Acta Cytologica

Gynecologic Cytopathology

Acta Cytologica 2026;70:52–71
DOI: 10.1159/000546182

Received: February 21, 2025
Accepted: March 26, 2025
Published online: July 8, 2025

Artificial Intelligence in Gynecologic Cytology

Lakshmi Harinath[a] Xinru Bai[b] Jeremy Minkowitz[a] Xianxu Zeng[b]
Sarah Harrington[c] Chengquan Zhao[a] Liron Pantanowitz[a, d]

[a]Department of Pathology, University of Pittsburgh, UPMC Magee-Womens Hospital, Pittsburgh, PA, USA; [b]Department of Pathology, The Third Affiliated Hospital of Zhengzhou University, Zhengzhou, China; [c]Scientific Affairs, Hologic, Inc., Marlborough, MA, USA; [d]Department of Pathology, Computational Pathology and AI Center of Excellence (CPACE), University of Pittsburgh, Pittsburgh, PA, USA

Keywords
Artificial intelligence · Cervical cancer · Cytology · Digital imaging · Pap test · Screening

Abstract
Background: Cervical cancer is the fourth most common cancer in women globally with highest incidence and mortality identified in less developed and medically underserved areas in the world. The diminishing cytology workforce, unavailability of expert consultation, and the high volume of Pap tests needing manual screening are the main reasons for exploring innovative solutions to help mitigate the negative effects resulting from the dearth of timely cervical cancer screening in certain population groups. ***Summary:*** Developments in whole slide imaging and artificial intelligence (AI) have enabled the emergence of new computer-assisted systems that have the potential for transforming traditional cytopathology practice. However, AI-based systems are relatively new with limited published data on their validation and clinical utility in clinical practice. Our article aims to increase awareness of the availability of such systems, explores the history and development of AI-assisted screening platforms for Pap tests, compares the performance characteristics of various systems, elaborates on technical challenges associated with conducting clinical trials employing this technology, and discusses considerations around deploying such systems in routine cytopathology practice. ***Key Message:*** Revolutionary AI-based systems are being developed and utilized in cytopathology practice to screen Pap tests. Some of these systems have good performance characteristics and provide opportunities to combat various issues such as workload and standardization faced by cytology laboratories globally. However, judicious review of these systems using evidence-based studies is imperative to promote widespread adoption and maintain high-quality standards for patient safety.

Introduction

Cervical cancer is the fourth most common cancer among women with an estimated 660,000 women diagnosed with cervical cancer and about 350,000 women succumbing to this disease worldwide in 2022 [1]. Almost all cases of cervical cancer are linked to the sexually transmissible high-risk human papilloma virus (HPV)

Lakshmi Harinath and Xinru Bai are co-first authors and contributed equally to this work.

karger@karger.com
www.karger.com/acy

Correspondence to:
Lakshmi Harinath, harinathl@upmc.edu
Liron Pantanowitz, pantanowitzl2@upmc.edu

Karger

with persistent infection, leading to the development of cervical cancer in affected populations. As early as 1928, Dr. George Papanicolaou proposed using the Pap smear cytology screening test for the detection of cervical cancer. The efficacy of this Pap test was proven around the 1940s with routine annual screening drastically reducing the incidence and prevalence of cervical cancer in the USA, earmarking this test as one of the most significant clinical tools in the control of cancer in the 20th century [2, 3].

Although primary HPV vaccination, early detection with cervical cytology screening, and appropriate management can prevent and help treat this disease, factors such as lack of routine care in underdeveloped regions, ineffective vaccination programs, and the scarcity of skilled cytology personnel in both developed and underdeveloped regions contribute to the high incidence and prevalence of this disease [1, 4–6]. Currently, the dire cytologist staffing shortages experienced by multiple large laboratories are being compounded by the decreasing enrollment reported by cytotechnology schools [7, 8]. Moreover, conventional cervical cytology screening relies on cytologists manually reviewing slides with a traditional microscope that can be a very tedious, labor-intensive, and error-prone task requiring significant time and effort by screeners. The arrival of liquid-based cytology (LBC) solutions in the 1990s with the discovery of automated or computer-assisted LBC screening systems like the ThinPrep® Imaging System (TIS) in the early 2000s was a technological breakthrough that provided some reprieve from these difficulties [9]. However, despite these advances, cytology laboratories continue to face a myriad of additional challenges such as inaccurate results, significant interobserver variability with manual screening, increased turnaround times, medicolegal issues, fatigue, and other ergonomic health conditions faced by personnel due to prolonged manual microscope use [9–13]. These challenges have prompted relevant stakeholders in the cytology community to discover newer innovative means to redress these difficulties [11].

The field of artificial intelligence (AI) has rapidly expanded due to high performance computing, deep learning, and access to large digital datasets for training algorithms. Coupled with the transformation occurring in the field of digital pathology (DP), many novel AI-assisted screening platforms are being designed and tested for both gynecologic and non-gynecologic cytology practice. The AI-assisted platforms for non-gynecologic cytology are still in the research phase with no commercial product currently available in the market for routine clinical use. However, there are several AI-based screening systems focused on cervical cytology with some of them currently available for commercial use to help screen Pap tests in clinical practice.

It is important to note that the concept of utilizing AI systems coupled with whole slide imaging (WSI) for cervical cancer detection is relatively recent. Hence, there are limited standardized guidelines available for their deployment with minimal published data focusing on their validation and clinical utility in routine practice. This article aims to increase awareness of the availability of such AI-based systems, explores the history and development of AI screening platforms utilized in cervical cytology, analyzes and compares the performance characteristics of some of these systems, discusses the potential for improvement in accuracy, efficiency, and accessibility seen with such systems, and considers potential barriers and drivers that would impact the successful integration of such systems in routine cytopathology practice.

Digital Cytopathology and AI: Past and Present

The Digital Pathology Association defines digital pathology (DP) as a "dynamic, image-based environment that enables the acquisition, management, and interpretation of pathology information generated from a digitized glass slide" [14]. In this context, a Digital Pathology System includes a whole slide scanner, image acquisition software, image viewing and database software (or image management system), possible image analysis software, and the necessary IT infrastructure to support the Digital Pathology System [14].

AI, as a term, encompasses machine learning and deep learning algorithms that mimic human intelligence and decision-making. Machine learning has the capacity for computer-based algorithms to learn from and make decisions based on data. On the other hand, deep learning is a subset of machine learning that is based on artificial neural networks of many layers designed to mimic the way a human brain learns from large datasets, processes information, and makes complex decisions [15]. In 2023, Muhtasim et al. [16] conducted a critical analysis of machine learning and deep learning methods for cervical cancer screening wherein they looked at a total of 50 studies, 25 of which used machine learning and 22 studies that utilized deep learning. They found that deep learning techniques, including convolutional neural networks, have the potential to increase classification accuracy and decrease false-positive rates in cervical cancer screening (sensitivity – 0.90, specificity – 0.95), and that these AI

systems displayed high sensitivity and specificity rates (sensitivity – 0.86, specificity – 0.92). However, the caliber of the input data and settings applied, as well as the availability of sizable datasets of cervical image annotations for algorithm training, impacted the results from these studies [17]. AI algorithms for cervical cytology screening based on WSI and LBC have been studied widely in the past decade. Knowledge and understanding of deep learning algorithms for automatic processing, recognition, feature extraction, and classification of images have increased by leaps and bounds and enabled AI to analyze images, identify patterns, and interpret malignant characteristics especially in the field of cervical cytology screening [18]. Thus, AI-assisted screening applied to digitized Pap test slides has the potential to mitigate the shortage of skilled cytology personnel, standardize interpretations, and enhance access to expert consultations with the help of remote digital education platforms [18].

Machine learning has been integrated into cervical cytology screening since it was developed in the late 1990s. The US Food and Drug Administration (FDA) has approved at least three DP systems (PAPNET, FocalPoint GS, and Hologic ThinPrep Imaging System) for use in clinical practice in the past [19, 20]. In 1995, the US FDA approved PAPNET (Neuromedical Systems, Inc., Suffern, NY, USA), the first commercially available computer-assisted screening system for cervical cytology [21]. This system used neural networking to differentiate abnormal epithelial cells from normal cells in Pap tests and was designed to reduce the rate of false negatives, and it was used mainly for performing quality control [9, 14, 22]. Later on, in the early 2000s, the TIS (Hologic, Bedford, MD, USA) and the FocalPoint GS Imaging System (Becton Dickinson, Burlington, NC, USA) were introduced for primary screening to be used to interpret ThinPrep and SurePath Pap test preparations for final reporting according to the Bethesda system [9, 14]. The predecessor of the FocalPoint GS, AutoPap 300 QC System (Tripath Imaging, Inc., Burlington, NC, USA), was approved by the US FDA in the mid-1990s [23]. FocalPoint GS used a robotic microscope and screened and marked slides, highlighting the most concerning cells enabling cytologists to identify abnormal cells quickly and reliably [24–26]. FocalPoint GS was acquired and renamed by Becton, Dickinson (BD) Diagnostics with the US FDA approving their system in 2010. Compared with manual screening, BD FocalPoint GS has been shown to have the same or even higher sensitivity when diagnosing cases of low-grade squamous intraepithelial lesion and above (LSIL+) and high-grade squamous intraepithelial lesion [27, 28] (Fig. 1).

The TIS was developed by Cytyc and FDA approved in 2003. In 2007, Hologic acquired Cytyc and the ThinPrep product line [29, 30]. The TIS used an algorithm to detect areas of interest on unscreened Pap liquid-based slides for further examination by cytologists. The selection of the field of view and positioning guidance are provided by a combination of a machine learning algorithm and robotic controller, yet these computer-aided Pap test screening systems still require cytologists to manually use glass slides and optical microscopes. According to some studies, TIS-assisted screening approximately doubled slide reading efficiency, while keeping sensitivity equal or exceeding that of manual primary screening, without adversely affecting specificity [31, 32].

AI-Aided Diagnostic Systems

The development of a WSI-based AI system for cervical cytology screening ushered in a new era in cervical cancer screening. Since the COVID-19 pandemic in 2020, WSI has been recognized and utilized remotely by many more pathologists. The resultant broader adoption of WSI for diagnostic use provided an impetus for the development of WSI-based AI, including Pap cytology testing [20].

