Computational Biomedicine

Science Exploration

Open Access Review

Advances in the coordinated development of digital cytopathology testing equipment and intelligent interpretation systems

JunNan Wang[1], QianYao Lin[1], YiYi Fu[2], JianXiong Guan[3], HongTao Li[4], JingJing Fan[4], XiaoFan Deng[2], ShiQi Liu[5], Fa Zhang[6], DongXu Zhang[1,*]

[1]School of Public Health, Xiamen University, Xiamen 361102, Fujian, China.
[2]School of Mechanical Engineering, Tianjin University of Technology and Education, Tianjin 300222, China.
[3]Xiamen Mimai Technology Co., Ltd., Xiamen 361102, Fujian, China.
[4]The Affiliated Cancer Hospital of Xinjiang Medical University, Urumqi 830011, Xinjiang, China.
[5]Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China.
[6]School of Medical Science and Engineering, Beijing Institute of Technology, Beijing 100081, China.

***Correspondence to:** DongXu Zhang, School of Public Health, Xiamen University, Xiamen 361102, Fujian, China. E-mail: zhangdongxu@xmu.edu.cn

Received: August 14, 2026 **Accepted:** October 30, 2026 **Published:** October 30, 2026

Cite this article: Wang J, Lin Q, Fu Y, Guan J, Li H, Fan J, et al. Advances in the coordinated development of digital cytopathology testing equipment and intelligent interpretation systems. Comput Biomed. 2026;1:202626. https://doi.org/10.70401/cbm.2026.0026

Abstract

Digital cytopathology links specimen preparation, staining, imaging, quality control, and artificial intelligence (AI)-assisted interpretation. This narrative review compares how equipment conditions, annotation levels, and aggregation mechanisms affect model performance and its clinical interpretation. Recent studies cover cervical screening as well as thyroid, urinary, pancreatic, and effusion cytology, but differences in specimen types, reference standards, and validation designs limit direct comparisons. We explain detection and segmentation, weakly supervised aggregation, self-supervised adaptation, and multimodal fusion, distinguishing technical endpoints from reader performance and patient benefit. Preparation-dependent morphology, sparse abnormal cells, overlapping groups, and focal-plane selection constrain transfer between laboratories. An equipment–quality control–AI–human review framework can organize error-specific corrective actions, but its overall safety and cost-effectiveness require prospective evaluation. Deployment must match the validated specimen, population, and assistive or autonomous role.

Keywords: Cytopathology testing equipment, digital cytopathology, artificial intelligence, AI-assisted interpretation, human-AI collaboration

1. Introduction

Cytopathology examines dispersed cells, clusters, and background material from exfoliative, body-fluid, and aspiration specimens. Cervical screening illustrates its relationship with complementary human papillomavirus (HPV) testing[1,2]. Automated preparation, staining, and whole-slide imaging (WSI) increasingly connect these specimens to artificial intelligence (AI)-assisted interpretation[3-5]. Equipment determines which cells and morphological features reach a model. Cell recovery, fixation, staining, focal coverage, and retained material for ancillary tests therefore belong within algorithm evaluation, rather than serving merely as acquisition details.

Existing reviews cover cervical deep learning[6], diagnostic accuracy in digital pathology[7], and public evidence for histology AI products[8]. Direct cytology applications include pleural-effusion slide classification[9], patch classification[10], and malignant-cell enrichment[11], which have distinct endpoints. We examine how equipment shapes model inputs, supervision and aggregation, and how acquisition failures affect validation and human review.

Figure 1 presents the proposed workflow. Table 1, Table 2, and Table 3 summarize equipment variables, measurements, and failure-specific corrective actions.

1.1 Literature search and review strategy

This narrative review was supplemented by targeted PubMed and Web of Science searches, last conducted on 29 October 2026, without publication-year or language restrictions. Search terms combined cytopathology/cytology with artificial intelligence, deep

© The Author(s) 2026. This is an Open Access article licensed under a Creative Commons Attribution 4.0 International License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, sharing, adaptation, distribution and reproduction in any medium or format, for any purpose, even commercially, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

learning, machine learning, and foundation models, with additional searches addressing preparation, digital imaging, quality control, and clinical validation. Sources were selected for their relevance to cytology preparation, staining, digital imaging, image quality control, AI-assisted interpretation, or clinical validation and governance. Relevant primary cytopathology studies were prioritized; reviews, official guidance, and selected histopathology studies provided contextual or methodological comparisons. Sources outside these topics and redundant reports supporting the same claim were excluded. These searches were targeted rather than exhaustive and did not constitute a systematic review. Evidence was compared by subtopic, study design, and endpoint. Table S1 assigns each of the 105 cited sources to one of six mutually exclusive themes according to its principal role. The sources include empirical studies, datasets, methodological comparators, reviews, guidance, and regulatory or product information. Descriptive counts of the cited papers are based on the 61 original empirical cytology studies and on each paper's bibliographic publication year; they describe this cited sample and do not estimate publication volume or study quality in the field.

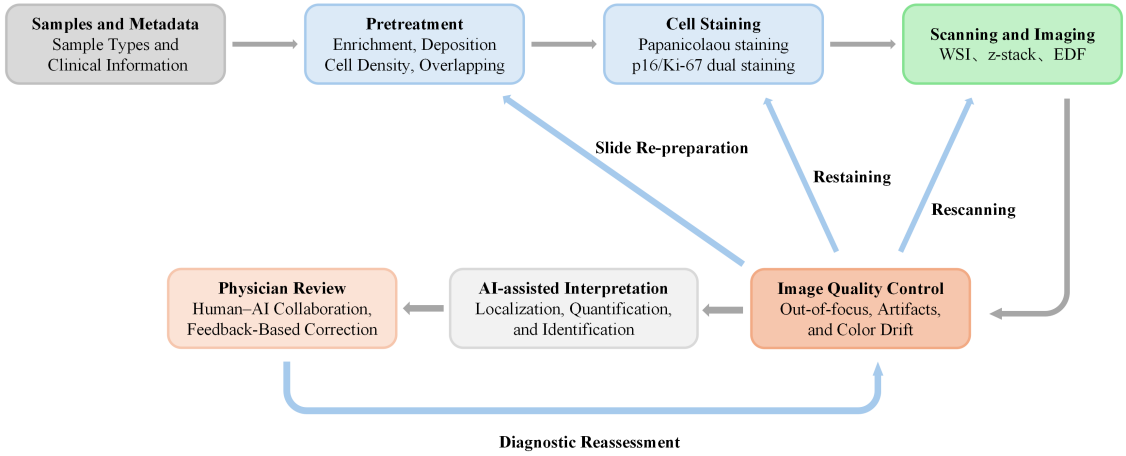

Figure 1. Proposed workflow integrating cytopathology preparation, staining, digital imaging, quality control, AI-assisted interpretation, and physician review. Feedback pathways indicate repeat preparation, restaining, rescanning, and diagnostic reassessment as appropriate. The framework is conceptual. AI: artificial intelligence; WSI: whole-slide imaging; EDF: extended depth of field.

Table 1. Major equipment technology approaches and their impact on AI-based interpretation in digital cytopathology.

Workflow module	Approach	Equipment-dependent features	Implications for AI	Validation priorities
Preparation	Conventional smears; filtration- or sedimentation-based LBC[3,12]	Cell recovery, density, overlap, background	Preparation-dependent inputs affect retrieval and transfer	Abnormal-cell retention; overlap; residual material
ICC staining	Lung biomarker ICC[13]	Assay-dependent signal and background	Signal variation can bias marker quantification	Assay validation; controls; reagent traceability
Dual staining	p16/Ki-67[14]	Same-cell p16/Ki-67 coexpression	Positive-cell counting requires platform validation	Platform, threshold, reproducibility
Ancillary testing	Cell blocks; lung ICC[13]; pancreatic ancillary assays[4]; HPV testing[3]	Material allocation and preservation	Linked, prediction-time measurements can support fusion	Assay compatibility; linkage; provenance
Imaging	Single plane; z-stack; EDF[15]	Resolution, focus, coverage, depth	Focal coverage constrains retrieval and quantification	Speed/storage trade-off; retained source planes

References support the specified preparations, tests, and imaging approaches; AI implications and validation priorities reflect the authors' synthesis. AI: artificial intelligence; LBC: liquid-based cytology; HPV: human papillomavirus; ICC: immunocytochemistry.

2. Cytopathology Sample Preprocessing, Staining, and Ancillary Testing

2.1 Cytopathology sample preprocessing and automated slide preparation equipment

2.1.1 Preprocessing equipment and cytopathology data quality

Preprocessing determines cellular morphology, background, distribution, and specimen adequacy, affecting both manual review and

AI-assisted interpretation. Conventional smears are inexpensive and familiar, but operator-dependent thickness, fixation, and obscuring blood or mucus can complicate diagnosis and reduce model generalizability. Preprocessing equipment should therefore be evaluated as part of the image data generation pipeline.

Table 2. Reported quantitative evidence and recommended reporting metrics for cytopathology equipment and AI inputs.

