ in device classification, post-market expectations, and adaptive AI governance.

Under FDA framework, scanners and image management systems are classified as medical devices, while AI modules are considered as SaMD. The Food and Drug Administrations Good Machine Learning Practice emphasizes documenting model updates, validation datasets, and rollback procedures. The European Union's *In Vitro* Diagnostic Regulation requires conformity assessment with notified bodies and detailed post-market performance follow-up, whereas the MHRA uses a hybrid model integrating device and AI lifecycle oversight.

For networked laboratory systems, cybersecurity has become essential. Both FDA and IVDR require encrypted communication, access control, audit trails, and continuous vulnerability monitoring, particularly for devices operating within hospital networks or cloud annotation platforms [49]. These measures extend quality assurance beyond image accuracy to include data integrity and system resilience.

A consolidated regulatory-technical matrix is provided in Supplementary Table S5, mapping scanners, image management systems, diagnostic AI tools, and integrated DP platforms to their respective regulatory pathways, post-market requirements, Predetermined Change Control Plan elements, and cybersecurity expectations.

Conclusions

The integration of HI with DP marks a transition from image-centric digitization to a traceable, data-driven diagnostic workflow. Embedding sensors and structured metadata capture across the histology and imaging pipeline enables laboratories to move from retrospective quality checks to real-time, predictive QC. In parallel, aligning these infrastructures with SaMD, GMLP, and IVDR requirements ensures that safety, transparency, and accountability extend from mechanical calibration through to clinical interpretation.

Future development should prioritize:

- unified interoperability standards across devices and institutions,
- closed-loop QC systems that integrate mechanical, optical, and computational data,
- regulatory readiness for adaptive, networked, and continuously learning diagnostic platforms,
- ethical governance that maintains privacy, fairness, and meaningful human oversight.

Ultimately, the future of DP will be determined not only by AI performance but also by the mechanical reliability, metadata traceability, and regulatory maturity of the systems that support it. In summary, this systematized review underscores that HI form the physical foundation upon which reliable digital and AI-assisted pathology is built. By integrating quantitative QC parameters, standardized metadata, and regulatory alignment, laboratories can advance toward a "hardware-aware AI" paradigm-linking device precision with computational accuracy to achieve reproducible and clinically safe DP workflows.

Disclosures

1. Institutional review board statement: Not applicable.
2. Assistance with the article: None.
3. Financial support and sponsorship: None.
4. Conflicts of interest: None.

References

1. Waqas A, Bui MM, Glassy EF, El Naqa I, Borkowski P, Borkowski AA, et al. Revolutionizing digital pathology with the power of generative artificial intelligence and foundation models. Lab Invest 2023; 103: 100255.
2. Deman F, Broeckx G, Declercq S, Degotte Q, Raymaekers J, Salgado R, et al. Practical implementation of AI in a non-academic, non-commercial Pathology laboratory: real world experience and lessons learned. Histopathology 2025; 87: 635-646.
3. Sandeep F, Kiran N, Rahaman Z, Devi P, Bendari A. Pathology in the age of artificial intelligence (AI): redefining roles and responsibilities for tomorrow's practitioners. Cureus 2024; 16: e56040.
4. Botros M, de Boer OJ, Cardenas B, Bekkers EJ, Jansen M, van der Wel MJ, et al. Deep learning for histopathological assessment of esophageal adenocarcinoma precursor lesions. Mod Pathol 2024; 37: 100531.
5. Wu S, Wang Y, Hong G, Luo Y, Lin Z, Shen R, et al. An artificial intelligence model for detecting pathological lymph node metastasis in prostate cancer using whole slide images: a retrospective, multicentre, diagnostic study. EClinicalMedicine 2024; 71: 102580.
6. Geread RS, Sivanandarajah A, Brouwer ER, Wood GA, Androutsos D, Faragalla H, et al. piNET-an automated proliferation index calculator framework for Ki67 breast cancer images. Cancers (Basel) 2020; 13.
7. Mohanty S, Shivanna DB, Rao RS, Astekar M, Chandrashekar C, Radhakrishnan R, et al. Building automation pipeline for diagnostic classification of sporadic odontogenic keratocysts and non-keratocysts using whole-slide images. Diagnostics (Basel) 2023; 13.
8. Yuenyong S, Boonsakan P, Sripodok S, Thuwajit P, Charngkaew K, Pongpaibul A, et al. Detection of centroblast cells in H&E stained whole slide image based on object detection. Front Med (Lausanne) 2024; 11: 1303982.
9. Yu H, Yu W, Enwu Y, Ma J, Zhao X, Zhang L, et al. Enhancing head and neck cancer detection accuracy in digitized whole-slide histology with the HNSC-classifier: a deep learning approach. Front Mol Biosci 2025; 12: 1652144.
10. Parra-Medina R, Zambrano-Betancourt C, Peña-Rojas S, Quintero-Ortiz L, Caro MV, Romero I, et al. Detection of helicobacter pylori infection in histopathological gastric biopsies using deep learning models. J Imaging 2025; 11: 226.
11. Tsuneki M, Kanavati F. Weakly supervised learning for poorly differentiated adenocarcinoma classification in gastricendoscopic submucosal dissection whole slide images. Technol Cancer Res Treat 2022; 21: 15330338221142674.
12. Cossio M, Wiedemann N, Torres ES, Sole EB, Igual L. AI-augmented pathology: the experience of transfer learning and intra-domain data diversity in breast cancer metastasis detection. Front Oncol 2025; 15: 1598289.
13. Kindler C, Elfwing S, Öhrvik J, Nikberg M. A deep neural network-based decision support tool for the detection of lymph node metastases in colorectal cancer specimens. Mod Pathol 2023; 36: 100015.