Currently, multiple WSI-based AI platforms are commercially available for screening Pap tests. There is increasing interest in these AI-powered systems, specifically for metrics of their performance such as specificity, sensitivity, accuracy, negative predictive value, and positive predictive value when interpreting Pap tests. Further studies are needed on whether these AI platforms can actually reduce workload by acting as an initial screener, provide interpretive results, and possibly even assign cases a category of "no further review" (NFR). Essentially, cases triaged as NFR would not need any further diagnostic assessment by a human. Many of these pioneering studies have been conducted in Europe and Asia [33].

Nearly all modern cervical cytology screening AI platforms allow for whole slide image or cell spot viewing in addition to displaying objects of interest (OOI) detected in the form of a gallery of static images or folder. The OOI gallery shows abnormal cervical cells in an organized (e.g., rows) and labeled format. Platforms such as CytoProcessor™ sort and screen abnormal cells enable the reviewer to make a diagnosis simply by examining the first 100 cells picked by the system [34, 35]. The BestCyte® cell sorter imaging system streamlines hybrid digital screening by efficiently categorizing and sorting

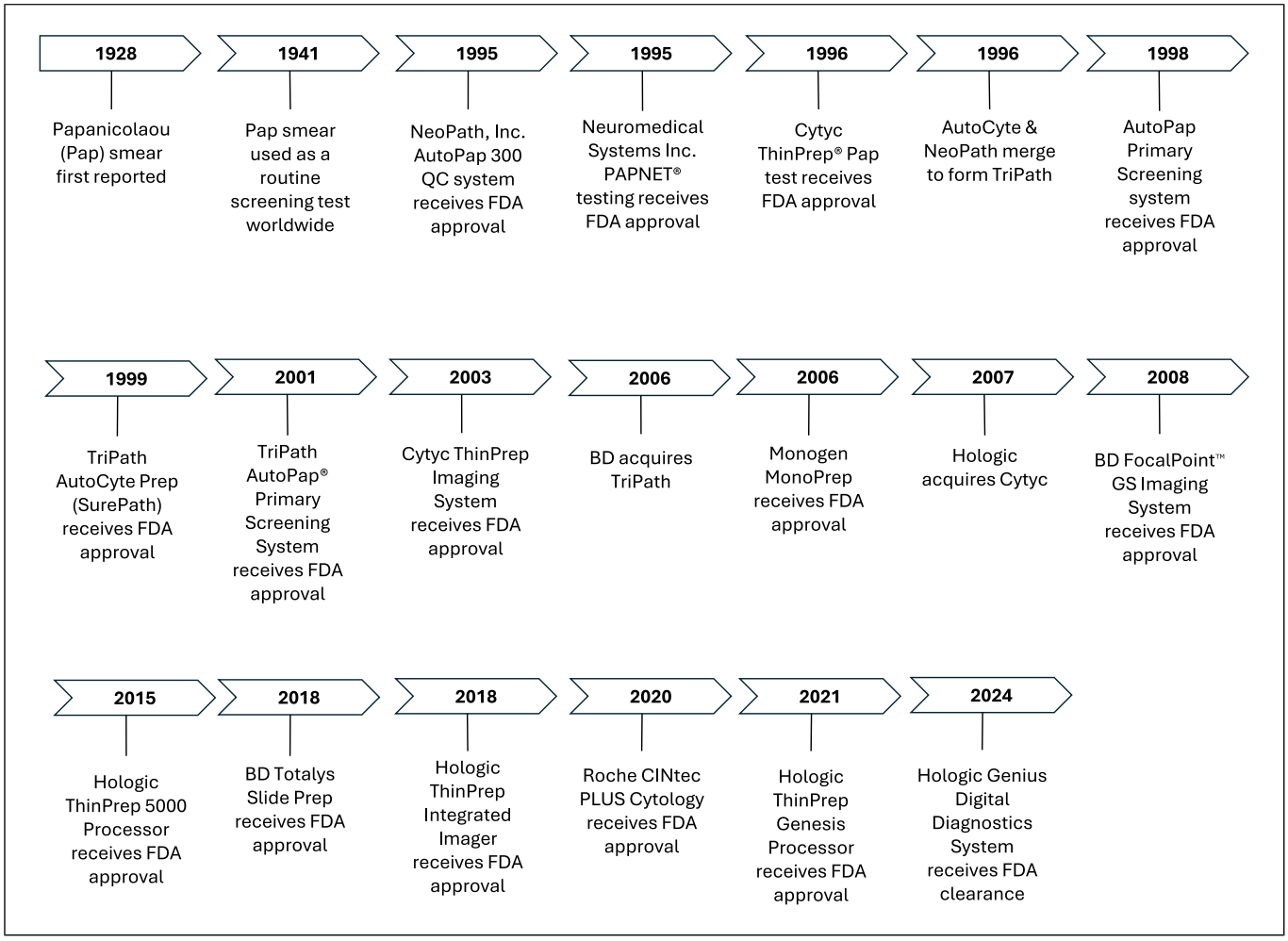

Fig. 1. PapTest: historic timeline.

cell tiles within 8 galleries [19, 36, 37]. The Hologic Genius Diagnostic System presents 30 OOIs (arranged in 5 rows with 6 OOIs per row), and an additional 30 OOIs if needed, for diagnosis [38]. The Landing CytoScanner system ranks all image patches based on assigned scores and highlights the 20 most significant OOIs for easy viewing [39, 40]. AICyte scans Pap test slides and recommends only 20 OOIs to cytologists, thereby assisting in the diagnostic process [41]. Finally, Cytologist-In-The-Loop Artificial Intelligence (CITL-AI), Artificial Intelligence Cervical Cancer Screening (AICCS), and Cell Image Analysis System (CIAS) platforms all provide cytologists with the location of suspicious OOIs for a detailed review [42–44].

There is no standard benchmark for monitoring speed when utilizing such platforms, making it a variable that is challenging to compare. Reading times for each slide by personnel are different due to the types of cases selected and different validation study designs used. Reported speeds for CytoProcessor's AI system surpass traditional microscopy by 1.6 times, while the Hologic Genius Digital Diagnostic System (HGDDS) boasts a notably faster screening time than the TIS (44.8 s versus 89.9 s) [35, 38]. The AI-assistive TBS (AIATBS) system processes slides in under 180 s, while the CIAS processes slides in 80 s [44, 45]. AICyte has the fastest review speeds, slashing the average slide review time down to just 22.23 s [41].

AI Platforms for Cytology Cervical Pap Tests

After analyzing stained Pap test slides, many platforms provide AI-assisted interpretive results. These diagnoses are either binary or picked from a variety of choices. For

Table 1. Overview of study characteristics of AI-related studies

Lead author	Institution (year)	AI system tested	Patient population	Reviewers	Type of study/aim	Type of cytology slide
Elishaev (manuscript forthcoming)	University of Pittsburgh	Genius Digital Diagnostic System (HGDDS)	USA	3 CT 3 CP	Validation	TP
Cantley et al. [11]	University of Michigan (2024)	Genius Digital Diagnostic System (HGDDS)	USA	6 CT 3 CP	Validation	TP
Ikenberg et al. [38]	MVZ CytoMol Commercial Lab (2021)	Genius Digital Diagnostic System (HGDDS)	Germany	4 CT 2 CP (+1 CP for consensus)	Validation	TP
Crowell et al. [34]	Research Division DATEXIM (2019)	CytoProcessor	France	9 investigators	Comparison with TIS	TP
Crowell et al. [35]	Research Division DATEXIM (2019)	CytoProcessor	France	4 investigators	Comparison with a microscope for challenging slides	Nova
Chantziantoniou et al. [36]	CellPathology Plus (2022)	BestCyte	USA	1 CT	Comparison with manual microscopy	TP
Delga et al. [37]	University of Liege (2014)	BestCyte	Belgium	6 CT 2 CP	Comparison with manual microscope	TP and BP
Bai et al. [41]	The Third Affiliated Hospital of Zhengzhou University (2024)	AICyte	China	2 CP (+1 CP for consensus)	Assessment of efficacy and accuracy	TP
Zeng et al. [47]	Collaborative Study (2024)	AICyte	China	0 (AI-alone)	Comparison with original interpretation	TP
Wang et al. [43]	Collaborative Study (2024)	Artificial Intelligence Cervical Cancer Screening (AICCS) system	China	6 CP	Model development and validation	LBC
Yang et al. [44]	Collaborative Study (2024)	Cell Image Analysis System (CIAS)	China	2 CT for ASC-US+	Comparison of with LBC, HPV testing combinations in high-risk population	TP
Zhu et al. [45]	Collaborative Study (2021)	Artificial Intelligence-Assistive TBS (AIATBS) Diagnostic Solution	China	20 CT 3 CP	Model development and validation	LBC
Xue et al. [42]	Collaborative Study (2023)	Cytologist-In-The-Loop Artificial Intelligence (CITL-AI)	China	2 expert CT for GT (+1 expert CT for consensus) 9 Junior CT 9 Senior CT	Model development and validation	SP TP LBC

BP, BestPrep; CP, cytopathologists; CT, cytologists; LBC, liquid-based preparation system; Nova, NovaPrep; SP, SurePath; TBS, The Bethesda System; TP, ThinPrep.

instance, the CITL-AI and AICyte systems offer two possible outcomes: negative or positive [41, 42, 46]. The Landing CytoScanner system provides a 3-option classification: normal, abnormal, or unsatisfactory [39, 40]. The AICCS system gives 5 possible diagnoses: negative for intraepithelial lesion or malignancy, atypical squamous cells of unknown significance, LSIL, high-grade intraepithelial lesion and above, and atypical glandular cells (NILM, ASC-US, LSIL, high-grade squamous intraepithelial lesion, and AGC) [43]. The AIATBS follows the 2014 The Bethesda System (TBS) classification for specific cervical diagnoses, including squamous, glandular, and infectious lesions [45]. CIAS automatically provides analysis-specific results such as negative and positive, TBS grades, a description of the interpreted results, and assessment of slide scan quality [44]. In contrast to the above, systems like CytoProcessor™, BestCyte®, and the Hologic Genius Diagnostic System serve as an aid but do not directly provide diagnostic classifications [19, 34–38].