Part A. Quantitative findings reported by individual studies		
Comparison/study	Study material and technical setting	Quantitative findings
SurePath versus ThinPrep[3]	2,005 women; preparations/residual HPV material	Unsatisfactory: 0.2% versus 1.5%; insufficient residual HPV material: 0% versus 18.2%.
Single plane versus EDF[16]	P1000; dry 3-plane EDF versus water-immersion single plane; 40×, 0.12 µm/pixel	Mean dry EDF/water single-plane time: ThinPrep 7 min 51 s/2 min 40 s; SurePath 3 min 28 s/1 min 7 s. Mean size: ThinPrep 825.92/764.30 MB; SurePath 443.89/412.87 MB. Both immersion and plane count changed.
Whole-slide edge tomography[17]	Experimental setup; 10/20/40 planes	Approximately 3/4.5/8 min for 10/20/40 planes, excluding loading/setup. Optical resolution: 0.22 µm lateral, 1 µm axial.
Urine scanner comparison[18]	Cytospin/ThinPrep; Aperio AT2/NanoZoomer S360; timed scans: 21 planes, 1 µm spacing, manual focus	Median Cytospin/ThinPrep time: AT2 1,010/5,357 s; S360 275/1,429 s. Median size: AT2 2.0/13.1 GB; S360 4.4/25.5 GB. Single-plane default-focus success 85.7-92.9%; manual focus 14/14 in each scanner–preparation combination.
MCAS prototype[19]	16 lung FNA slides; 54 × 72 mm field	Three slides: approximately 5 min at 0.3 NA or 40 min at 0.5 NA, with different axial steps. Resolution: 1.2 µm (0.3 NA), 0.6 µm (0.5 NA).
Part B. Recommended reporting metrics		
Reporting domain	**Recommended metrics and documentation**	
Adequacy and recovery	Unsatisfactory rate; abnormal-cell recovery/loss; overlap; diagnostic-area coverage; adequacy at operating speed	
Throughput and reliability	Slides/hour; hands-on/loading/setup time; downtime; failures; repeat preparation; missed fields; rescans	
Focal strategy	Thickness/task; plane count/spacing; focus confidence; fusion method; retained planes	
Resolution and storage	Pixel size; optical resolution; usable/scanned area; compression; raw/compressed storage	
Quality control	Defocus; bubbles; stitching/fusion/stain artifacts; false rejection; corrective actions	
Metadata and exchange	Specimen/operator; fixation/timing; platform/lot/volume; scanner/software; calibration; focus/scan logs; provenance; DICOM/export	
AI and domain shift	Rare-cell recall; background robustness; QC sensitivity; calibration; repeatability; sensitivity at throughput; preparation/stain/scanner/focus/software strata	

Part A reports study-specific measurements; Part B synthesizes reporting domains from implementation and validation guidance[5,15], without prescribing thresholds. Scanning success used paired preparations from 14 specimens; time/file-size comparisons used 10 quality-passing pairs[18]. Single-plane success is not 21-plane success. CI: confidence interval; DICOM: digital imaging and communications in medicine; EDF: extended depth of field; FNA: fine-needle aspiration; LBC: liquid-based cytology; MCAS: multi-camera array scanner; NA: numerical aperture; QC: quality control; WSI: whole-slide image.

2.1.2 Liquid-based cytology slide preparation systems

Liquid-based cytology (LBC) uses preservation, mixing, enrichment, and controlled deposition to produce concentrated cell preparations and cleaner backgrounds while retaining material for HPV testing, dual staining, or other ancillary tests[3,12]. HPV testing is molecular; cytology and HPV testing can use the same preserved specimen.

ThinPrep uses filtration and membrane transfer. SurePath uses enrichment and sedimentation. The platforms differ in cell recovery, background removal, cluster preservation, and compatibility with ancillary tests[3]. Preparation area, cell density, and background

affect scanning and detection. Disease detection requires evaluation alongside adequacy.

Table 3. Image-quality failure modes, AI consequences, and corrective actions.

Phase	Failure	AI consequence	QC signal	Corrective action	Escalation/audit
Preanalytical	Low cellularity/inadequacy	Missed targets; unstable case score	Adequacy, counts, metadata	Repeat collection/preparation if feasible; otherwise non-assessable	Pathologist disposition; adequacy reason
Preanalytical	Overlap, mucus, blood, debris	Segmentation/retrieval error	Overlap/background fraction; QC alert	Reprocess or repeat; targeted review if recoverable	Rejection/repeat/override rates
Analytical/staining	Stain drift/nonspecific deposition	Domain shift; intensity/calibration error	Batch controls; color statistics	Restain or validated normalization	Batch/lot and normalization version
Imaging	Defocus, bubbles, coverslip defect	Morphology loss; missed cells	Focus/artifact/confidence maps	Correct coverslip; rescan; reprepare if unrecoverable	Abstain/review; retain both images
Imaging	Missed field/misregistration	Omitted cells; distorted relationships	Coverage; residuals; coordinates	Targeted rescan/re-stitch	Retain coordinates; review prior result
Computational	Compression, enhancement or EDF artifacts	Artificial morphology; biased measurements	Provenance; raw-plane comparison	Validated reprocessing or original-plane review	Retain source, processed image, version, decision
Model output	OOD, miscalibration, uncertainty	Unsafe threshold/prioritization	Uncertainty/OOD flags; drift	Abstain, review, escalate	Model/QC/evidence/override/final diagnosis

Author-proposed framework informed by implementation, validation, and image-quality studies[15,20-27]. Thresholds require intended-use validation. Normalization cannot restore absent staining signals or remedy nonspecific deposits. Preserve source images, corrective-action history, overrides, and final disposition; monitor false rejection, repeat preparation, and rescanning. OOD: out of distribution; QC: quality control; EDF: extended depth of field; AI: artificial intelligence.

2.1.3 Cell enrichment, filtration, and controlled deposition

Centrifugation, sedimentation, and membrane filtration can concentrate scarce target cells in the area examined. Enrichment must balance abnormal-cell recovery and background clearance against cell loss, overlap, cluster disruption, and filter clogging.

Dense deposits complicate segmentation, while sparse deposits require a wider search for abnormal cells. Deposition area, cell density, and background control should reflect specimen type and cellularity.

2.1.4 Multipurpose sample preservation and ancillary testing

Limited cytology specimens, including fine-needle aspirates and effusions, must be allocated among morphological examination, cell blocks, immunocytochemistry, and molecular testing. Cell blocks preserve cell clusters and some spatial relationships while supporting immunomarker assessment and selected molecular tests. In lung-cancer cytology, specimen allocation and preservation should account for the requirements of predictive biomarker testing[13]. Liquid-based cytology can improve sample utilization after endoscopic ultrasound-guided pancreatic sampling[4]. Preprocessing should preserve material for both diagnosis and ancillary testing.

2.1.5 Summary of preprocessing equipment

Table 1 compares preparation approaches. Relevant measures include adequacy, abnormal-cell recovery, overlap, batch consistency, and residual material for ancillary tests. Table 2 distinguishes reported measurements (Part A) from recommended metrics (Part B).

Adaptive preparation could use specimen cellularity and process measurements to adjust enrichment and deposition. This remains a development proposal that requires testing for abnormal-cell retention and downstream diagnostic performance. Processing-fluid and augmentation experiments in the MKN45 gastric cancer cell line illustrate controllable input variation, not patient-level validation[28].

2.2 Automated staining, dual staining, and ancillary testing equipment

2.2.1 Automated staining equipment and image consistency

Papanicolaou staining, hematoxylin and eosin (H&E) staining of cell blocks, and immunocytochemistry shape the color and texture seen by a model. Automated systems standardize staining, washing, dehydration, and coverslipping to reduce variation between operators.

Staining consistency affects model inputs: variation across batches or centers may encode laboratory identity rather than lesion characteristics. Platforms should be evaluated for color stability, inter-batch consistency, reagent traceability, quality-control alerts, and scanner compatibility. Integrating staining equipment with image quality-control modules could identify fading, excessive background, or nonspecific deposition before interpretation. Endometrial work evaluated cell-clump segmentation and classification using H&E and Papanicolaou images[29]; this does not establish general cross-stain equivalence.

2.2.2 p16/Ki-67 dual staining and cell-level risk stratification

p16/Ki-67 co-expression marks dysregulated cell-cycle activity and supports triage after a positive HPV test. World Health Organization (WHO) and American Society for Colposcopy and Cervical Pathology recommendations define this role[30,31] within risk-based management[32,33]. The PALMS screening study[34], the IMPACT trial of HPV-positive triage[35], and longitudinal risk assessment after dual-stain testing[36] address different aspects of clinical validity.

Compared with conventional staining, dual staining converts some subjective morphological judgments into localizable, countable, and reviewable cell-level signals. Clinical validity concerns the association between assay results and disease risk[31,34-36]. Analytical validity of automated interpretation concerns positive-cell detection, classification, and reproducibility, whereas clinical utility concerns the consequences of using the reader in the intended workflow. Wentzensen *et al*. developed CYTOREADER for automated p16/Ki-67 dual-stain assessment, using manually annotated image tiles to train and validate positive-cell recognition, followed by slide-level evaluation against disease endpoints[14]. The study used preparation-specific models: a four-layer convolutional neural network performed well on ThinPrep slides but poorly on SurePath slides, prompting the development of a separate Inception-v3 model for SurePath. In the Kaiser validation cohort of 3,095 HPV-positive women evaluated using SurePath preparations, a threshold of at least two dual-stain-positive cells yielded similar sensitivity for cervical intraepithelial neoplasia grade 3 or worse (CIN3+) to Pap cytology (88.1% versus 85.8%) and a lower colposcopy referral proportion (41.9% versus 60.1%). These findings support automated dual-stain assessment in the evaluated settings. Robustness to staining drift, overlap, and sparse positive cells, and the effects on routine patient management, still require evaluation for each intended implementation.

2.2.3 Integration of immunocytochemistry, cell blocks, and molecular testing

Immunocytochemistry and cell blocks provide information on cell origin, lesion characteristics, and treatment-related biomarkers, particularly in limited effusion and lung, lymph-node, or pancreatic aspiration samples[4,13]. Equipment should support routine morphology, cell block sections, immunomarker testing, and material preservation for molecular analysis.

Future models could combine Papanicolaou or cell block H&E images with special stains, immunomarkers, HPV results, genetic variants, and clinical history available at prediction time. Linking biomarker images to positive-signal locations, stain batches, and quality-control records would support multimodal analysis.

3. Digital Cytopathology Imaging and Image Data Quality

3.1 Digital cytopathology scanners and multifocal-plane imaging equipment

Cytology digitization must capture sparse diagnostic cells across large areas and uneven focal depths. Dispersed cells and clusters may be obscured by mucus, inflammatory cells, erythrocytes, or necrotic debris, placing high demands on focus, coverage, and local image quality. The 2022 College of American Pathologists WSI validation guideline excludes cytology[20]; a cytology-specific review instead documents validation and multifocal-imaging challenges[21]. Single-plane WSI therefore requires validation for each specimen and task.

Z-stacks can improve visualization of thick or overlapping groups at the cost of acquisition time and storage[15,21]. Single-plane imaging may suffice for thin preparations after task-specific validation. Extended-depth-of-field fusion simplifies navigation but can alter morphology and depth relationships. Lee *et al*. combined multifocal imaging on three scanners with focus fusion and color normalization for thyroid fine-needle aspiration (FNA) classification[37]. This design could not isolate the contribution of multifocal acquisition.

A multi-camera array scanner (MCAS) demonstrated rapid three-dimensional cytology imaging through parallel acquisition[19]. Stability, cost, and workflow compatibility remain relevant to adoption. Differences in numerical aperture and lateral and axial sampling limit direct speed comparisons.