14. Cai M, Zhao K, Wu L, Huang Y, Zhao M, Hu Q, et al. Artificial intelligence-based analysis of tumor-infiltrating lymphocyte spatial distribution for colorectal cancer prognosis. Chin Med J (Engl) 2024; 137: 421-430.
15. Zhou YF, Koyuncu C, Lu C, Grobholz R, Katz I, Madabhushi A, et al. Multi-site cross-organ calibrated deep learning (MuSClD): automated diagnosis of non-melanoma skin cancer. Med Imag Analysis 2023; 84: 102702.
16. Slotman A, Xu M, Lindale K, Hardy C, Winkowski D, Baird R, et al. Quantitative nuclear grading: an objective, artificial intelligence-facilitated foundation for grading noninvasive papillary urothelial carcinoma. Lab Invest 2023; 103: 100155.
17. Anghel A, Stanisavljevic M, Andani S, Papandreou N, Rüschoff JH, Wild P, et al. A high-performance system for robust stain normalization of whole-slide images in histopathology. Front Med 2019; 6: 193.
18. Chlipala EA, Butters M, Brous M, Fortin JS, Archuletta R, Copeland K, et al. Impact of preanalytical factors during histology processing on section suitability for digital image analysis. Toxicol Pathol 2021; 49: 755-772.
19. Abd El-Aziz YM, Sobh MS, Saad HM, Morsy MHE, El-Haroun E, Abass KS, et al. Mastering the art of sectioning: a comprehensive guide to slide-microtome technology and histological applications. Front Vet Sci 2025; 12: 1635706.
20. Fera A, He Q, Zhang G, Leapman RD. Quantitative method for estimating stain density in electron microscopy of conventionally prepared biological specimens. J Microsc 2020; 277: 71-78.
21. Yagi Y. Color standardization and optimization in whole slide imaging. Diagn Pathol 2011; 6: S15.
22. Kumar N, Gupta R, Gupta S. Whole slide imaging (WSI) in pathology: current perspectives and future directions. J Digit Imaging 2020; 33: 1034-1040.
23. Faghani S, Codipilly DC, Moassefi M, Iyer PG, Erickson BJ. Optimizing storage and computational efficiency: an efficient algorithm for whole slide image size reduction. Mayo Clin Proc Digit Health 2023; 1: 419-424.
24. Salvi M, Michielli N, Salamone L, Mogetta A, Gambella A, Molinaro L, et al. Computational synthesis of histological stains: a step toward virtual enhanced digital pathology. Int J Imaging Syst Technol 2024; 34.
25. Eccher A, Marletta S, Sbaraglia M, Guerriero A, Rossi M, Gambaro G, et al. Digital pathology structure and deployment in Veneto: a proof-of-concept study. Virchows Arch 2024; 485: 453-460.
26. Belloni E, Bonoldi E, Bovo G, Buoro S, Cerati M, Cribiú FM, et al. Digital transition in pathology lab: a survey from the Lombardy region. Pathologica 2024; 116: 232-241.
27. Bouyssoux A, Jarnouen K, Lallement L, Fezzani R, Olivo-Marin JC. Automated staining analysis in digital cytopathology and applications. Cytometry A 2022; 101: 1068-1083.
28. Humphries MP, Kaye D, Stankeviciute G, Halliwell J, Wright AI, Bansal D, et al. Development of a multi-scanner facility for data acquisition for digital pathology artificial intelligence. J Pathol 2024; 264: 80-89.
29. Health C for D and R. Fda. FDA 2025. Software as a medical device (SaMD). Available from: https://www.fda.gov/medical-devices/digital-health-center-excellence/software-medical-device-samd (accessed: 10.10.2025).
30. Iso [Internet]. Iso 13485:2016. Available from: https://www.iso.org/standard/59752.html (accessed: 10.10.2025).
31. Iso [Internet]. Iec 62304:2006. Available from: https://www.iso.org/standard/38421.html (accessed: 10.10.2025).
32. Available from: https//www.tuvsud.com/en-us/industries/healthcare-and-medical-devices/medical-devices-and-ivd/quality-management-and-quality-control-for-medical-devices (accessed: 10.10.2025).