Performance Metrics of Different AI-Assisted Systems

This review examined 13 randomly selected English language publications available in PubMed that studied 8 AI-related cytology diagnostic tools, including Artificial Intelligence-Assistive TBS diagnostic solution (AIATBS), AICCS, Artificial Intelligence-Assistive Cytology diagnostic system (AICyte), BestCyte, cervical cancer CIAS, CITL-AI, CytoProcessor, and HGDDS. We included studies that focused on different systems and different countries with a robust number of cases. An overview of the study characteristics of these 13 publications is shown in Table 1. A total of 6 of the studies involved a patient population from various locations in China, while the remaining 7 were from Europe and the USA. In all, 6 of the studies were framed as a validation study, and 7 were comparison studies with the AI tool and conventional (manual) Pap test review. One study (Zeng et al. [47], study of AICyte) included an AI-only review with no human reviewer, and two studies referenced reviewers as investigators without noting their credentials. The remaining studies included cytologists and/or cytopathologists (CPs). Ten of the 13 studies included ThinPrep slides (8 of those 13 were solely ThinPrep, while the remaining 5 had multiple slide preparation types included), 1 included NovaPrep, 1 included BestPrep, 1 included SurePath, and 3 included liquid-based Paps without further designation (2 of the 3 were solely liquid-based Paps, while the remaining one study included cases with other preparation methods).

The case overview, including total number of cases, diagnoses examined, availability of corresponding or subsequent histology, and ground truth reference, is summarized in Table 2. In total, more than 245,500 cases were included in the verification phases of this group of publications, with an additional 101,195 cases used for AI model development. Histology for ground truth diagnosis was available for more than 4,600 cases (available in 7 of the 13 studies). The ground truth in the remainder of the studies was the original signed out Pap diagnosis or a consensus diagnosis involving an additional reviewer for discordant cases. TBS was the diagnostic system most often used, although there was variation on which diagnoses were included in the study. For example, cases that were unsatisfactory for diagnosis were included in 3 of the 13 publications, and malignancy (adenocarcinoma or squamous cell) was included in 9 of the 13 studies.

Diagnostic accuracy assessments are presented in Table 3. Several publications used the signed-out Pap diagnosis as the ground truth, while others used the histology follow-up, if applicable. In order to aid with comparisons, diagnostic accuracy assessments (including sensitivity [Se], specificity [Sp], positive predictive value, negative predictive value, and accuracy) were recalculated for the studies that provided data tables with the original signed-out diagnosis, follow-up histology diagnosis, or both.

For the 3 HGDDS studies, an overall sensitivity was estimated using the original Pap test interpretation as the ground truth, since the histology follow-up diagnosis was not provided for one of the studies [11]. Assessing positivity as ASC-US diagnoses or higher, excluding unsatisfactory specimens, there were 1,246 true positives and 189 false negatives, resulting in a sensitivity estimate of 86.8%.

The combined estimated sensitivity for both CytoProcessor studies includes 542 true positives and 24 false negatives, resulting in a sensitivity estimate of 95.8%. The BestCyte sensitivity was similarly high, 92.7–100% (a combined estimated sensitivity for both BestCyte studies was not able to be calculated from the data reported). AICyte sensitivity for AI-alone assessments (at a 50% negative cutoff) was 90.8% (Zeng et al. [47]) and 99.3% (Bai et al. [41]), while the AICyte-assisted was 94.9% and the conventional pap test screening was 95.1% (a combined estimated sensitivity for both AICyte studies was not able to be calculated from the data reported). AICCS showed increased sensitivity with AICCS-assisted CP (100.0%) compared to CP alone (92.9%) or AICCS-alone (92.3%). CIAS sensitivity for AI-alone assessment was 95.8%, higher than the 70.0% sensitivity reported for the

Table 2. Case overview of AI-related studies

Lead author	AI system tested	Cases, *n*	Diagnostic categories	GT	HC
Elishaev (manuscript forthcoming)	HGDDS	890	NILM, ASC-US, LSIL, ASC-H, HSIL, AGC, ADC, SCC	Histology (575 of 890 cases)	Y
Cantley et al. [11]	HGDDS	319	Unsat, NILM, ASC-US, LSIL, ASC-H, AGC, HSIL, malignant	Original Pap (all 319)	N
Ikenberg et al. [38]	HGDDS	1994	NILM, ASC-US, AGC (NOS), LSIL, ASC-H, AGC (FN), HSIL, AIS, SCC, ADC	Histology (555 of 1994 cases)	Y
Crowell et al. [34]	CytoProcessor	1,352 (including 1,360 diagnoses due to 8 cases diagnosed twice)	NILM, ASC-US, LSIL, ASC-H, HSIL, SCC, AGC, AGC (neo), AIS, ADC	Consensus diagnosis between TIS and CytoProcessor (discordances reviewed by 7-member consensus committee)	N
Crowell et al. [35]	CytoProcessor	309 (216 for diagnostic performance [with 210 included in the analysis] and 93 for diagnostic duration)	NILM, ASC-US, LSIL, ASC-H, HSIL, SCC, AGC, AGC (neo), AIS, ADC	Consensus diagnosis between TIS and CytoProcessor (discordances reviewed by 4-member consensus committee)	N
Chantziantoniou [36]	BestCyte	500	UNSAT, NILM, ASC-US, ASC-H, AGUS, LSIL, HSIL, CA	Adjudicative WSI rescreening	N
Delga et al. [37]	BestCyte	105	NILM, ASC-US, AGC, LSIL, ASC-H, HSIL, Unsat	Consensus diagnosis between 2 CPs for ThinPrep and BestPrep collection methods and HC when available	N
Bai et al. [41]	AICyte	32,451	NILM, ASC-US, LSIL, ASC-H, HSIL/SCC, AGC/AIS/ADC	Histology for 1,156 cases	Y
Zeng et al. [47]	AICyte	163,848	NILM, ASC-US, LSIL, ASC-H, HSIL, AGC	Original Pap (all 163,848) HC for 26 SO LSIL+ AI-neg cases (49 cases reviewed, 26 had available histo)	Y
Wang et al. [43]	AICCS	4,588 for validation; 11,468 for model development (16,056 total)	NILM, ASC-US, LSIL, ASC-H, HSIL, SCC, AGC	Histology for 165 of 2,152 cases from internal validation set; expert CP for 608 randomized observational trials	Y
Yang et al. [44]	CIAS	1,231	NILM, LSIL, HSIL, AIS, cervical cancer	Histology for all 1,231 patients	Y
Zhu et al. [45]	AIATBS	34,403 for validation; 81,727 for model development (116,130 total)	NILM, EMC, TRI, CAN, HSV, CC, ACTINO, ASC-US, LSIL, ASC-H, HSIL, SCC, AGC NOS, AGC FN, AIS, ADC	Histology for 980 cases with biopsy results	Y
Xue et al. [42]	CITL-AI	3,514 for validation (8,000 for model development)	Negative, ASC-US, LSIL, ASC-H, HSIL	Expert cytology consensus	N

ACTINO, actinomycetes; ADC, adenocarcinoma; AGC, atypical glandular cells; AGC (FN)/AGC (neo), atypical glandular cells, favor neoplastic; AGC (NOS), atypical glandular cells (not otherwise specified); AIS, adenocarcinoma in situ; ASC-H, atypical squamous cell, high grade not excluded; ASC-US, atypical squamous cells of unknown significance; CAN, *Candida albicans*; CC, clue cell; EMC, endometrial cells; GT, ground truth; HC, histological comparison; HSIL, high-grade squamous intraepithelial lesion; HSV, herpes simplex virus; LSIL, low-grade squamous intraepithelial lesion; NILM, negative for intraepithelial lesion or malignancy; SCC, squamous cell carcinoma; TRI, *Trichomonas vaginalis*; Unsat, unsatisfactory.