Focus failures, bubbles, incomplete coverage, and compression artifacts may remove information needed to detect sparse abnormal cells. FDA guidance on WSI technical performance describes component- and system-level testing for surgical pathology; it does not establish cytology-specific clinical validity[22]. Cytopathology implementation recommendations and proficiency testing provide the relevant specimen-specific context[15,21,23]. We recommend recording focus maps, coverage, compression settings, and rescanning events so that failures can be linked to corrective actions.

3.2 Computational imaging algorithms and intelligent quality-control models

Whole-slide imaging uses stitching, focus control, and quality assessment. Multifocal workflows may also use depth fusion. Registration, illumination correction, and blending join adjacent fields. Drift, uneven lighting, or sparse landmarks can duplicate cells or distort clusters. Flag uncertain regions and retain their original coordinates.

Image stitching commonly uses feature matching, phase correlation, and random sample consensus (RANSAC) to estimate translation, rotation, or local deformation, followed by global coordinate optimization. Evaluation should address cell locations, cluster morphology, and background relationships as well as visible seams.

Focus assessment includes handcrafted sharpness measures and learned defocus classifiers. Hosseini *et al.* developed a no-reference focus-quality method for WSI[24]. ConvFocus, a separate convolutional neural network (CNN)-based method trained with semi-synthetic defocus, localized out-of-focus regions and examined their effect on cancer detection in histopathology[25]. Neither study alone establishes cytopathology performance. Focus maps should therefore be evaluated against diagnostically relevant cells, including sparse targets and overlapping groups, before they are used to trigger rescanning.

Extended-depth-of-field imaging combines sharp regions across planes using multiscale, frequency-domain, or neural methods. Fusion can ease navigation but distort morphology, obscure depth, or sharpen debris. Preserve original planes, parameters, and uncertainty maps for complex clusters, accounting for storage of both stacks and fused images. Processing cannot recover features absent at acquisition.

3.3 Image quality control and data standardization

Image quality control should connect each failure to a detectable signal, a possible model error, and a corrective action. Table 3 presents our proposed failure-mode framework, informed by digital cytology implementation recommendations[15,21]. Kang *et al.* quantified staining quality, cell counts, and cell-cluster proportions and used XGBoost to generate comprehensive quality scores for cervical cytology WSIs[26]. Evaluation on 402 WSIs demonstrated improved efficiency and consistency in quality assessment.

Rules assess sharpness, coverage, background, staining, and scan failures. Learned detectors can flag defocus, bubbles, overlap, and staining or coverslip defects but may reject diagnostically valuable cases. Color standardization, augmentation, and domain adaptation may reduce the effects of acquisition variability. Table 3 links each failure to a corrective action. Harinath *et al.* reported slide-related errors in 21 of 909 archived ThinPrep slides (2.3%); eight were successfully rescanned, whereas 13 remained unscannable[27].

Digital imaging and communications in medicine (DICOM) WSI supports image storage and exchange[38]. We recommend linking images to specimen type, fixation and preparation method, stain batch, scanner and software version, focal-plane settings, compression, and quality-control outcomes. These metadata can support stratified validation and investigation of failures after equipment or workflow changes.

4. Intelligent Cytopathology Interpretation Models and Closed-Loop Applications

4.1 Model frameworks, evidence levels, and reliability of intelligent interpretation

Model choice starts with the input unit and label. Detection maps cells or patches to locations and classes; segmentation assigns pixels or instance boundaries; classification assigns a label to a selected cell or patch[6,39-42]. Kurita *et al.* use abnormality scores to prioritize review[41]. Feature encoders feed box/class predictors or mask decoders, with instance separation needed for overlapping cells. Nuclear masks, boxes, and case labels provide different supervision. Shape-constrained evolution addresses overlapping cytoplasm, whereas solo_GAN separates single-cell identification from boundary generation[43,44]. Generated occluded contours remain estimates, not observed boundaries. Hierarchical and detection datasets supply complementary labels[45,46]; SCAC uses labeled and unlabeled images for semi-supervised abnormal-cell detection[47]. Segmentation Dice, abnormal-cell recall with candidates per slide, and patient-level sensitivity measure different stages; none substitutes for the others.

Slide-level weak supervision, often implemented as multiple instance learning (MIL), encodes instances and aggregates their features using a bag label. Mean pooling can dilute rare positive cells; max or top-k pooling is vulnerable to artifacts. Attention-based multiple instance learning (ABMIL) learns normalized instance weights[48]; clustering-constrained attention MIL (CLAM) adds instance-clustering constraints, originally evaluated in histopathology[49]. Attention weights are neither calibrated abnormal-cell probabilities nor causal explanations. Kanavati's EfficientNet–gated recurrent unit (GRU) model uses recurrent aggregation[50];

pleural-effusion weak supervision provides another specimen-specific application[9]. Transformer aggregation models instance relationships[42]. No downstream aggregator can retrieve cells discarded before encoding. Foundation models can supply the encoder within MIL, so these categories overlap.

Self-supervision learns representations from unlabeled images through a pretraining objective; downstream adaptation then uses task labels. DINOv2-style training matches student representations to teacher targets across views and masked patches. A frozen encoder with a linear probe tests feature separability; few-shot adaptation, a trained aggregator, and full fine-tuning update different components and have different label and computing costs. Both UniCAS and CROWN use DINOv2-based pretraining; UniCAS, pretrained on 48,532 cervical WSIs, uses a shared multi-task slide aggregator, whereas CROWN evaluates a broader cytology task range[51,52]. UniCAS evaluation lacked a histological reference standard; differing task sets and adaptation protocols also preclude a simple ranking. Cervical self-supervised/MIL work addresses low-data triage[53]; contrastive normal-cell learning provides an anomaly-detection alternative[54]. CytoFM combines cytology pretraining with a frozen encoder and attention MIL[55]; thyroid self-supervision supplies organ-specific evidence[56]. Pretraining–test overlap, preparation transfer, and independent patient separation remain central checks. A histopathology review provides methodological context[57]. UNI[58], the image–text model CONCH[59], Virchow[60], and Prov-GigaPath[61] illustrate different pretraining and downstream evaluation settings. NGNet[62] and DANet[63] provide additional tissue-based methodological comparators; these results do not establish cytology accuracy.

Multimodal systems can concatenate features early, learn intermediate cross-modal interactions, or combine predictions late[64]. Thyroid ultrasound–cytology fusion is a direct application[65]. The multimodal extension mPUCAS-M combines urine-cytology model scores, radiological assessments, and clinical variables using LightGBM[66]. In contrast, a thyroid system that predicts BRAF from cytomorphology does not thereby demonstrate fusion of an independently measured BRAF result[67]. Specimen allocation for immunomarker testing[13] and image-based classification or enrichment[9-11] do not by themselves establish multimodal fusion. Fusion requires modality-specific comparators, case/time linkage, missing-input handling, and exclusion of unavailable follow-up information. Morphological knowledge–image alignment[68] is a different form of multimodality from patient-data fusion.

Nongynecologic studies address multicenter thyroid FNA diagnosis[69], AI-assisted thyroid review[70], urinary malignant-cell recognition[71], and Paris-system categorization[72]. Pancreatic studies separately investigate semi-supervised classification of endoscopic ultrasound-guided fine-needle aspiration (EUS-FNA) images[73] and two-stage rapid on-site evaluation (ROSE)[74]. The semi-supervised system uses teacher-generated pseudo-labels and moving-average teacher updates[73]; the ROSE approach localizes candidates before classification[74]. The multicenter thyroid study combines retrospective external testing with a single-center prospective cohort and mixed histology/cytology reference standards[69]. Urine studies separately evaluate AI-assisted digital review[75] and upper-tract candidate retrieval[76]. Lung classification on LBC prepared from surgical specimens cannot establish performance on routine lung FNA or brushings[77]. TORCH combines selected effusion images with clinical variables to predict organ-system origin[78], whereas effusion CLAM predicts slide malignancy[79]. Cervical studies also evaluate paired manual–digital reading[80], candidate-gallery screening[81], and diagnostic strategies in high-risk referral populations[82]. A multi-modality cervical meta-analysis is not a cytology-only estimate[83]; a nongynecologic review found thyroid and urinary studies predominated[84]. Evidence from cervical or pleural-fluid tasks should not be generalized to all aspiration specimens.

Figure 2 maps data and annotation levels to modeling approaches; Table 4 compares their mechanisms and validation needs, and Table 5 summarizes representative study designs, performance, and endpoints.

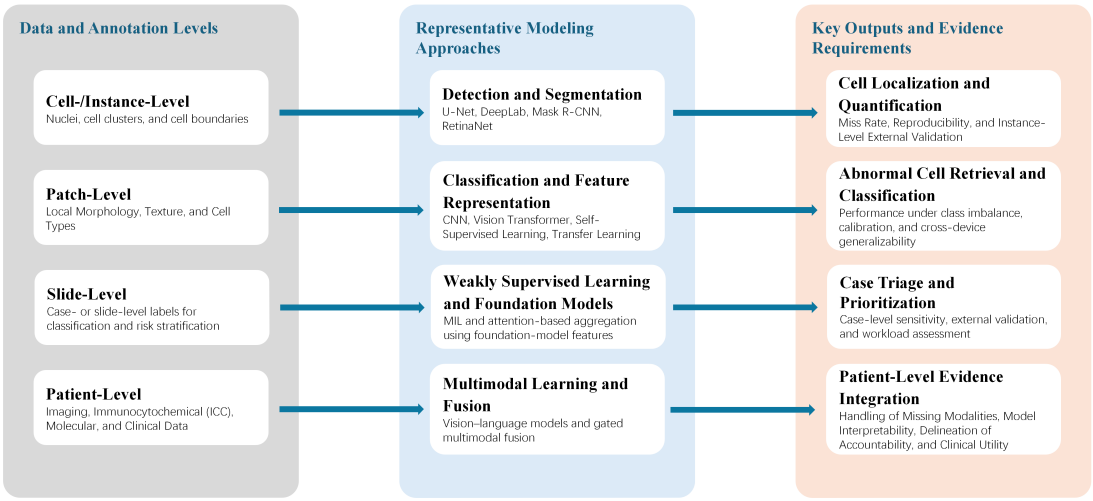

Figure 2. Data and annotation levels, representative AI modeling approaches, and key outputs and evidence requirements. Arrows indicate representative, non-exclusive associations. Self-supervised learning and foundation-model pretraining can support multiple task levels and are not synonymous with multiple instance learning.

Multimodal integration and transfer from histopathology require validation in the intended cytology setting. Model selection should reflect the task and consequences of errors. CNN: convolutional neural network; MIL: multiple instance learning; ICC: immunocytochemistry.