33. Available from: https//www.fda.gov/medical-devices/software-medical-device-samd/good-machine-learning-practice-medical-device-development-guiding-principles (accessed: 10.10.2025).
34. Anghel A, Stanisavljevic M, Andani S, Papandreou N, Rüschoff JH, Wild P, et al. A high-performance system for robust stain normalization of whole-slide images in histopathology. Front Med 2019; 6: 193.
35. Rosenzweig C. FDA guidance on cybersecurity for medical devices [Internet]. Regulatory knowledge for medical devices. Available from: https://blog.johner-institute.com/iec-62304-medical-software/fda-guidance-on-cybersecurity/ (accessed: 10.10.2025).
36. Iso [Internet]. Iso 14971: 2019. Available from: https://www.iso.org/standard/72704.html (accessed: 10.10.2025).
37. Health C for D and R. Predetermined change control plans for machine learning-enabled medical devices: guiding principles. Fda [Internet]. Available from: https://www.fda.gov/medical-devices/software-medical-device-samd/predetermined-change-control-plans-machine-learning-enabled-medical-devices-guiding-principles (accessed: 10.10.2025).
38. Iso [Internet]. Iso/iec 27001: 2022. Available from: https://www.iso.org/standard/27001 (accessed: 10.10.2025).
39. Iso [Internet]. Iec 80001-1: 2021. Available from: https://www.iso.org/standard/72026.html (accessed: 10.10.2025).
40. Matias-Guiu X, Temprana-Salvador J, Garcia Lopez P, Kammerer-Jacquet SF, Rioux-Leclercq N, Clark D, et al. Implementing digital pathology: qualitative and financial insights from eight leading european laboratories. Virchows Arch Int J Pathol 2025; 487: 815-826.
41. Makhlouf Y, Salto-Tellez M, James J, O'Reilly P, Maxwell P. General roadmap and core steps for the development of AI tools in digital pathology. Diagn Basel Switz 2022; 12: 1272.
42. Alajaji SA, Khoury ZH, Elgharib M, Saeed M, Ahmed ARH, Khan MB, et al. Generative adversarial networks in digital histopathology: current applications, limitations, ethical considerations, and future directions. Mod Pathol Off J U S Can Acad Pathol Inc 2024; 37: 100369.
43. Puget C, Ganz J, Ostermaier J, Conrad T, Parlak E, Bertram CA, et al. Artificial intelligence can be trained to predict c-KIT-11 mutational status of canine mast cell tumors from hematoxylin and eosin-stained histological slides. Vet Pathol 2025; 62: 152-160.
44. Bumgardner VKC, Armstrong S, Virodov A, Hickey C. Automated curation and AI workflow management system for digital pathology. AMIA Jt Summits Transl Sci Proc AMIA Jt Summits Transl Sci 2023; 2023: 71-80.
45. Vyberg M, Nielsen S, Bzorek M, Røge R. NordiQC assessments of synaptophysin immunoassays. Appl Immunohistochem Mol Morphol AIMM 2021; 29: 643-647.
46. College of American Pathologists [Internet]. Frequently asked questions from immunohistochemistry laboratories. Available from: https://www.cap.org/member-resources/councils-committees/immunohistochemistry-topic-center/frequently-asked-questions-from-immunohistochemistry-laboratories (accessed: 10.10.2025).
47. Redrup Hill E, Mitchell C, Brigden T, Hall A. Ethical and legal considerations influencing human involvement in the implementation of artificial intelligence in a clinical pathway: a multi-stakeholder perspective. Front Digit Health 2023; 5: 1139210.
48. 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: 79-85.
49. Eccher A, Marletta S, Sbaraglia M, Guerriero A, Rossi M, Gambaro G, et al. Digital pathology structure and deployment in veneto: a proof-of-concept study. Virchows Arch Int J Pathol 2024; 485: 453-460.

Address for correspondence

Tiantian Li, PhD
Department of Pathology and Experimental Medicine
Graduate School of Medicine
Dentistry and Pharmaceutical Sciences
Okayama University
2-5-1 Shikata, Kita-ku
Okayama, 700-8558, Japan
e-mail: opheliali2012@163.com