Table 3. Diagnostic accuracy assessments of AI-related studies

Study	Metrics							Calculation notes
	Se	Sp	PPV	NPV	Accuracy	AUC	Concordance	
HGDDS – Elishaev (manuscript forthcoming)	HC as GT ASC-US+ Pap, CIN 1+ bx: • HGDDS 97.5% (TP = 465, FN = 12) • SO 99.8% (TP = 476, FN = 1) SO as GT ASC-US+: • HGDDS 90.9% (TP = 510, FN = 51)	HC as GT ASC-US+ Pap, CIN 1+ bx: • HGDDS 48.0% (TN = 47, FP = 51) • SO 13.3% (TN = 13, FP = 85) SO as GT ASC-US+: • HGDDS 76.9% (TN = 253, FP = 76)	HC as GT ASC-US+ Pap, CIN 1+ bx: • HGDDS 90.1% (TP = 465, FP = 51) • SO 84.8% (TP = 476, FP = 85) SO as GT ASC-US+: • HGDDS 87.0% (TP = 510, FP = 76)	HC as GT ASC-US+ Pap, CIN 1+ bx: • HGDDS 79.7% (TN = 47, FN = 12) • SO 92.9% (TN = 13, FN = 1) SO as GT ASC-US+: • HGDDS 83.2% (TN = 253, FN = 51)	HC as GT ASC-US+ Pap, CIN 1+ bx: • HGDDS 89.0% (512/575) • SO 85.0% (489/575) SO as GT ASC-US+: • HGDDS 85.7% (763/890)	NP	• Rx CP κ = 0.909 • HGDDS and TP overall κ = 0.53 • HGDDS and TP 2-category (NILM vs. ASC-US+) κ = 0.689	Se, Sp, PPV. NPV, accuracy: • HC as GT calculated from manuscript Table 4 • SO as GT calculated from manuscript Table 3
HGDDS – Cantley et al. [11]	SO as GT ASC-US+ • HGDDS 73.0% (TP = 108, FN = 40) • LM 68.2% (TP = 103, FN = 48)	SO as GT ASC-US+: • HGDDS 88.3% (TN = 128, FP = 17) • LM 79.9% (TN = 115, FP = 29)	SO as GT ASC-US+: • HGDDS 86.4% (TP = 108, FP = 17) • LM 78.0% (TP = 103, FP = 29)	SO as GT ASC-US+: • HGDDS 76.2% (TN = 128, FN = 40) • LM 70.6% (TN = 115, FN = 48)	SO as GT 8-category: • HGDDS = 62.1% (198/319) • LM = 55.8% (178/319)	NP	3-category (Unsat, NILM, ASC-US+): • HGDDS 76.8% (245/319) • LM 71.5% (228/319) CT: • HGDDS 75.5% (241/319) • LM 72.7% (232/319) CP: • HGDDS 88.3% (n = 106) • LM 81.7% (n = 98)	Se, Sp, PPV, NPV, accuracy calculated from manuscript Tables 2 and 3 excluding Unsat
HGDDS – Ikenberg et al. [38]	HC as GT ASC-US+ Pap, CIN 2+ bx: • HGDDS 99.4% • TIS 98.8% SO as GT ASC-US+: • HGDDS 86.5% (TP = 628, FN = 98)	HC as GTASC-US+ Pap, CIN 2+ bx: • HGDDS 6.7% • TIS 2.9% SO as GT ASC-US+: • HGDDS 93.4% (TN = 1183, FP = 83)	HC as GT ASC-US+ Pap, CIN 2+ bx: • HGDDS 63.9% • TIS 62.8% SO as GT ASC-US+: • HGDDS 88.3% (TP = 628, FP = 83)	HC as GT ASC-US+ Pap, CIN 2+ bx: • HGDDS 87.5% • TIS 60.0% SO as GT ASC-US+: • HGDDS 92.3% (TN = 1183, FN = 98)	Concordance of HGDDS and TIS with HC 90.4%	NP	• HGDDS and TIS overall: 81.2% (1618/1992) • 3-category (neg, low severity, high severity): 86.7% (1727/1992)	For SO as GT, Se, Sp, calculated from manuscript Table 3 (excluded 2 HGDDS Unsat)PPV, NPV calc
CytoProcessor (TP) Crowell et al. [34]	ASC-US+: • CytoProcessor 95.9% (TP = 488, FN = 21) • TIS 89.4% (TP = 455, FN = 54)	ASC-US+: • CytoProcessor 80.8% (TN = 688, FP = 163) • TIS 87.9% (TN = 748, FP = 103)	ASC-US+: • CytoProcessor 75.0% (TP = 488, FP = 163) • TIS 81.5% (TP = 455, FP = 103)	ASC-US+: • CytoProcessor 97.0% (TN = 688, FN = 21) • TIS 93.3% (TN = 748, FN = 54)	2-category (normal/abnormal): • CytoProcessor 86.5% (1,176/1,360) • TIS 88.5% (1,203/1,360)	NP	CytoProcessor and TIS concordant in 68.8% (936/1,360)	Se, Sp, PPV, NPV calculated from manuscript Tables 6 and 7
CytoProcessor (Nova) – Crowell et al. [35]	ASC-US+: • CytoProcessor 94.7% (TP = 54, FN = 3) • MM 93.0% (TP = 53, FN = 4)	ASC-US+: • CytoProcessor 85.6% (TN = 131, FP = 22) • MM 92.8% (TN = 142, FP = 11)	ASC-US+: • CytoProcessor 71.1% (TP = 54, FP = 22) • MM 82.8% (TP = 53, FP = 11)	ASC-US+: • CytoProcessor 97.8% (TN = 131, FN = 3) • MM 97.3% (TN = 142, FN = 4)	2-category (normal/abnormal): • CytoProcessor 88.1% (185/210) • MM 92.9% (195/210)	NP	NP	Se, Sp, PPV, NPV calculated from manuscript Tables 5 and 6

Acta Cytologica 2026;70:52–71
DOI: 10.1159/000546182

Table 3 (continued)

Study	Metrics							Calculation notes
	Se	Sp	PPV	NPV	Accuracy	AUC	Concordance	
BestCyte – Chantziantoniou [36]	ASC-US+ 92.7% (TP = 127, FN = 10)	ASC-US+ 99.4% (TN = 344, FP = 2)	ASC-US+ 98.4% (TP = 127, FP = 2)	ASC-US+ 97.2% (TN = 344, FN = 10)	8-category diagnostic match 96.6% (483/500)	NP	Overall κ = 0.93	Se, Sp, PPV, NPV, accuracy calculated from manuscript Tables 1 and 2 (excluded Unsat from Se, Sp, PPV, NPV)
BestCyte – Delga et al. [37]	ASC-US+, HSIL+ as truth: • BestPrep: – BestCyte 100% – Manual 100% • ThinPrep: – BestCyte 100% – Manual 95.6% ASC-US+, LSIL+ as truth: • BestPrep: – BestCyte 85.7% – Manual 85.7% • ThinPrep: – BestCyte 80.0% – Manual 80.0%	ASC-US+: • BestPrep: – BestCyte 92.8% – Manual 88.1% • ThinPrep: – BestCyte 95.2% – Manual 88.1%	NP	NP	NP	NP	NP	Data for GT in publication refers to 42 confirmed negative, 23 HSIL, 35 LSIL
AICyte – Bai et al. [41]	Detection of CIN I+: • AICyte 94.9% (TP = 779, FN = 42) • SO 95.1% (TP = 781, FN = 40) Detection of CIN 2+ • AICyte-alone at 50% neg cutoff 99.3%	Detection of CIN I+: • AICyte 31.3% (TN = 105, FP = 230) • SO 7.5% (TN = 25, FP = 310) Detection of CIN 2+ • AICyte-alone at 50% neg cutoff 9.9%	Detection of CIN I+: • AICyte 77.2% (TP = 779, FP = 230) • SO 71.6% (TP = 781, FP = 310) Detection of CIN 2+ • AICyte-alone at 50% neg cutoff 26.5%	Detection of CIN I+: • AICyte 71.4% (TN = 105, FN = 42) • SO 38.5% (TN = 25, FN = 40) Detection of CIN 2+ • AICyte-alone at 50% neg cutoff 97.7%	Detection of CIN I+: • AICyte 76.5% (884/1156) • SO 69.7% (806/1156) Detection of CIN 2+ • AICyte-alone at 50% neg cutoff 31.9% Dx category match • AICyte 63.6% • SO 57.4%	NP	• AICyte Rx CP κ = 0.65 • AICyte 2-category (neg, ASC-US+) Rx CP κ = 0.74 • AICyte and SO κ = 0.61 • 2-category (neg, ASC-US+) AICyte and SO κ = 0.68	Dx category match for accuracy calculated from manuscript Tables 3 and 4, (ASC-US/LSIL Pap concordant with CIN I Bx, ASC-H/HSIL/SCC Pap concordant with CIN 2/3/SCC Bx)
AICyte – Zeng et al. [47]	AICyte-alone at 50% neg cutoff for ASC-US+ 90.8%	AICyte-alone at 50% neg cutoff for ASC-US+ 49.7%	AICyte-alone at 50% neg cutoff for ASC-US+ 10.6%	AICyte-alone at 50% neg cutoff for ASC-US+ 98.8%	NP	NP	NP	
AICCS – Wang et al. [43]	Rand Obs Tr: • AICCS-alone 92.3% • AICCS-assisted CP 100.0% • CP-alone 92.9% HC as GT • AICCS-alone 96.5% • CP-alone 100.0%	Rand Obs Tr: • AICCS-alone 85.4% • AICCS-assisted CP 98.9% • CP-alone 98.4%	NP	Rand Obs Tr: • AICCS-alone 99.4% • AICCS-assisted CP 100.0% • CP-alone 98.9%	Rand Obs Tr: • AICCS-alone 85.9% • AICCS-assisted CP 99.0% • CP-alone 98.0% HC as GT • AICCS-alone 83.6% • CP-alone 86.1%	Rand Obs Tr: • AICCS-alone 94.4% • AICCS-assisted CP 99.5% • CP-alone 95.9%	NP	HC as GT for subset of validation (1 location)
CIAS – Yang et al. [44]	HSIL+: • AI 95.8% • LBC 70.0% • HPV 95.0% • LBC+HPV 89.9% • AI+LBC 68.3% • AI+HPV 93.7% • HPV followed by LBC 87.8%	HSIL+: • AI 28.3% • LBC 41.7% • HPV 8.8% • LBC+HPV 24.0% • AI+LBC 44.3% • AI+HPV 21.8% • HPV followed by LBC 26.7%	HSIL+: • AI 28.9% • LBC 26.8% • HPV 24.1% • LBC+HPV 26.5% • AI+LBC 27.1% • AI+HPV 26.7% • HPV followed by LBC 26.7%	HSIL+: • AI 95.7% • LBC 82.1% • HPV 86.5% • LBC+HPV 88.7% • AI+LBC 82.1% • AI+HPV 92.0% • HPV followed by LBC 87.8%	HSIL+: • AI 44.0% • LBC 48.3% • HPV 29.0% • LBC+HPV 39.4% • AI+LBC 50.0% • AI+HPV 38.6% • HPV followed by LBC 40.9%	HSIL+: • AI 62.1% • LBC 59.7% • HPV 52.1% • LBC+HPV57.0% • AI+LBC 56.2% • AI+HPV 57.8% • HPV followed by LBC 57.2%	NP	

Table 3 (continued)

Study	Metrics							Calculation notes
	Se	Sp	PPV	NPV	Accuracy	AUC	Concordance	
AIATBS – Zhu et al. [45]	Overall: • Intraepithelial lesions 92.0% • Other lesions 83.0% 6 Sr CT: Intraepithelial lesion: • AIATBS 90.8% • CT manual 86.1% Other lesions: • AIATBS 84.2% • CT manual 76.3%	Overall: • Intraepithelial lesions 84.4% • Other lesions 97.8% 6 Sr CT: • AIATBS 81.9% • CT manual 90.2%	NP	NP	Squamous intraepithelial lesions • AIATBS 75.2% • CT manual 80.2%	NP	NP	
CITL-AI – Xue et al. [42]	• Overall stand-alone AI 89.4% • Jr CT 53.1% • Jr CT CITL-AI 81.6% • Sr CT 84.4% • Sr CT CITL-AI 87.2%	• Overall stand-alone AI 66.4% • Jr CT 66.2% • Jr CT CITL-AI 78.9% • Sr CT 89.9% • Sr CT CITL-AI 91.5%	NP	NP	NP	• Overall stand-alone AI 77.9% • Jr CT 59.7% • Jr CT CITL-AI 80.2% • Sr CT 87.1% • Sr CT CITL-AI 89.3%	• Sr CT and Jr CT κ = 0.791 • Stand-alone AI and Jr CT κ = 0.707 • Stand-alone AI and Jr CT CITL-AI κ = 0.822 • Stand-alone AI and Sr CT κ = 0.806 • Stand-alone AI and Sr CT CITL-AI κ = 0.818	

Bx, biopsy; GT, ground truth; HC, histological comparison; K, kappa; NP, not provided; Rand Obs Tr, randomized observational trial; SO, signed-out Pap diagnosis.