4.2 Coordinated design of equipment and intelligent interpretation systems and closed-loop cytopathology platforms

The US Food and Drug Administration (FDA)-authorized Genius system links ThinPrep slides, digital imaging, AI, and interpretation by qualified personnel[87]. Techcyte SureView has been evaluated for 100% quality control (QC)[16]; the manufacturer's August 2026 statement describes European IVDD CE marking, ongoing IVDR certification, and US research-only use[88]. These dated, specified uses do not validate additional feedback actions. Our proposed framework links image failures to rescanning, reprocessing, or human review (Table 3).

Local defocus may warrant rescanning, stain drift a batch investigation, and poor cell recovery reprocessing or recollection. Route unresolved failures to human review. Weigh recovery of diagnostic cells against false alerts, repeats, and delays. Track missed cells, calibration, outcomes by device, rescans, workload, disagreement, and clinical concordance.

TRIPOD+AI, STARD-AI, CONSORT-AI, SPIRIT-AI, and DECIDE-AI address prediction, diagnostic accuracy, trials, protocols, and early clinical evaluation, respectively[89-93]. These reporting guidelines support transparency but do not replace task-specific validation.

FDA lifecycle recommendations cited here remain draft guidance; the 2025 predetermined-change-control-plan guidance is final[94,95]. Monitoring should link preparation, scanner and software versions, calibration, and case mix to QC failures, overrides, and final diagnoses. Device or software changes should activate prespecified investigation and targeted revalidation, with unresolved failures routed to routine review.

Authorization is device- and use-specific; the FDA AI-device list does not extend labeling to other specimens or autonomous roles[96]. WHO guidance supports accountability and human oversight[97].

Table 4. Task matching, advantages, and limitations of different AI model approaches.

Approach	Examples	Input and supervision	Process/output	Strength	Failure/validation
Detection and segmentation	RetinaNet; YOLO; U-Net; shape constraints[6,40,42-44]	Images; boxes/masks/classes	Locations, classes, pixel/instance boundaries	Localized review; counting	Overlap; rare-cell misses; annotation variability; recall and candidate burden
Cell and patch classification	CNN; vision transformer[6,45]	Selected cells/patches; class labels	Morphology encoding to local label	Fine-grained recognition	Selection bias; limited context; no inherent slide diagnosis
Weak supervision and MIL	ABMIL; CLAM; GRU[48-50]	Instance bags; slide/case labels	Encoded features to weighted/recurrent bag prediction	Less localization annotation	Rare-positive dilution; artifact weights; discarded cells unrecoverable
Self-supervision and foundation models	UniCAS; CROWN; histology comparators[51,52,58-61]	Unlabeled images/image-text; downstream labels	Pretrain; freeze/probe, adapt, or fine-tune	Reusable features	Pretraining/test overlap; specimen transfer; calibration; computing cost
Multimodal fusion	Ultrasound-cytology; clinical/radiological fusion [64-66]	Linked images and measured patient data	Feature, interaction, or prediction fusion	Complementary information	Missing inputs; timing/leakage; image-only comparator; utility

These categories overlap: a pretrained encoder can feed MIL, and a detector can precede weak aggregation. AI: artificial intelligence; ABMIL: attention-based multiple-instance learning; CLAM: clustering-constrained attention multiple-instance learning; CNN: convolutional neural network; R-CNN: region-based convolutional neural network; GRU: gated recurrent unit.

4.3 Human–AI collaboration and responsibility

The intended role of AI determines responsibility, escalation, and audit requirements. The following safeguards are proposed for evaluation in the local workflow[16,93,98]. Triage changes case priority while preserving planned human review; the laboratory should monitor delayed or deprioritized positives and allow staff to override queue order. Autonomous negative screening permits eligible cases to bypass routine review and therefore requires a defined release authority, validated thresholds, exclusions, fallback review, and surveillance of false negatives. Low confidence or failed QC should prevent automatic release.

Table 5. Representative cytopathology AI studies and their validation evidence.

Study	Specimen	Preparation and imaging	Units and cohorts	Model and task	Validation and reference	Performance	Endpoint
Cheng 2021[39]	Cervical WSI	Membrane/sedimentation LBC; site-dependent staining; five scanners; 20×/40×, 0.180-0.293 μm/pixel	3,545 patient-wise WSIs; independent evaluation 1,170; candidate evaluation 447 positive slides	79,911 lesion-cell annotations; two ResNet50 stages + RNN; cell recommendation/WSI classification	Retrospective multicenter testing; some test centers contributed development data; three-cytopathologist consensus	Separate test (1,170): AUC 0.979; sensitivity/specificity 95.1%/93.5%. Top-10 true-lesion candidate proportion 88.5% (original TPR; 447 positive slides).	Cell retrieval/slide classification; no prospective workflow endpoint
Kanavati 2022[50]	Cervical LBC	ThinPrep Pap; Aperio AT2, 20×; SVS/JPEG2000	Train 1,503; validation 150. Clinical/equal-balance tests: 750/750 WSIs (rereview 525/643). Full-agreement test: 300 WSIs.	WSI labels; EfficientNet-B0 + GRU; neoplastic/non-neoplastic	One Japanese laboratory; retrospective; cytology labels, including rereview; no independent-center/prospective test	AUC; sensitivity/specificity: Clinical: 0.774; 81.6%/61.9% (rereview 0.890; 88.6%/90.4%). Equal-balance: 0.827; 62.4%/89.3% (rereview 0.915; 83.9%/92.0%). Full-agreement: 0.960; 85.0%/91.1%.	WSI screening; no prospective utility endpoint
Wang 2024[40]	Cervical WSI	Sedimentation LBC; stain not reported; PRECICE 600/KF-PRO-400-HI, 40×; 0.2529/0.2484 μm/pixel	16,056 participants: train 9,316; internal 2,152; external 600/600; prospective 2,780; observational trial 608	Cell boxes/WSI labels; ResNet18–RetinaNet + random forest; cytology grades	Internal/external multicenter; prospective population; randomized observational trial; cytology reference; separate histology analysis	Cytology ASC-US+ AUC: internal 0.923; external 0.929/0.879; prospective 0.947. Prospective sensitivity/specificity 94.6%/89.0%.	Assisted-reader/screening classification; not autonomous diagnosis
Xue 2025[42]	Cervical LBC	ThinPrep/SurePath/LBP; stain not reported; IBL-300/SQ-600/EasyScan NFC-60; 20×, 0.18 μm/pixel	Train 17,397; external 5,803 (468 CIN2+; reader subset 550); prospective 3,001 (41 CIN2+)/1,472 (70 CIN2+) women; one slide/woman	Cell/slide labels; YOLOv8-m + ResNet34 + Transformer MIL; CIN2+	Nine-hospital external test; readers; prospective cohorts; incomplete biopsy verification and poor-image exclusion	Internal CIN2+ clinical result not separately reported. External sensitivity 86.1% (95% CI 82.7-89.0), specificity 82.9% (81.9-83.9); prospective community 87.8%/83.1%; hospital junior-assisted 85.7%/84.0%.	CIN2+ detection/assisted workflow; verification and exclusion bias

Xie 2022[9]	Pleural effusion	CytoRich Non-gyn LBC; H&E; EasyScan 6, 40×, 0.25 μm/pixel	404 cases; one selected WSI/case; train/validation/test 266/78/60	WSI labels; weakly supervised ResNet18/MIL; benign/malignant	Retrospective internal split; cytology ± cell block/IHC; no external/prospective validation	Internal: accuracy 91.67%; sensitivity/specificity 87.50%/94.44%; AUC 0.9526.	Slide classification; no clinical follow-up
Mavropoulos 2023[11]	Body-fluid effusions	Live unstained brightfield; microfluidic sorting; immunolabeled FACS-purified cells supply labels, not inputs	37 specimen records: UCLA 34 + vendor 3; cell datasets also include lines/technical replicates	Reference cell labels; supervised CNN; carcinoma enrichment	In silico/cell validation, molecular concordance, enrichment; not diagnostic trial	Cell AUC 0.96; AI/flow-cytometry R^2 0.949; not patient sensitivity / specificity.	Enrichment/downstream testing
Nitta 2026[17]	Cervical LBC	ThinPrep/SurePath ; Pap specified at index center; 3D whole-slide edge tomography/compression	1,124 slides; four centers	Nuclear annotations; YOLOX + MaxViT; morphology-distribution triage	Multicenter evaluation includes center-specific development samples; cytological reference	Cell AUC approximately 0.99-1.00 (LSIL/HSIL/adenocarcinoma). Slide AUC across centers: LSIL+ 0.86-0.91; HSIL+ 0.89-0.97.	Autonomous-triage research; no clinical rule-out safety/outcome trial
Xue 2026[85]	Cervical LBC	ThinPrep/LBP; stain not reported; IBL-300/SQ-600; 40×, 0.18 μm/pixel	1,920 women; four centers; four non-expert readers	Cell/WSI labels; ResNet50–RetinaNet + Transformer MIL; abnormal-cell detection	Prospective randomized crossover; four-week washout; expert cytology consensus, not universal histology	Standalone AUC 0.879; sensitivity/specificity 91.8%/84.0%. Assisted sensitivity 85.7% (95% CI 81.1-89.3) versus manual 71.3% (65.8-76.3); difference 14.3 percentage points (7.6-21.1). Specificity 86.5%/85.1%.	Median reading 30/175 s; reader accuracy, not histological outcome
Wu 2024[86]	Urine WSI	LBC; Pap; five scanner types, 40×	Development 2,641; internal 1,132; external 929/127/147; prospective 400 patients (154 positive, 246 negative); discordant cohort 284	WSI labels; multistage PUCAS; urothelial carcinoma	Multicenter retrospective + prospective observational. Positive histology; negative histology or imaging/endoscopy + follow-up	AUC: internal 0.970; external 0.928-0.963; prospective 0.917 (95% CI 0.885-0.949). Prospective sensitivity 89.6% (83.7-93.9), specificity 89.4% (84.9-93.0).	Carcinoma diagnostic accuracy; improved patient outcomes unestablished
Wang 2024[69]	Thyroid FNA	Sedimentation LBC; Pap; WSI	Training: 11,254 WSIs/4,037 patients.	ThyroPower: cell classifier plus random-forest/top-	Retrospective internal/three external centers;	Prospective AUC: Bethesda III+ 0.977; V+ 0.981. Junior-	Categorization /assisted reading;

			Retrospective validation: 5,638 WSIs/2,914 patients. Prospective: 537 smears/469 patients.	feature fusion; Bethesda categorization	single-center prospective evaluation. Histology when available, otherwise expert cytology; external-center reports.	reader accuracy (537 smears): AI-assisted 94.8% (95% CI 92.6-96.5) versus manual 87.7% (84.6-90.4).	mixed references, sequential reassessment, and single-center prospective testing limit transfer.
Fang 2025[73]	Pancreatic EUS-FNA	Selected smear fields; Olympus BX53F2/Path QC; 40× objective (400× total), 2,592 × 1,944 pixels	210 patients/four hospitals. Images: training 2,567 (535 annotated); validation 134; internal/external tests 288/131.	Semi-supervised CNN; teacher/student pseudo-labels; PDAC/non-PDAC cell-cluster segmentation/classification	Retrospective image-level splits; external-center test; expert-consensus image labels; composite pathological/clinical final diagnosis. Image-level splits do not establish patient independence.	Internal/external AUC 0.97/0.99; sensitivity 93.82%/93.58%; specificity 97.27%/95.45%.	Selected-field classification; quality/diagnostic-discordance exclusions restrict applicability; not a prospective ROSE-utility trial.