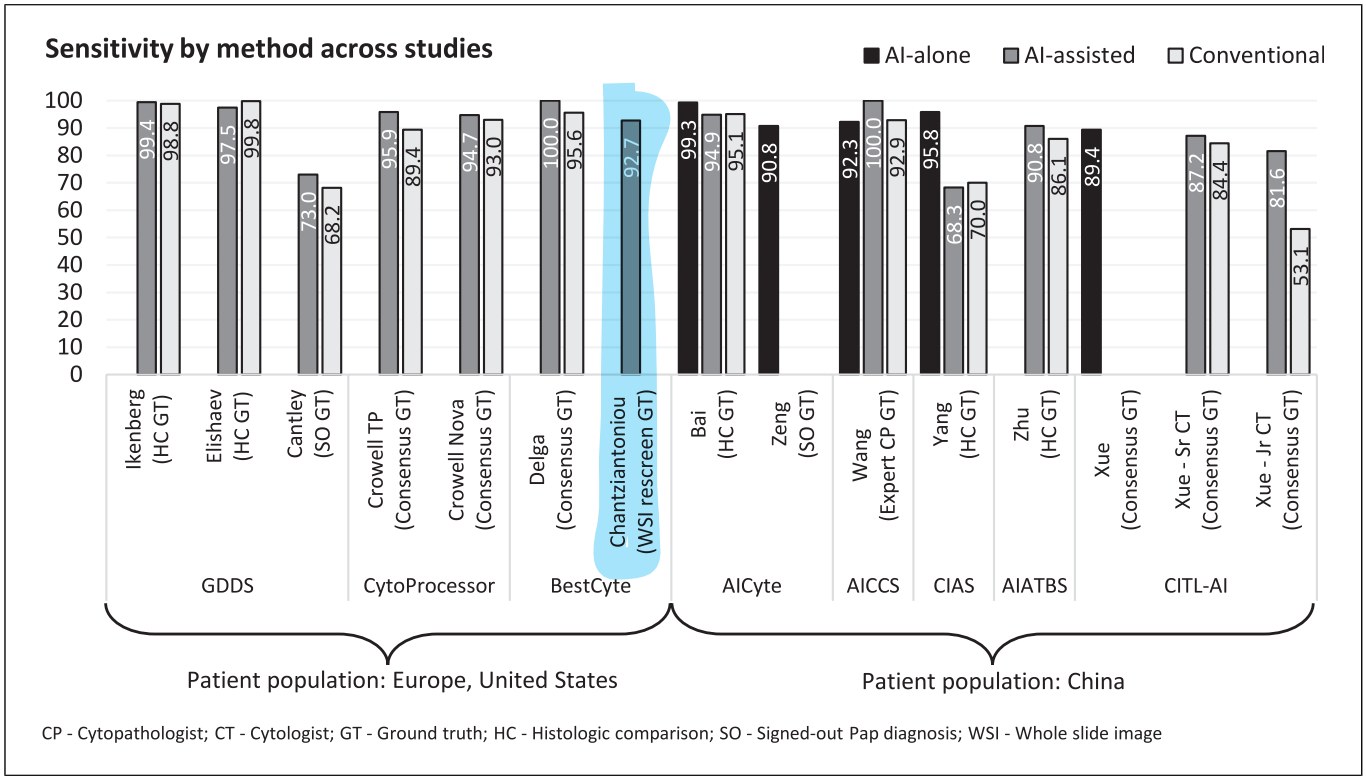

Fig. 2. Sensitivity by method across studies.

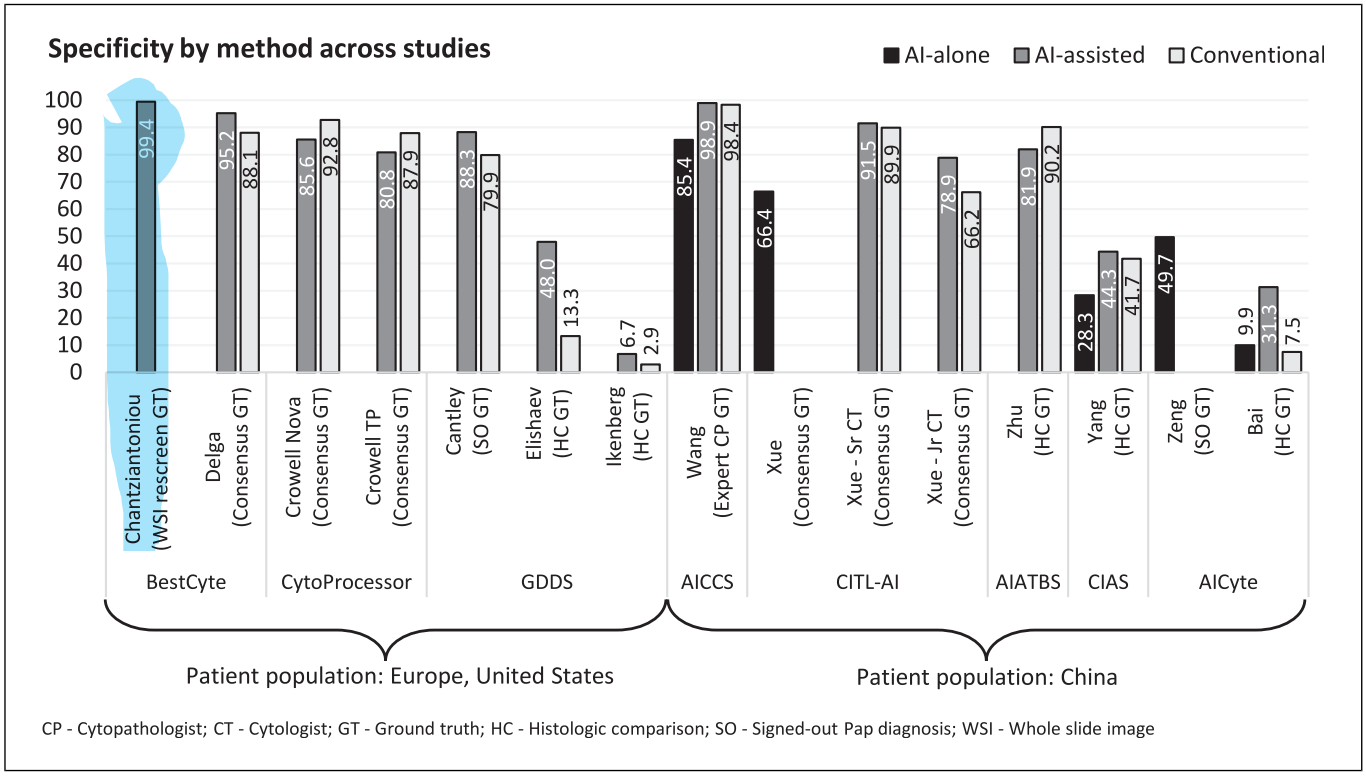

Fig. 3. Specificity by method across studies.

Table 4. Efficiency of AI-related studies

Study	Review time		Reduction %
	AI	conventional	
HGDDS – Elishaev (manuscript forthcoming)	• HGDDS CT M = 70 s • HGDDS CP M = 38 s	• TIS CT est. range = 180–300 s* • TIS CP est. range = 60–180 s*	• HGDDS CT 157% faster[1] • HGDDS CP 58% faster*
HGDDS – Cantley et al. [11]	• Overall HGDDS M = 192 s • CT HGDDS M = 228 s • CP HGDDS M = 72 s	• Overall light microscopy M = 354 s • CT light microscopy M = 426 s • CP light microscopy M = 132 s	• Overall HGDDS 84% faster • HGDDS CT 87% faster • HGDDS CP 83% faster
HGDDS – Ikenberg et al. [38]	• HGDDS 44.8 s	• TIS 89.9 s	• HGDDS 101% faster
CytoProcessor (TP) – Crowell et al. [34]	• Human + machine workflow: CytoProcessor 595 s/slide • Human only workflow: CytoProcessor 290 s/slide	• Human + machine workflow: TIS 594 s/slide • Human only workflow: TIS 436 s/slide	• Human + machine workflow CytoProcessor 0% faster • Human only workflow Cytoprocessor 50% faster
CytoProcessor (Nova) – Crowell et al. [35]	• CytoProcessor M = 99 s	• Manual microscopy M = 163 s	• CytoProcessor 65% faster
BestCyte – Chantziantoniou [36]	• Overall M = 82.8 s • Overall WSI rescreen M = 236.4 s	NP	NA
BestCyte – Delga et al. [37]	• BestCyte 30–180 s/slide[2]	• est. range 300–500 s/slide[2]	• BestCyte 281% faster[2]
AICyte – Bai et al. [41]	• AICyte M = 22.23 s/case	• Manual M = 180 s/case	• AICyte 710% faster
AICyte – Zeng et al. [47]	NP	NP	NA
AICCS – Wang et al. [43]	NP	NP	NA
CIAS – Yang et al. [44]	NP	NP	NA
AIATBS – Zhu et al. [45]	• Natural sedimentation M = 66.3 s • Membrane sedimentation M = 171.8 s • Centrifugal sedimentation M = 85.3 s	NP	NA
CITL-AI – Xue et al. [42]	• Reduced cytology workload by 37.5% with CITL-AI	NP	NA

Est, estimated; M, mean; NA, not applicable; NP, not provided; Sec, seconds; WSI, whole slide image. [1]For Elishaev conventional times when a range is provided, the lowest value of the range is used for the reduction rate calculation to demonstrate AI performance for most optimal conventional time. [2]For Delga review times, since both AI and conventional are reported as ranges, comparison performed on the mid-point.

conventional review of LBC, and AI+LBC sensitivity (68.3%). AIATBS sensitivity for detecting intraepithelial lesions was 92.0%. CITL-AI sensitivity had a notable increase for junior cytologists (81.6% CITL-AI compared to 53.1% for junior cytologists alone) and a small increase for senior cytologists (87.2% CITL-AI compared to 84.4% for senior cytologists alone), with the highest noted for stand-alone AI (89.4%).