AI: artificial intelligence; WSI: whole-slide imaging; LBC: liquid-based cytology; H&E: hematoxylin and eosin; AUC: area under the receiver operating characteristic curve; CNN: convolutional neural network; MIL: multiple instance learning; LBP: liquid-based preparation; TPR: true-positive rate; IHC: immunohistochemistry; FACS: fluorescence-activated cell sorting; RNN: recurrent neural network; GRU: gated recurrent unit; LSIL: low-grade squamous intraepithelial lesion; HSIL: high-grade squamous intraepithelial lesion; EUS-FNA: endoscopic ultrasound-guided fine-needle aspiration; PDAC: pancreatic ductal adenocarcinoma.

As a second reader, AI flags discrepancies for the responsible cytologist or pathologist. Log initial interpretations, alerts, decisions, and override reasons. Reviewers verify candidate cells and inspect additional areas when coverage or confidence is low. Record rejected candidates and coverage gaps. Unflagged areas may still contain abnormal cells.

For counts, ratios, and biomarkers, verify cell selection, segmentation, and denominators. Correct, repeat, or manually remeasure unreliable results and retain original and revised values. For case-level decisions integrating morphology with ancillary or clinical data, the responsible pathologist resolves discrepancies, seeks missing inputs or defers interpretation, and documents the reasons for accepting or overriding AI advice.

Nitta's 1,124-slide multicenter triage study includes center-specific development samples[17]. Its areas under the receiver operating characteristic curves (AUCs) do not establish independent validation at every center, clinical rule-out sensitivity, or prospective safety when human review is omitted.

Audit records should link preparation/scanner and model versions, QC decisions, supporting images, uncertainty, human actions, and final disposition. Evaluation should include review burden, missed detections, disagreement resolution, and patient-management consequences under the intended role.

5. Current Challenges and Future Directions

Feasibility, internal testing, independent external validation, prospective workflow evaluation, implementation, and surveillance answer different questions. Multicenter collection does not ensure independence, and prospective collection alone does not measure clinical utility. Table 5 separates design, performance, and endpoint; regulatory status remains a separate dimension.

5.1 Equipment heterogeneity and data distribution shifts

Acquisition and adaptation interact. The paired urine-scanner comparison changes preparation, scanner, and focus conditions[18], whereas MCAS protocols also differ in numerical aperture (NA) and axial step[19]; neither isolates a universal acquisition-speed advantage. The selected-field BMT ThinPrep dataset and cross-preparation adaptation study highlight preparation-specific morphology[99,100]. Evaluation after using target-domain data for adaptation or threshold selection differs from untouched-domain generalization. Pancreatic augmentation experiments compare transformation choices[101]; their effects on cross-center robustness

require separate evaluation. In archival thyroid FNA, seven-layer scanning did not improve agreement or binary accuracy over a single layer, although review time was further reduced[70]. Matched-patient comparisons across preparation/scanner settings should hold diagnostic-area coverage and resolution requirements constant and measure rare-cell recall, calibration, and threshold stability on untouched centers. Report end-to-end processing time alongside rescan, persistent-failure, and human-fallback rates, stating the denominator for each rate and retaining failed scans in the workflow assessment.

Paired or prospective studies should relate adequacy, material retention, and signal localization to disease detection, tracking abnormal-cell recovery and loss, unsatisfactory preparations, and residual material for ancillary testing[3,11-14,28,29,34-36]. Assess calibration before and after prespecified recalibration in separate temporal and cross-site tests. For uncertainty and out-of-distribution alerts, count missed abnormalities and unnecessary human referrals. For preparation, scanner, or software changes, apply the revalidation and fallback procedures in Section 4.2.

5.2 From model performance to clinical evidence

Clinical value depends on the reader and reference standard. Xue's randomized crossover trial improved non-expert cervical-reader sensitivity from 71.3% to 85.7% against expert cytology consensus, rather than histology in every participant[85]. A urine reader study included 200 cases: 98 had biopsy follow-up, of which 72 entered the ⩽ 3-month biopsy comparison; improved sensitivity/efficiency accompanied lower specificity[102]. A digital cervical workflow study likewise found faster review but lower agreement for negative for intraepithelial lesion or malignancy (NILM) cases against original sign-out[103]. Faster reading therefore need not mean net benefit. PUCAS detects urothelial carcinoma[86], whereas the muscle-invasion model concerns established carcinoma[66]; eligibility and endpoints prevent direct ranking. Prospective workflow studies or randomized crossover reader comparisons should prespecify the intended population, operating threshold, and reference standard, with procedures to address incomplete verification. Primary diagnostic or management endpoints should be selected for the intended role; missed disease, false-positive referrals, additional procedures, total workload, and failed scans should be reported alongside reading time. Assistive triage and autonomous negative release require separate comparisons with routine practice and explicit human fallback.

5.3 Foundation models, multimodal models, and tiered deployment

Model scale should match the available labels and computing budget. Frozen pretrained features may reduce task-specific annotation, but rare-cell detection, aggregation, and QC still need validation[51,52]. Thyroid self-supervision and pancreatic semi-supervision address different label shortages[56,73]. Comparisons among smaller supervised models, frozen pretrained encoders, and fine-tuning should use matched patient cohorts and label budgets, with patient-level separation and distinct reporting of within-dataset, untouched-center, and temporal tests. Calibration, rare-cell recall, storage, and inference time should accompany task accuracy. Multimodal studies should compare image-only, clinical-only, and fused models at a locked prediction time and test genuinely missing inputs. Workforce modeling estimates capacity effects[104], excluding several implementation costs; prospective comparisons should include scanning, information technology, training, QC, and downstream consequences. Tiered deployment should follow local preparation, scanning, staffing, and review capacity. Microfluidic enrichment adds a separate recovery, morphology, ancillary-test compatibility, and cost trade-off[105].

Figure 3 charts 61 cited empirical cytology papers by year and primary theme. Table S1 summarizes progress, gaps, and evaluation priorities across all 105 sources. The inclusion ledger in Data S1 identifies the papers included in the figure.

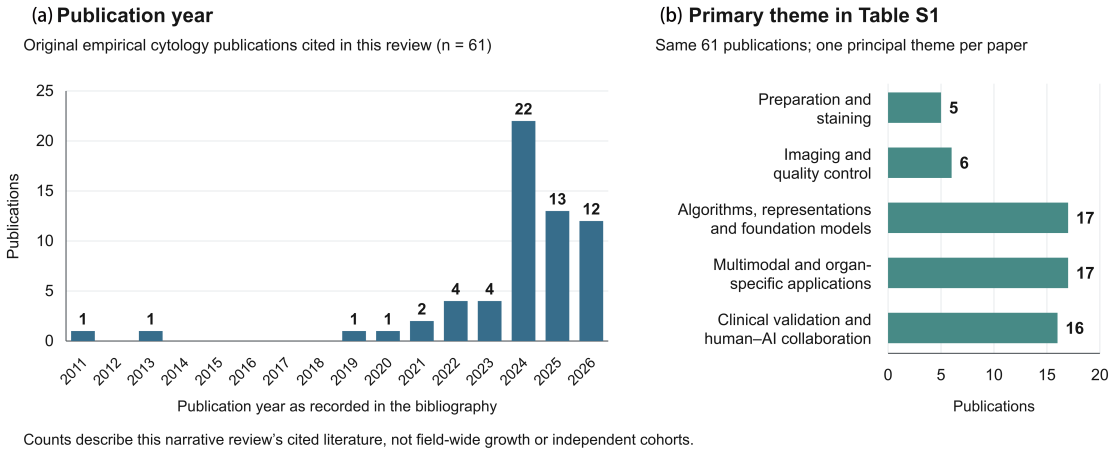

Figure 3. Distribution of original empirical cytology publications cited in this review. (a) Publication year as recorded in the bibliography; (b) Primary theme using the mutually exclusive assignment in Table S1. Both panels count the same 61 papers, which include preparation and assay studies, technical studies, datasets, and conference papers. Reviews, guidance and web resources, general or histopathology comparators without direct cytology experiments, and the workforce model are excluded. No paper falls in

the sixth theme of Table S1 (standards, reporting, regulation, and governance), so panel B shows five categories. Theme assignment is descriptive rather than a measure of validation quality. Source data and the full inclusion ledger are provided in Data S1. The 2026 bar is partial: literature searches closed on 29 October 2026.

5.4 Limitations of this review

Targeted searches may favor accessible and positive studies. We undertook no formal risk-of-bias assessment. Differences in specimens, reference standards, sample units, and designs limit pooling and ranking. Some datasets or cohorts may overlap. Figure 3 and Table S1 describe the cited literature. Findings and product status reflect the update date.

The proposed equipment–QC–AI–human feedback framework lacks end-to-end clinical validation. Evidence for individual components does not establish the safety, patient benefit, or cost-effectiveness of additional feedback actions. Image fusion may alter morphology; task-specific validation and access to acquired planes remain necessary.

6. Conclusions and Future Perspectives

Evidence supports selected abnormal-cell retrieval, slide classification, and assisted-reading applications, with expanding nongynecologic and representation-learning research[17,39-42,50-52,84,85]. Equipment conditions, supervision, and aggregation jointly determine the available evidence and likely failures.