A visual representation of the sensitivity values by method across studies is shown in Figure 2. When grouped by patient population, the average sensitivity among the China patient population (AICyte, AICCS, CIAS, AIATBS, CITL-AI) is 93.5% for AI-alone, 87.1% for AI-assisted, and 80.3% for manual Pap test review. The average sensitivity among the European and US patient populations (HGDDS, CytoProcessor, BestCyte) is 93.3% for AI-assisted and 90.8% for manual Pap test review. AI-alone results were not reported among the latter studies.

The BestCyte specificity was as high as 95.2%–99.4% (a combined estimated specificity for both BestCyte studies was not able to be calculated from the data reported). The combined estimated specificity for both CytoProcessor studies included 819 true negatives and 185 false positives, resulting in a specificity estimate of 81.6%. For the three HGDDS published studies, an overall specificity was estimated using the original Pap test interpretation as the ground truth since histology follow-up diagnoses were not provided for one of the studies [11]. Assessing positivity as ASC-US diagnoses or higher, excluding unsatisfactory specimens, there were 1,564 true negatives and 176 false positives, resulting in a specificity estimate of 89.9%.

AICCS showed increased specificity with AICCS-assisted CP (98.9%) compared to CP alone (98.4%) or AICCS-alone (84.4%). CITL-AI specificity increased for junior cytologists (78.9% CITL-AI compared to 66.2% for junior cytologists alone) and senior cytologists (91.5% CITL-AI compared to 89.9% for senior cytologists alone), with the stand-alone AI having the lowest specificity (66.4%).

AIATBS specificity for intraepithelial lesion detection was 84.4%. CIAS specificity for AI-alone assessment was 28.3%, lower than the 41.7% specificity reported for the conventional Pap test review of LBC cases, and AI+LBC sensitivity (44.3%). AICyte specificity for AI-alone assessments (at a 50% negative cutoff) were 49.7% (Zeng, ASC-US+) and 9.9% (Bai, detection of CIN 2+), while the AICyte-assisted was 31.3% and conventional was 7.5% (a combined estimated specificity for both AICyte studies was not able to be calculated from the data reported).

A visual representation of the specificity values by the method across studies is shown in Figure 3. When grouped by patient population, the average specificity among the China patient population (AICyte, AICCS, CIAS, AIATBS, CITL-AI) is 47.9% for AI-alone, 71.1% for AI-assisted, and 65.7% for manual Pap test review. The average specificity among the Europe and US patient populations (HGDDS, CytoProcessor, BestCyte) is 72.0% for AI-assisted and 60.8% for manual pap test review, with no AI-alone results reported among the latter studies.

Table 4 provides the workflow efficiency measures for these published studies. Among the 13 studies selected for analysis, 10 reported workflow assessments (no time comparison was provided for the publications from Zeng et al. [46] [AICyte], Wang et al. [43] [AICCS], or Yang et al. [44] [CIAS]). These studies varied on their time reporting; therefore for ease of comparison, all time was converted to seconds.

Among the HGDDS studies, the use of the AI tool reduced review time overall by more than 80%. Crowell et al. [34, 35] showed a reduction in review time for the CytoProcessor for screening ThinPrep Pap tests (50% reduction when excluding the machine aspects of the workflow) and NovaPrep (65% reduction).

AICyte [41] showed the most dramatic reduction in review time among the studies, with a 710% reduction in review time (from 180 s/case for a manual review to 22.2 s/case with AICyte). BestCyte [37] also showed a large reduction, with a 281% decrease (using the midpoint of the ranges, there was a reduction from 400 s/slide for the conventional review to 105 s/slide for BestCyte).

There were no comparative data to the conventional review method for Chantziantoniou's BestCyte study (reporting an average time of 82.8 s for patch review and 236.4 s for the WSI rescreen), for the AIATBS (reporting average time based on sediment type, varying from 66.3 s for natural sedimentation and 171.8 for membrane sedimentation), and for the CITL-AI system (citing a 37.5% reduction).

Important Considerations

For this review, we grouped study results in order to have comparable metrics to facilitate discussion. However, it is important to note that the original publications did not all provide diagnostic accuracy assessments, including sensitivity and specificity. When the original publication provided sufficient data to allow for a calculation, the authors completed the calculations as consistently as possible (using ASC-US+ as the positivity threshold, for example). However, values may be skewed for some patient populations depending on the focus of the publication. For example, some studies overrepresented positive cases to provide a sufficient sample of validation for the detection of intraepithelial lesions. Others included large volumes of negative cases to differentiate nonmalignant conditions such as infections.

Barriers and Drivers to Implementation of AI-Assisted Digital Systems in Routine Practice

Digital cytology and AI have the potential for revolutionizing cytopathology practice. Worsening staffing shortages in the face of increasing case volume and the possibility of reduced turnaround time without compromising the quality of diagnosis are all powerful drivers to explore using these new AI-based systems in routine practice. Additional advantages include standardization of diagnosis across different health systems with the opportunity for remote expert consultation, screening of cases remotely to allow end-users to work from home or while waiting outside the lab to perform rapid on-site evaluations, educational benefits including virtual education, ease of conducting image-based research, and easy storage of digital images as opposed to glass slide storage. Conducting quality assurance activities with well-maintained digital images is less expensive and likely requires less labor than the current manual method of performing this essential lab function. One of the most important drivers is that AI-assisted systems have the potential to improve accuracy and for the most part have good sensitivity (>80%) and specificity (>80%) [48].

On the other hand, the application of WSI technology in everyday clinical cytopathology practice has its own unique challenges. First and foremost, it is important to have leadership with a vision and strong motivation to tackle this challenging task of planning and implementing a new, disruptive system in everyday practice. Technical considerations such as available infrastructure and resources (cloud vs. on-premise deployment, IT analyst support) and mindset and participation of personnel working in the trenches (i.e., training, change management, and proficiency testing) are some of the main factors determining adoption or rejection of such systems.

Another important technical challenge is the fact that cytology slides often contain three-dimensional (3D) cell clusters and background artifactual materials such as blood and mucus. LBC, while having less background artifact and appearing more flat, may still have 3D cell aggregations. The AI-assisted digital systems, despite using volumetric imaging techniques in some devices, present 3D cell clusters in 2D digital format to the reviewer. The reviewer will have to adapt to this nuanced difference in morphology seen with this change. Cytology-specific WSI solutions, including cell detection algorithms, higher magnification scanning, multilayer/volumetric scanning, and easy-to-use navigation and annotation tools, can help address these concerns to some extent [49, 50]. However, this is a learning curve that CPs and cytologists will have to scale in order to accurately interpret different morphologies presented in a digital format. The cytology community had to overcome a similar transition and learn new cytomorphological criteria when moving from conventional smears to LBC.

AI-Assisted Workflow

The advent of DP and AI has necessitated the reassessment of routine workflows in cytology laboratories. Depending upon the type of AI-assisted system being implemented (interpretative with NFR feature as opposed to AI-assisted diagnostic systems), the cytology workflow process is going to be transformed, especially in the transition phase; this almost seems counterproductive. Recruiting tech-savvy cytology personnel willing to work with these newer systems, restructuring of personnel including information technology (IT) support that understands how to troubleshoot AI-related technology, and adjusting physical lab space to accommodate digital systems instead of microscopes, as well as additional time earmarked for implementation purposes including training and practice with these newer systems in the initial phase of deployment, have to be factored in the plan for the adoption of digital signout. Moreover, digitization will require additional steps in the workflow process with personnel needed for properly barcoding slides, making sure that slides are dry and free of sticky mounting medium, loading and unloading slides, assigning cases to screeners in a digital system, troubleshooting technical issues, and managing the system in general [20]. The transition between glass and AI-assisted digital signout may be challenging with some cases requiring simultaneous reviews with both methods (e.g., broken glass slide or foreign slide from another lab that is unable to be scanned), potentially leading to additional delays in turnaround time. The cost alone for this new technology, including data storage, may be prohibitive to some labs. Hence, a meticulous business plan with a calculated strategy along with a cost-benefit analysis is essential before commencing with integrating any new system in clinical practice.

In our opinion, developing trust with the digital system and confidently being able to use the system for clinical signout will depend on a multitude of factors including leadership support, adequate training, reliable validation to ensure accuracy and safety, evidence-based guidelines, endorsement by reputable national and international organizations and peer individuals, malpractice coverage in this context, and actual experience of working with these systems.

Challenges with Clinical Trials Using AI-Assisted Platforms

With the addition of new technologies, it is important to conduct studies and trials to ensure clinical performance and patient safety, either as clinical trials for regulatory approvals or as validation studies once the system is brought into the laboratory. The recently published white papers from the American Society of Cytopathology Digital Cytology Task Force provides guidance for clinical validation of whole slide imagers in cytology, but stopped short of providing guidance for the validation of AI given "insufficient data and wide variability in AI algorithms in cytology" [20, 48]. Results from the accompanying Task Force survey found that the majority of respondents believe that a validation study of AI algorithms in cytology would need between 100 and 200 cases. When conducting a clinical trial for regulatory approval, the sample size is significantly higher [51]. The number of cases needed for a study or trial is critically important, given that they are limited by the cytologist and pathologist resources available to review study cases outside of their clinical work. The workforce shortage compounded by the fact that studies should attempt to mimic clinical workflow and samples, reflecting that "the diversity and complexity of the samples encountered in the laboratory" are important issues that have a tremendous impact on the will and ability of the labs to conduct such studies [20, 48].

In cervical cancer screening, TBS for Reporting Cervical Cytology is widely used to classify Pap test results [52]. Including cases representative of all diagnostic categories in experimental studies presents the challenge of how to incorporate these cases given the low prevalence of cervical disease. For example, in the 2023 College of American Pathologist's (CAP) Cytopathology Checklist, the median reporting rate for an abnormal ThinPrep® Pap test was 8.8%. The majority of these were ASC-US cases, accounting for 5.4% of all cases. The more severe diagnostic categories (atypical squamous cells cannot exclude a high-grade squamous intraepithelial lesion [ASC-H] and above) only account for 1.0% of cases in clinical practice [53]. Thus, in a study with 200 cases that are intended to be representative of clinical practice, it would be expected that more than 180 of them would be negative, 11 would be ASC-US, and only 2 would be ASC-H or more severe. Because this would limit the ability to properly assess new technology, particularly AI algorithms built to detect disease, these studies are often seeded with abnormal cases to ensure appropriate representation of the most severe disease cases [11, 54, 55]. It is common to seed with approximately 50% abnormal cases with the resulting study set having a larger percentage of abnormal cases than would be observed in a screening population, and also potentially have a different percentage of each disease category. For example, it would be quite rare to observe cervical carcinoma in routine screening; but several examples may be included in a study set in order to properly test an algorithm's ability to detect squamous cell, endocervical and endometrial adenocarcinoma, and other cancers. While this approach deviates from what would be observed in routine clinical practice, it enables researchers to challenge the algorithm with several examples of each diagnostic category without being too burdensome for study reviewers.

It is well established that diagnostic biases exist in all aspects of medicine, although this is reported to be lower within visual specialties such as pathology [56, 57]. While cytologists and pathologists are trained to objectively review and approach each case without being influenced by prior case reviews, biases can still influence results. Confirmation and heuristic biases could impact cytological screening diagnoses and have the potential to influence study results. Confirmation bias refers to the tendency to make a diagnosis by interpreting cytological features in a way that confirms one's preconceptions, and heuristic bias refers to the tendency to make a diagnosis based on how easily the diagnosis comes to mind [58]. Larson et al. [58] hypothesized that if there was a diagnostic bias in the signout of Pap tests, they would expect to see a deviation toward the mean of their typical diagnostic patterns. This was not observed in their study, leading them to conclude that cytologists are not influenced by an expected number of abnormal cases per day. This may be due to the fact that deviations in abnormal prevalence within routine screening are not significant enough to demonstrate bias. Evans et al. [59] compared study sets with varying disease prevalence, mimicking routine screening (5% abnormal) and typical experimental testing scenarios (50% abnormal). While this study was not conducted using typical microscopic screening methods, the authors found that the prevalence of disease in a screening set can influence the behavior of the reviewer and affect overall accuracy [59]. Additional studies have also demonstrated that knowledge of a patient's HPV-positive status can bias the cytologist's diagnosis, which resulted in the overcalling of ASC-US in these cases [60]. These studies highlight the importance of understanding how behavioral responses to changes in clinical information or disease prevalence can influence outcomes among highly trained cytology experts.

Reviewers participating in studies and clinical trials using new AI algorithm technologies for cytology could also be influenced by limited experience after receiving training and may participate prior to overcoming the learning curve. Inexperience associated with the novel digital presentation of diagnostically relevant material, such as in explainable cell galleries, could result in performance data differing from that observed in clinical use [61].

To combat these limitations, it is advantageous to include a control arm (current standard of care) and minimize extraneous variables. For example, a robust study design may include the same cytologist and pathologist reviewing the same cases using a traditional light microscope versus using a digital cytology system with an AI algorithm with a washout period between reviews. Depending on the study objectives, it may also be useful to curate the study cases to minimize interobserver variability, which is often observed in cytology and histology studies, particularly for equivocal cases [62, 63]. This would reduce the real-world applicability of the results and conclusions that could be made, but would also reduce some of the potential variability associated with this field and truly test the AI algorithm's performance compared to a microscopic review. It may also be useful to analyze data according to clinical management rather than by a pathology-based classification system. For example, within cervical cytology, this would mean analyzing negative vs. ASC-US cases and more severe disease to demonstrate that disease is being detected, or analyzing at the ASC-H threshold as these patients would be referred to colposcopy regardless of HPV status [64]. Once the study is complete, conducting a discordant analysis is useful to review any false-negative and false-positive cases to determine the source of the error and whether the AI presented the appropriate diagnostic cells for interpretation. While an exhaustive list is not presented, being cognizant of potential challenges when designing studies to test new AI algorithms will increase the likelihood of a successful study that more accurately captures data reflective of the true performance of the new technology and not simply document confounding human factors [64].

Considerations When Integrating AI in Everyday Clinical Practice

Information Technology Requirements

Sufficient IT resources and experts in this field are invaluable for a smooth and streamlined implementation, including integration of these digital and AI-assisted platforms into the existing laboratory information system. Large institutions and private labs may have the means and access to such resources enabling successful integration, but smaller labs may face difficulties in availing these resources. Furthermore, once DP and AI-assisted platforms are functional, it is important to establish quality control and quality assurance methods. For instance, the FDA has particular specifications that are taken into consideration when clearing medical equipment to be used for clinical practice. Recently, the HGDDS has received clearance from US FDA. However, a Barco monitor is an essential part of the HGDDS, since usage of any other replacement monitor in order to control costs would break the pixel pipeline [65]. Additionally, issues related to data privacy, storage, and data security are paramount when dealing with patient health information. National organizations need to formulate guidelines and recommendations on the optimal storage specifications for AI-related systems, the amount of time digital data as well as scanned glass slides need to be stored, safe methods for discarding data at the end of the recommended storage period, and safe and easy retrieval of digital files or OOIs for clinical practice and legal cases.

Legal and Ethical Considerations

With some AI platforms promoting "NFR" for some cases, it is imperative to think about legal issues that may arise with the widespread adoption of these newer systems. Does an NFR result mean that no human (cytologist or CP) will review the negative results diagnosed by AI? What happens in case of litigation involving a particular case that was interpreted with an AI-assisted system? Who would be responsible for the diagnostic accuracy and litigation claims? Moreover, with newer versions of different systems being developed and made commercially available, what impact will be seen with diagnostic accuracy and how will this work in solving legal queries arising with the diagnoses made with the previous versions of the same systems? With the current proposal for primary HPV testing and direct dual p16/ki67 testing, would it be feasible to implement such AI-assisted diagnostic platforms in labs when future guidelines may change or just recommend molecular testing?

These are some thought provoking controversial issues that would need to be tackled in order to reassure both the cytology community and patient populations before the adoption of these systems. The "blackbox" nature of deep learning algorithms, their continuous adaptation to new data, lack of guidelines on how to perform QA and proficiency testing, and limited literature on liability

issues will need to be addressed. Liability for interpretation will likely remain with the CP who should probably sign out their reports disclosing that AI assistance was used. A supportive legal framework that supports physicians is needed, since cytology personnel as well as companies responsible for developing these algorithms could all be potentially legally responsible for their AI-aided interpretations [48, 66]. In case of utilization of devices with the NFR feature, assigning responsibility becomes more complicated.

Reimbursement Considerations

In the USA, there are different current procedue terminology codes (CPT) used specifically for cervical or vaginal pap tests when using automated screening systems, manual screening or review under physician supervision. However, there is no code for AI-assisted platforms currently when such systems are used in daily cytopathology practice. This means that new codes would likely have to be assigned to cases that are signed out with digital/AI-assisted technology. With the reduction in turnaround time and increased productivity, is a reduction in reimbursement justifiable? On the other hand, would that not deter laboratories from buying this expensive technology, that requires maintenance and cloud storage of digital data that would further increase costs for the lab? Furthermore, some of the challenging cases would need to be reviewed digitally as well as by glass slides. Would the code be different in these instances?

Cytology-Histology Correlation, Training, and Proficiency Testing

Changing the method of signout will impact cyto-histo correlation activities that a cytology laboratory has to conduct to maintain CAP and or/other accreditation status. Would this correlation still be feasible and accurate especially during the transition period? Successful integration of any new system requires training of all personnel involved in utilizing the system. The entire cytology workforce will need training before widespread adoption occurs for these new systems in clinical practice. Trainees (residents, fellows, cytologists) are an important part of our community since they are the future of pathology practice. Should we recommend training with both digital systems as well as glass slides for them? Moreover, how will proficiency testing (PT) for the cytology personnel who have gone digital be conducted? Should the PT be undertaken with glass slides and/or with digital slides? Would we allow retraining and remediation only with AI/digital platforms or allow glass slide interpretations for people with no inclination or aptitude for AI-assisted diagnosis? Several of these questions would need to be answered satisfactorily, as evidenced by the survey results conducted by the ASC digital cytology Task Force where Kim et al. highlighted the need for guidance on the safe implementation, validation, and maintenance of digital cytology technology for cytopathology practice [20, 48]. Lastly, it is important to remember that any change in healthcare processes, especially those involving emerging technologies, takes time and effort as well as a strong desire to overcome various obstacles encountered during the whole process [9].

Conclusion and Future Directions

Judicious review of evidence-based scientific data, an open mindset with long-term vision toward leveraging new technology, and a desire to improve healthcare processes in today's challenging climate of staffing shortages may help drive widespread adoption of new innovative technologies applying AI to screen Pap tests. Additional research and contemplation of issues inherent to utilizing such systems in everyday practice are needed. Some of these include practical solutions for successful training, validation, implementation, continued quality assurance, PT monitoring, and reimbursement. Proper guidance and recommendations for successful integration and a collective collaborative effort by various stakeholders are required to successfully embark on this exciting and new frontier in cytopathology practice.

Acknowledgments

The authors would like to acknowledge the contribution of Colleen Vrbin from Analytical Insights LLC for help with statistical analysis.

Conflict of Interest Statement

Lakshmi Harrinath received grant funding from the American Society of Cytopathology to work on AI-assisted Pap test screening. Sarah Harrington is an employee of Hologic, Inc. Liron Pantanowitz is on the medical advisory board of rIbex and NTp, consults for Hamamatsu and AiXMed, and a cofounder of LeanAP and Placenta AI.

Funding Sources

This study was not supported by any sponsor or funder.