Prospective comparison with routine practice should measure errors, workload, costs, and management consequences. Deployment should match the validated specimen and role, with explicit responsibility, fallback procedures, and surveillance. Assistive reading and autonomous rule-out require their own evidence.

Supplementary materials

The supplementary material for this article is available at: Supplementary materials.

Declarations

Acknowledgements

The authors would like to thank all colleagues who provided constructive suggestions and technical support during the preparation of this manuscript.

The authors declare that GPT-6 was used solely to polish the language during manuscript preparation. All scholarly content, including the study design, data analysis, interpretations, figures, and tables, is the authors' original work and was developed without the use of AI tools. The authors reviewed, revised, and approved the final manuscript and take full responsibility for its content.

Authors contribution

Wang JN: Conceptualization, writing-original draft.
Lin QY, Fu YY: Investigation, formal analysis.
Guan JX: Visualization.
Li HT, Fan JJ: Formal analysis, writing-review & editing.
Deng XF, Liu SQ, Zhang F, Zhang DX: Conceptualization, supervision, writing-review & editing.

Conflicts of interest

DongXu Zhang is an Editorial Board member of *Computational Biomedicine* . The other authors declare no conflicts of interest.

Ethical approval

Not applicable.

Consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data and materials

All data generated or analysed during this review are included in this article and its supplementary information files. Per-paper source data for Figure 3 (reference number, publication year, and primary theme) and the inclusion ledger for all 105 references (source type, primary

theme, bibliographic publication year, figure inclusion, and inclusion or exclusion reason) are provided in Data S1.

Funding

This work was supported by the National Natural Science Foundation of China (NSFC) under Grant No. 62472365, for the project "Intelligent Immunogen Design of Broad-Spectrum Vaccines Against Highly Variable Pathogens"; the Xiamen Natural Science Foundation under Grant No. 3502Z202473003, for the project "Artificial Intelligence-Based Immunogen Design of Viral Vectors for Broad-Spectrum Influenza Vaccines"; the Central Government-Guided Local Science and Technology Development Fund of Xinjiang Uygur Autonomous Region under Grant No. ZYYD2026QY25, for the project "Development and Clinical Application of Intelligent Digital Pathology Diagnostic Instruments for Breast Surgical Diseases"; and the Xinjiang Uygur Autonomous Region Science and Technology Innovation Leading Talent Program under Grant No. 2023TSYCLJ0039.

Copyright

© The Author(s) 2026.

References

1. World Health Organization. WHO guideline for screening and treatment of cervical pre-cancer lesions for cervical cancer prevention. 2nd ed. Geneva: World Health Organization; 2021.
2. Nayar R, Wilbur DC. The Bethesda system for reporting cervical cytology: Definitions, criteria, and explanatory notes. Cham: Springer; 2015. [DOI]
3. Zhao FH, Hu SY, Bian JJ, Liu B, Peck RB, Bao YP, et al. Comparison of ThinPrep and SurePath liquid-based cytology and subsequent human papillomavirus DNA testing in China. Cancer Cytopathol. 2011;119(6):387-394. [DOI][PubMed]
4. Itonaga M, Ashida R, Kitano M. The usefulness of liquid-based cytology for endoscopic ultrasound-guided tissue acquisition of solid pancreatic masses. Front Med. 2022;9:943792. [DOI][PubMed][PMC]
5. Pantanowitz L, Sinard JH, Henricks WH, Fatheree LA, Carter AB, Contis L, et al. Validating whole slide imaging for diagnostic purposes in pathology: Guideline from the College of American Pathologists Pathology and Laboratory Quality Center. Arch Pathol Lab Med. 2013;137(12):1710-1722. [DOI][PubMed][PMC]
6. Jiang P, Li X, Shen H, Chen Y, Wang L, Chen H, et al. A systematic review of deep learning-based cervical cytology screening: From cell identification to whole slide image analysis. Artif Intell Rev. 2023;56(2):2687-2758. [DOI]
7. McGenity C, Clarke EL, Jennings C, Matthews G, Cartlidge C, Freduah-Agyemang H, et al. Artificial intelligence in digital pathology: A systematic review and meta-analysis of diagnostic test accuracy. NPJ Digit Med. 2024;7(1):114. [DOI][PubMed][PMC]
8. Matthews GA, McGenity C, Bansal D, Treanor D. Public evidence on AI products for digital pathology. NPJ Digit Med. 2024;7(1):300. [DOI][PubMed][PMC]
9. Xie X, Fu CC, Lv L, Ye Q, Yu Y, Fang Q, et al. Deep convolutional neural network-based classification of cancer cells on cytological pleural effusion images. Mod Pathol. 2022;35(5):609-614. [DOI][PubMed][PMC]
10. Park HS, Chong Y, Lee Y, Yim K, Seo KJ, Hwang G, et al. Deep learning-based computational cytopathologic diagnosis of metastatic breast carcinoma in pleural fluid. Cells. 2023;12(14):1847. [DOI][PubMed][PMC]
11. Mavropoulos A, Johnson C, Lu V, Nieto J, Schneider EC, Saini K, et al. Artificial intelligence-driven morphology-based enrichment of malignant cells from body fluid. Mod Pathol. 2023;36(8):100195. [DOI][PubMed]
12. Ito K, Kimura R, Konishi H, Ozawa N, Yaegashi N, Ohashi Y, et al. A comparison of liquid-based and conventional cytology using data for cervical cancer screening from the Japan Cancer Society. Jpn J Clin Oncol. 2020;50(2):138-144. [DOI]
13. Roy-Chowdhuri S. Immunocytochemistry of cytology specimens for predictive biomarkers in lung cancer. Transl Lung Cancer Res. 2020;9(3):898-905. [DOI][PubMed][PMC]
14. Wentzensen N, Lahrmann B, Clarke MA, Kinney W, Tokugawa D, Poitras N, et al. Accuracy and efficiency of deep-learning-based automation of dual stain cytology in cervical cancer screening. J Natl Cancer Inst. 2021;113(1):72-79. [DOI][PubMed][PMC]
15. 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][PubMed]
16. del Mar Rivera Rolon M, Gustafson E, Cole R, Matos J, Hicken K, Hicks J, et al. Implementing 100% quality control in a cervical cytology workflow using whole slide images and artificial intelligence provided by the Techcyte SureView™ System. Cancer Cytopathol. 2025;133(6):e70019. [DOI]
17. Nitta N, Sugiyama Y, Sugimura T, Ito T, Ikebata K, Abe H, et al. Clinical-grade autonomous cytopathology through whole-slide edge tomography. Nature. 2026;651(8105):472-481. [DOI]
18. Hang JF, Ou YC, Yang WL, Tsao TY, Yeh CH, Li CB, et al. Comparative evaluation of slide scanners, scan settings, and cytopreparations for digital urine cytology. J Pathol Inform. 2024;15:100346. [DOI]
19. Kim K, Chaware A, Cook CB, Xu S, Abdelmalak M, Cooke C, et al. Rapid 3D imaging at cellular resolution for digital cytopathology with a multi-camera array scanner (MCAS). npj Imaging. 2024;2(1):39. [DOI][PubMed][PMC]

20. Evans AJ, Brown RW, Bui MM, Chlipala EA, Lacchetti C, Milner DA, et al. Validating whole slide imaging systems for diagnostic purposes in pathology. Arch Pathol Lab Med. 2022;146(4):440-450. [DOI][PubMed]
21. Antonini P, Santonicco N, Pantanowitz L, Girolami I, Rizzo PC, Brunelli M, et al. Relevance of the College of American Pathologists guideline for validating whole slide imaging for diagnostic purposes to cytopathology. Cytopathology. 2023;34(1):5-14. [DOI][PubMed][PMC]
22. Food and Drug Administration (US). Technical performance assessment of digital pathology whole slide imaging devices: guidance for industry and Food and Drug Administration staff [Internet]. Silver Spring (MD): Food and Drug Administration (US); 2016 [cited 2026 Sep 28]. Available from: https://www.fda.gov/regulatory-information/search-fda-guidance-documents/technical-performance-assessment-digital-pathology-whole-slide-imaging-devices
23. Chong Y, Hong SA, Oh HK, Jung SJ, Kim BS, Jeong JY, et al. Diagnostic proficiency test using digital cytopathology and comparative assessment of whole slide images of cytologic samples for quality assurance program in Korea. J Pathol Transl Med. 2023;57(5):251-264. [DOI][PubMed][PMC]
24. Hosseini MS, Brawley-Hayes JAZ, Yueyang Z, Chan L, Plataniotis K, Damaskinos S. Focus quality assessment of high-throughput whole slide imaging in digital pathology. IEEE Trans Med Imaging. 2020;39(1):62-74. [DOI]
25. Kohlberger T, Liu Y, Moran M, Chen PC, Brown T, Hipp JD, et al. Whole-slide image focus quality: Automatic assessment and impact on AI cancer detection. J Pathol Inform. 2019;10:39. [DOI][PubMed][PMC]
26. Kang L, Wang J, Qin J, Liang Y, He Y. Automated quality evaluation of Cervical Cytopathology Whole Slide Images based on content analysis. Biomed Signal Process Control. 2026;115:109419. [DOI]
27. Harinath L, Harrington S, Matsko J, Colaizzi A, Elishaev E, Khader S, et al. Technical considerations during validation of the genius® digital diagnostic system. J Pathol Inform. 2026;20:100532. [DOI][PubMed][PMC]
28. Sakabe N, Yoshizaki Y, Fukuda K, Sato S, Ikeda K. Relationship between artificial intelligence-based cell detection and cytomorphological variations induced by cell processing solutions: Usefulness of data augmentation in artificial intelligence cytology. Acta Cytol. 2025;69(6):603-612. [DOI][PubMed]
29. Wang R, Li Q, Shi G, Li Q, Zhong D. A deep learning framework for predicting endometrial cancer from cytopathologic images with different staining styles. PLoS One. 2024;19(7):e0306549. [DOI][PubMed]
30. World Health Organization. WHO guideline for screening and treatment of cervical pre-cancer lesions for cervical cancer prevention: use of dual-stain cytology to triage women after a positive test for human papillomavirus (HPV) [Internet]. Geneva: World Health Organization; 2024. Available from: https://www.who.int/publications/i/item/9789240091658
31. Clarke MA, Wentzensen N, Perkins RB, Garcia F, Arrindell D, Chelmow D, et al. Recommendations for use of p16/Ki67 dual stain for management of individuals testing positive for human papillomavirus. J Low Genit Tract Dis. 2024;28(2):124-130. [DOI][PubMed][PMC]
32. Perkins RB, Guido RS, Castle PE, Chelmow D, Einstein MH, Garcia F, et al. 2019 ASCCP risk-based management consensus guidelines for abnormal cervical cancer screening tests and cancer precursors. J Low Genit Tract Dis. 2020;24(2):102-131. [DOI]
33. Perkins RB, Guido RS, Castle PE, Chelmow D, Einstein MH, Garcia F, et al. 2019 ASCCP risk-based management consensus guidelines: Updates through 2023. J Low Genit Tract Dis. 2024;28(1):3-6. [DOI][PubMed][PMC]
34. Ikenberg H, Bergeron C, Schmidt D, Griesser H, Alameda F, Angeloni C, et al. Screening for cervical cancer precursors with p16/Ki-67 dual-stained cytology: Results of the PALMS study. J Natl Cancer Inst. 2013;105(20):1550-1557. [DOI][PubMed][PMC]
35. Wright TC Jr, Stoler MH, Ranger-Moore J, Fang Q, Volkir P, Safaeian M, et al. Clinical validation of p16/Ki-67 dual-stained cytology triage of HPV-positive women: Results from the IMPACT trial. Int J Cancer. 2022;150(3):461-471. [DOI][PubMed][PMC]
36. Clarke MA, Cheung LC, Castle PE, Schiffman M, Tokugawa D, Poitras N, et al. Five-year risk of cervical precancer following p16/Ki-67 dual-stain triage of HPV-positive women. JAMA Oncol. 2019;5(2):181-186. [DOI][PubMed][PMC]
37. Lee Y, Alam MR, Park H, Yim K, Seo KJ, Hwang G, et al. Improved diagnostic accuracy of thyroid fine-needle aspiration cytology with artificial intelligence technology. Thyroid. 2024;34(6):723-734. [DOI][PubMed]
38. National Electrical Manufacturers Association. DICOM whole slide imaging (WSI) [Internet]. Arlington (VA): National Electrical Manufacturers Association. Available from: https://dicom.nema.org/dicom/dicomwsi/
39. Cheng S, Liu S, Yu J, Rao G, Xiao Y, Han W, et al. Robust whole slide image analysis for cervical cancer screening using deep learning. Nat Commun. 2021;12(1):5639. [DOI][PubMed][PMC]
40. Wang J, Yu Y, Tan Y, Wan H, Zheng N, He Z, et al. Artificial intelligence enables precision diagnosis of cervical cytology grades and cervical cancer. Nat Commun. 2024;15(1):4369. [DOI]
41. Kurita Y, Meguro S, Kosugi I, Enomoto Y, Kawasaki H, Kano T, et al. Enhancing cervical cancer cytology screening via artificial intelligence innovation. Sci Rep. 2024;14(1):19535. [DOI][PubMed][PMC]
42. Xue P, Dang L, Kong LH, Tang HP, Xu HM, Weng HY, et al. Deep learning enabled liquid-based cytology model for cervical precancer and cancer detection. Nat Commun. 2025;16(1):3506. [DOI][PubMed][PMC]
43. Song Y, Zhang A, Zhou J, Luo Y, Lin Z, Zhou T. Overlapping cytoplasms segmentation via constrained multi-shape evolution for cervical cancer screening. Artif Intell Med. 2024;148:102756. [DOI][PubMed]
44. He Z, Jia D, Zhang C, Li Z, Wu N. A two-stage approach solo_GAN for overlapping cervical cell segmentation based on single-cell identification and boundary generation. Appl Intell. 2024;54(6):4621-4645. [DOI]
45. Cai D, Chen J, Zhao J, Xue Y, Yang S, Yuan W, et al. HiCervix: An extensive hierarchical dataset and benchmark for cervical cytology classification. IEEE Trans Med Imaging. 2024;43(12):4344-4355. [DOI]

46. Zhang X, Ji J, Zhang Q, Zheng X, Ge K, Hua M, et al. A large annotated cervical cytology images dataset for AI models to aid cervical cancer screening. Sci Data. 2025;12(1):23. [DOI]
47. Zhang Z, Yao P, Chen M, Zeng L, Shao P, Shen S, et al. SCAC: A semi-supervised learning approach for cervical abnormal cell detection. IEEE J Biomed Health Inform. 2024;28(6):3501-3512. [DOI]
48. Ilse M, Tomczak JM, Welling M. Attention-based deep multiple instance learning. In: Dy J, Krause A, editors. Proceedings of the 35th International Conference on Machine Learning; 2018 Jul 10-15; Stockholm, Sweden. Cambridge: PMLR; 2018. p. 2127-2136. Available from: https://proceedings.mlr.press/v80/ilse18a.html
49. Lu MY, Williamson DFK, Chen TY, Chen RJ, Barbieri M, Mahmood F. Data-efficient and weakly supervised computational pathology on whole-slide images. Nat Biomed Eng. 2021;5(6):555-570. [DOI][PubMed][PMC]
50. Kanavati F, Hirose N, Ishii T, Fukuda A, Ichihara S, Tsuneki M. A deep learning model for cervical cancer screening on liquid-based cytology specimens in whole slide images. Cancers. 2022;14(5):1159. [DOI][PubMed][PMC]
51. Jiang H, Cai J, Shen Z, Xu M, Fei M, Huang H, et al. UniCAS: A foundation model for cervical cytology screening. Cell Rep Med. 2026;7(1):102570. [DOI]
52. Zheng X, Zheng K, Wang J, Zhang X, Li K, Zhao Z, et al. A universal visual foundation model for computational cytopathology. Nat Cancer. 2026;7(10):1643-1658. [DOI][PubMed]
53. Stegmüller T, Abbet C, Bozorgtabar B, Clarke H, Petignat P, Vassilakos P, et al. Self-supervised learning-based cervical cytology for the triage of HPV-positive women in resource-limited settings and low-data regime. Comput Biol Med. 2024;169:107809. [DOI][PubMed]
54. Chun J, Yu A, Ko S, Chong G, Park J, Han H, et al. Automated screening of precancerous cervical cells through contrastive self-supervised learning. Life. 2024;14(12):1565. [DOI][PubMed][PMC]
55. Ivezić V, Radhachandran A, Redekop E, Athreya S, Lee D, Sant V, et al. CytoFM: The first cytology foundation model. In: 2025 IEEE/CVF conference on computer vision and pattern recognition workshops (CVPRW). IEEE; 2025. p. 4749-4757. [DOI]
56. Abe M, Suzuki A, Kamada K, Kanno T, Shimasaki M, Mori T, et al. Automated acquisition of thyroid cytology image features using self-supervised learning and application to image classification. Commun Med. 2026. [DOI]
57. Song AH, Jaume G, Williamson DFK, Lu MY, Vaidya A, Miller TR, et al. Artificial intelligence for digital and computational pathology. Nat Rev Bioeng. 2023;1(12):930-949. [DOI]
58. Chen RJ, Ding T, Lu MY, Williamson DFK, Jaume G, Song AH, et al. Towards a general-purpose foundation model for computational pathology. Nat Med. 2024;30(3):850-862. [DOI][PubMed][PMC]
59. Lu MY, Chen B, Williamson DFK, Chen RJ, Liang I, Ding T, et al. A visual-language foundation model for computational pathology. Nat Med. 2024;30(3):863-874. [DOI][PubMed][PMC]
60. Vorontsov E, Bozkurt A, Casson A, Shaikovski G, Zelechowski M, Severson K, et al. A foundation model for clinical-grade computational pathology and rare cancers detection. Nat Med. 2024;30(10):2924-2935. [DOI][PubMed][PMC]
61. Xu H, Usuyama N, Bagga J, Zhang S, Rao R, Naumann T, et al. A whole-slide foundation model for digital pathology from real-world data. Nature. 2024;630(8015):181-188. [DOI][PubMed][PMC]
62. Yan R, Ren F, Li J, Rao X, Lv Z, Zheng C, et al. Nuclei-guided network for breast cancer grading in HE-stained pathological images. Sensors. 2022;22(11):4061. [DOI][PubMed][PMC]
63. Yan R, Yang Z, Li J, Zheng C, Zhang F. Divide-and-attention network for HE-stained pathological image classification. Biology. 2022;11(7):982. [DOI][PubMed][PMC]
64. Qiao Y, Zhao L, Luo C, Luo Y, Wu Y, Li S, et al. Multi-modality artificial intelligence in digital pathology. Brief Bioinform. 2022;23(6):bbac367. [DOI][PubMed][PMC]
65. Yang D, Li T, Li L, Chen S, Li X. Multi-modal convolutional neural network-based thyroid cytology classification and diagnosis. Hum Pathol. 2025;161:105868. [DOI]
66. Shen R, Jiang F, Huang X, Hong G, Luo Y, Wan H, et al. Artificial intelligence-assisted urine cytology for noninvasive detection of muscle-invasive urothelial carcinoma: A multi-center diagnostic study with prospective validation. Adv Sci. 2025;12(39):e08977. [DOI][PubMed][PMC]
67. Lou Y, Su Y, Lu H, Li W, Yin W, Li S, et al. Cytological classification diagnosis for thyroid nodules via multimodal model deep learning. Adv Sci. 2025;12(48):e11369. [DOI][PubMed][PMC]
68. Fei M, Shen Z, Liu M, Song Z, Sun Y, Han X, et al. Refining cervical cell classification with Cytological knowledge and Optimal attribute descriptor matching. In: Gee JC, Alexander DC, Hong J, Iglesias JE, Sudre CH, Venkataraman A, Golland P, Kim JH, Park J, editors. Medical Image Computing and Computer Assisted Intervention – MICCAI 2025; 2025 Sep 23-27; Daejeon, South Korea. Cham: Springer; 2026. p. 519-529. [DOI]
69. Wang J, Zheng N, Wan H, Yao Q, Jia S, Zhang X, et al. Deep learning models for thyroid nodules diagnosis of fine-needle aspiration biopsy: A retrospective, prospective, multicentre study in China. Lancet Digit Health. 2024;6(7):e458-e469. [DOI]
70. Satturwar S, Li Z, Yang CS, Lin YJ, Yang WL, Lin MY, et al. Artificial intelligence-assisted digital thyroid FNA cytology: Improved agreement and sensitivity for higher-risk Bethesda categories with enhanced screening efficiency. Cancer Cytopathol. 2026;134(10):e70149. [DOI][PubMed][PMC]
71. Dey P, Muralidaran C. Comparative evaluation of transfer learning models for detecting malignant cells in urinary cytology. Cytopathology. 2026;37(3):232-239. [DOI]
72. Chen L, Cao X, Lu Z, Liu L, Sun M, Wu Y, et al. Artificial intelligence–assisted urine cytology based on the Paris System for Reporting

Urothelial Carcinoma. Cancer Cytopathol. 2026;134(6):e70120. [DOI]
73. Fang D, Huang Y, Li S, Shi C, Bao J, Du D, et al. A semi-supervised convolutional neural network for diagnosis of pancreatic ductal adenocarcinoma based on EUS-FNA cytological images. BMC Cancer. 2025;25(1):495. [DOI]
74. Tian Y, Zhang J, Rao X, Zhou B, Zhao X, He Y, et al. AI-assisted rapid on-site evaluation ROSE of EUS-FNA cytopathology for pancreatic solid lesions: A two-stage deep learning approach. Endosc Ultrasound. 2025;14(6):314-327. [DOI][PubMed][PMC]
75. Liu TJ, Yang WC, Huang SM, Yang WL, Wu HJ, Ho HW, et al. Evaluating artificial intelligence–enhanced digital urine cytology for bladder cancer diagnosis. Cancer Cytopathol. 2024;132(11):686-695. [DOI]
76. Chang KY, Yang CS, Lai JY, Lin SJ, Li JR, Liu TJ, et al. An artificial intelligence-assisted diagnostic system improves upper urine tract cytology diagnosis. . 2024;38(6):3016-3021. [DOI][PubMed][PMC]
77. Tanaka R, Tsuboshita Y, Okodo M, Settsu R, Hashimoto K, Tachibana K, et al. Artificial intelligence recognition model using liquid-based cytology images to discriminate malignancy and histological types of non-small-cell lung cancer. Pathobiology. 2025;92(1):52-62. [DOI] [PubMed]
78. Tian F, Liu D, Wei N, Fu Q, Sun L, Liu W, et al. Prediction of tumor origin in cancers of unknown primary origin with cytology-based deep learning. Nat Med. 2024;30(5):1309-1319. [DOI][PubMed][PMC]
79. Kim D, Lee J, Jung M, Yim K, Hwang G, Yoon H, et al. Whole slide image-level classification of malignant effusion cytology using clustering-constrained attention multiple instance learning. Lung Cancer. 2025;204:108552. [DOI]
80. Cantley RL, Jing X, Smola B, Hao W, Harrington S, Pantanowitz L. Validation of AI-assisted ThinPrep® Pap test screening using the GeniusTM Digital Diagnostics System. J Pathol Inform. 2024;15:100391. [DOI]
81. Bai X, Wei J, Starr D, Zhang X, Wu X, Guo Y, et al. Assessment of efficacy and accuracy of cervical cytology screening with artificial intelligence assistive system. Mod Pathol. 2024;37(6):100486. [DOI]
82. Yang W, Jin X, Huang L, Jiang S, Xu J, Fu Y, et al. Clinical evaluation of an artificial intelligence-assisted cytological system among screening strategies for a cervical cancer high-risk population. BMC Cancer. 2024;24(1):776. [DOI][PubMed][PMC]
83. Liu L, Liu J, Su Q, Chu Y, Xia H, Xu R. Performance of artificial intelligence for diagnosing cervical intraepithelial neoplasia and cervical cancer: A systematic review and meta-analysis. EClinicalMedicine. 2025;80:102992. [DOI][PubMed][PMC]
84. Byambadorj T, Alam MR, Chong Y. Artificial intelligence in nongynecologic cytology: A systematic review of current research and commercial tools. Cancer Cytopathol. 2026;134(4):e70092. [DOI][PubMed][PMC]
85. Xue P, Tang H, Weng H, Xu H, Kong L, Pan Q, et al. Deep learning-assisted versus manual reading in routine cervical cytopathology: A multicentre randomised crossover trial. npj Digit Med. 2026;9(1):422. [DOI]
86. Wu S, Shen R, Hong G, Luo Y, Wan H, Feng J, et al. Development and validation of an artificial intelligence-based model for detecting urothelial carcinoma using urine cytology images: A multicentre, diagnostic study with prospective validation. EClinicalMedicine. 2024;71:102566. [DOI]
87. Food and Drug Administration (US). Evaluation of automatic class III designation for Genius™ Digital Diagnostics System with the Genius™ Cervical AI algorithm: decision summary DEN210035 [Internet]. Silver Spring (MD): Food and Drug Administration (US); 2024 [cited 2026 Oct 20]. Available from: HYPERLINK "https://www.accessdata.fda.gov/cdrh_docs/reviews/DEN210035.pdf" https://www.accessdata.fda.gov/cdrh_docs/reviews/DEN210035.pdf
88. Techcyte, Inc. Cytolab moves its cervical cytology screening to Techcyte SureView [Internet]. Orem (UT): Techcyte, Inc.; 2026 [cited 2026 Oct 20]. Available from: https://techcyte.com/cytolab-moves-to-techcyte-sureview/
89. Collins GS, Moons KGM, Dhiman P, Riley RD, Beam AL, Van Calster B, et al. TRIPOD+AI statement: Updated guidance for reporting clinical prediction models that use regression or machine learning methods. BMJ. 2024;385:e078378. [DOI][PubMed][PMC]
90. Sounderajah V, Guni A, Liu X, Collins GS, Karthikesalingam A, Markar SR, et al. The STARD-AI reporting guideline for diagnostic accuracy studies using artificial intelligence. Nat Med. 2025;31(10):3283-3289. [DOI][PubMed]
91. Liu X, Cruz Rivera S, Moher D, Calvert MJ, Denniston AK, SPIRIT-AI and CONSORT-AI Working Group. Reporting guidelines for clinical trial reports for interventions involving artificial intelligence: The CONSORT-AI extension. Nat Med. 2020;26(9):1364-1374. [DOI] [PubMed][PMC]
92. Cruz Rivera S, Liu X, Chan AW, Denniston AK, Calvert MJ, SPIRIT-AI and CONSORT-AI Working Group, et al. Guidelines for clinical trial protocols for interventions involving artificial intelligence: The SPIRIT-AI extension. Nat Med. 2020;26(9):1351-1363. [DOI][PubMed] [PMC]
93. Vasey B, Nagendran M, Campbell B, Clifton DA, Collins GS, Denaxas S, et al. Reporting guideline for the early stage clinical evaluation of decision support systems driven by artificial intelligence: DECIDE-AI. BMJ. 2022;377:e070904. [DOI][PubMed][PMC]
94. Food and Drug Administration (US). Artificial intelligence-enabled device software functions: lifecycle management and marketing submission recommendations: draft guidance for industry and Food and Drug Administration staff [Internet]. Silver Spring (MD): Food and Drug Administration (US); 2025 [cited 2026 Oct 8]. Available from: https://www.fda.gov/regulatory-information/search-fda-guidance-documents/artificial-intelligence-enabled-device-software-functions-lifecycle-management-and-marketing
95. Food and Drug Administration (US). Marketing submission recommendations for a predetermined change control plan for artificial intelligence-enabled device software functions: guidance for industry and Food and Drug Administration staff [Internet]. Silver Spring (MD): Food and Drug Administration (US); 2025 [cited 2026 Oct 8]. Available from: https://www.fda.gov/regulatory-information/search-fda-guidance-documents/marketing-submission-recommendations-predetermined-change-control-plan-artificial-intelligence
96. Food and Drug Administration (US). List of artificial intelligence-enabled medical devices [Internet]. Silver Spring (MD): Food and Drug

Administration (US) [cited 2026 Sep 28]. Available from: https://www.fda.gov/medical-devices/artificial-intelligence-enabled-medical-devices/list-artificial-intelligence-enabled-medical-devices
97. World Health Organization. Ethics and governance of artificial intelligence for health: WHO guidance [Internet]. Geneva: World Health Organization; 2021. Available from: https://iris.who.int/handle/10665/341996
98. Xue P, Xu HM, Tang HP, Weng HY, Wei HM, Wang Z, et al. Improving the accuracy and efficiency of abnormal cervical squamous cell detection with cytologist-in-the-loop artificial intelligence. Mod Pathol. 2023;36(8):100186. [DOI][PubMed]
99. Welch EC, Lu C, Sung CJ, Zhang C, Tripathi A, Ou J. BMT: A cross-validated ThinPrep pap cervical cytology dataset for machine learning model training and validation. Sci Data. 2024;11(1):1444. [DOI]
100. Karasu Benyes Y, Welch EC, Singhal A, Ou J, Tripathi A. A comparative analysis of deep learning models for automated cross-preparation diagnosis of multi-cell liquid pap smear images. Diagnostics. 2022;12(8):1838. [DOI][PubMed][PMC]
101. Fujii Y, Uchida D, Sato R, Obata T, Akihiro M, Miyamoto K, et al. Effectiveness of data-augmentation on deep learning in evaluating rapid on-site cytopathology at endoscopic ultrasound-guided fine needle aspiration. Sci Rep. 2024;14(1):22441. [DOI][PubMed][PMC]
102. Lajara S, Cuda J, Geisler DL, Staniszewski C, Yang WL, Atkison K, et al. Enhancing urothelial carcinoma diagnosis with artificial intelligence-integrated urine cytology: Biopsy-validated accuracy and efficiency gain. Cancer Cytopathol. 2026;134(6):e70110. [DOI][PubMed][PMC]
103. Murphy KM, Weatherhead K, Chenault C, Nguyen C, Sefcik K, Harrington S, et al. Impact of the Genius Digital Diagnostics System on workflow and accuracy compared with the ThinPrep Imaging System for review of ThinPrep Papanicolaou tests. Am J Clin Pathol. 2025;164(5):746-751. [DOI][PubMed][PMC]
104. Wilson A, Cropper A, Ma Y, Adams EJ, Mehew S, Morgan M, et al. Improving laboratory workforce efficiency using AI-assisted digital cytology within an HPV-based cervical screening programme: A model-based evaluation for the NHS Cervical Screening Programmes. BMJ Open. 2026;16(7):e113298. [DOI][PubMed][PMC]
105. Mach AJ, Adeyiga OB, di Carlo D. Microfluidic sample preparation for diagnostic cytopathology. Lab Chip. 2013;13(6):1011-1026. [DOI][PubMed][PMC]

Publisher's Note

Science Exploration remains a neutral stance on jurisdictional claims in published maps and institutional affiliations. The views expressed in this article are solely those of the author(s) and do not reflect the opinions of the Editors or the publisher.