Author Contributions

Lakshmi Harinath, MD, MPH: writing – original draft, methodology, review and editing, data curation, and conceptualization. Xinru Bai, MD: writing – original draft, methodology, review and editing, data curation, and conceptualization. Jeremy Minkowitz, MD and Xianxu, Zeng, MD, PhD: writing – review and editing, and data curation. Sarah Harrington, PhD: writing – review and editing, methodology, and data curation. Chengquan Zhao, MD: writing – review and editing, methodology, conceptualization, and data curation. Liron Pantanowitz, MD, PhD, MHA: writing – review and editing, methodology, conceptualization, supervision, data curation, and resource allocation.

References

1 WHO. Cervical cancer (accessed February 18, 2025).
2 Vilos GA. The history of the papanicolaou smear and the odyssey of George and Andromache papanicolaou. Obstet Gynecol. 1998;91(3):479–83. https://doi.org/10.1016/s0029-7844(97)00695-9
3 Wingo PA, Cardinez CJ, Landis SH, Greenlee RT, Ries LA, Anderson RN, et al. Long-term trends in cancer mortality in the United States, 1930-1998. Cancer. 2003;97(12 Suppl):3133–275. https://doi.org/10.1002/cncr.11380
4 Zhao XL, Hu SY, Hu JW, Wang HH, Wen TM, Feng YS, et al. Tackling barriers to scale up human papillomavirus vaccination in China: progress and the way forward. Infect Dis Poverty. 2023;12(1):86. https://doi.org/10.1186/s40249-023-01136-6
5 Tao X, Austin RM, Kong L, Sun Q, Lv Q, Xu H, et al. Nationwide survey of cervical cytology laboratory practices in China. J Am Soc Cytopathol. 2019;8(5):250–7. https://doi.org/10.1016/j.jasc.2019.02.002
6 Satturwar: American Society of Cytopathology's Cytopathology Workforce Survey in the United States J Am Soc Cytopathology:1–13. Article in Press.xx.
7 Rockson L. Cytotechnology schools: future outlook. Diagn Cytopathol. 2024;52(8):407–12. https://doi.org/10.1002/dc.25281
8 Balachandran I, Friedlander M. Cytology workforce study: a report of current practices and trends in New York State. Am J Clin Pathol. 2011;136(1):108–18. https://doi.org/10.1309/AJCPUDFRALJY32XM
9 Lew M, Wilbur DC, Pantanowitz L. Computational cytology: lessons learned from pap test computer-assisted screening. Acta Cytol. 2021;65(4):286–300. https://doi.org/10.1159/000508629
10 Pantanowitz L, Harrington S. Experience reviewing digital pap tests using a gallery of images. J Pathol Inform. 2021;12:7. https://doi.org/10.4103/jpi.jpi_96_20
11 Cantley RL, Jing X, Smola B, Hao W, Harrington S, Pantanowitz L. Validation of AI-assisted ThinPrep® pap test screening using the Genius™ digital diagnostics system. J Pathol Inform. 2024;15:100391. https://doi.org/10.1016/j.jpi.2024.100391
12 Pulkkinen J, Huhtala H, Krogerus LA, Hollmén S, Laurila M, Kholová I. Endocervical cytology: inter- and intra-observer variability in conventional pap smears. Acta Cytol. 2022;66(3):206–15. https://doi.org/10.1159/000522212
13 Hwang H, Follen M, Guillaud M, Scheurer M, MacAulay C, MacAulay C, et al. Cervical cytology reproducibility and associated clinical and demographic factors. Diagn Cytopathol. 2020;48(1):35–42. https://doi.org/10.1002/dc.24325
14 Lowe A, Chlipala E, Elin J, Kawano Y, Long RE, Tillman D, et al. Validation of digital pathology in a healthcare environment. Digital Pathol Assoc. 2011. DPA-Healthcare-White-Paper: Second Draft-v2 (accessed 13 March 2025).
15 Janiesch C, Zschech P, Heinrich K. Machine learning and deep learning. Electron Markets. 2021;31(3):685–95. https://doi.org/10.1007/s12525-021-00475-2
16 Muhtasim, Rahman M, Khan J, Mostafizur Rahman ASM, Haque R, Ali S. A Critical Analysis of Machine Learning and Deep Learning Methods for Cervical Cancer Screening. J Comput Commun. 2023;11:64–85. https://doi.org/10.4236/jcc.2023.1112005
17 Rahman M, Khan J, Rahman ASMM, Haque R, Ali MS. A critical analysis of machine learning and deep learning methods for cervical cancer screening. J Comput Commun. 2023;11(12):64–85. https://doi.org/10.4236/jcc.2023.1112005
18 Wu T, Lucas E, Zhao F, Basu P, Qiao Y. Artificial intelligence strengthens cervical cancer screening - present and future. Cancer Biol Med. 2024;21(10):864–79. https://doi.org/10.20892/j.issn.2095-3941.2024.0198
19 Chantziantoniou N. BestCyte® Cell Sorter Imaging System: primary and adjudicative whole slide image rescreening review times of 500 ThinPrep Pap test thin-layers – an intra-observer, time-surrogate analysis of diagnostic confidence potentialities. J Pathol Inform. 2022;13:100095. https://doi.org/10.1016/j.jpi.2022.100095
20 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. https://doi.org/10.1016/j.jasc.2023.11.006
21 FaD PreAdministration. Premarket Approval of Neuromedical sciences, incorporated's PAPNET testing system. 1995. https://www.accessdata.fda.gov/cdrh_docs/pdf/p940029.pdf (accessed February 18, 2025).
22 Jenny J, Isenegger I, Boon ME, Husain OA. Consistency of a double PAPNET scan of cervical smears. Acta Cytol. 1997;41(1):82–7. https://doi.org/10.1159/000332310
23 FaD Administration. AUTOPAP(R) 300 QC AUTOMATIC PAP SCREENER/QC SYSTEM; FDA premarket approval P950009. https://fda.report/PMA/P950009 (accessed February 18, 2025).
24 Chang AR, Lin WF, Chang A, Chong KS. Can technology expedite the cervical cancer screening process? A Hong Kong experience using the AutoPap primary screening system with location-guided screening capability. Am J Clin Pathol. 2002;117(3):437–43. https://doi.org/10.1309/JC8E-4QNU-U60N-X612
25 Confortini M, Bonardi L, Bulgaresi P, Cariaggi MP, Cecchini S, Ciatto S, et al. A feasibility study of the use of the AutoPap screening system as a primary screening and location-guided rescreening device. Cancer. 2003;99(3):129–34. https://doi.org/10.1002/cncr.11077
26 Wertlake P. Results of AutoPap system-assisted and manual cytologic screening. A comparison. J Reprod Med. 1999;44(1):11–7.
27 Colgan TJ, Bon N, Clipsham S, Gardiner G, Sumner J, Walley V, et al. A validation study of the FocalPoint GS imaging system for gynecologic cytology screening. Cancer Cytopathol. 2013;121(4):189–96. https://doi.org/10.1002/cncy.21271
28 Wilbur DC, Black-Schaffer WS, Luff RD, Abraham KP, Kemper C, Molina JT, et al. The Becton Dickinson FocalPoint GS Imaging System: clinical trials demonstrate significantly improved sensitivity for the detection of important cervical lesions. Am J Clin Pathol. 2009;132(5):767–75. https://doi.org/10.1309/AJCP8VE7AWBZCVQT
29 Hologic and Cytyc complete merger. https://investors.hologic.com/press-releases/press-release-details/2007/Hologic-and-Cytyc-Complete-Merger/default.aspx (accessed February 18, 2025).

30 Administration FaD. Thinprep Imaging System; premarket approval (PMA). https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfpma/pma.cfm?ID=P020002S003 (accessed February 18, 2025).
31 Biscotti CV, Dawson AE, Dziura B, Galup L, Darragh T, Rahemtulla A, et al. Assisted primary screening using the automated ThinPrep Imaging System. Am J Clin Pathol. 2005;123(2):281–7. https://doi.org/10.1309/agb1mj9h5n43megx
32 Roberts JM, Thurloe JK, Bowditch RC, Hyne SG, Greenberg M, Clarke JM, et al. A three-armed trial of the ThinPrep imaging system. Diagn Cytopathol. 2007;35(2):96–102. https://doi.org/10.1002/dc.20600
33 The pathologist: digital pathology adoption trends: Europe and Asia; 2024 (accessed February 18, 2025).
34 Crowell EF, Bazin C, Saunier F, Brixtel R, Caillot Y, Lesner B, et al. CytoProcessorTM: a new cervical cancer screening system for remote diagnosis. Acta Cytol. 2019;63(3):215–23. https://doi.org/10.1159/000497111
35 Crowell EF, Bazin C, Thurotte V, Elie H, Jitaru L, Olivier G, et al. Adaptation of CytoProcessor for cervical cancer screening of challenging slides. Diagn Cytopathol. 2019;47(9):890–7. https://doi.org/10.1002/dc.24213
36 Chantziantoniou N. BestCyte® primary screening of 500 ThinPrep Pap Test thin-layers: 3 Cytologists' Interobserver diagnostic concordance with predicate manual microscopy relative to Truth Reference diagnoses defining NILM, ASCUS+, LSIL+, and ASCH+ thresholds for specificity, sensitivity, and equivalency grading. J Pathol Inform. 2023;14:100182. https://doi.org/10.1016/j.jpi.2022.100182
37 Delga A, Goffin F, Kridelka F, Marée R, Lambert C, Delvenne P. Evaluation of CellSolutions BestPrep® automated thin-layer liquid-based cytology Papanicolaou slide preparation and BestCyte® cell sorter imaging system. Acta Cytol. 2014;58(5):469–77. https://doi.org/10.1159/000367837
38 Ikenberg H, Lieder S, Ahr A, Wilhelm M, Schön C, Xhaja A. Comparison of the hologic genius digital diagnostics system with the ThinPrep imaging system-A retrospective assessment. Cancer Cytopathol. 2023;131(7):424–32. https://doi.org/10.1002/cncy.22695
39 Bao H, Bi H, Zhang X, Zhao Y, Dong Y, Luo X, et al. Artificial intelligence-assisted cytology for detection of cervical intraepithelial neoplasia or invasive cancer: a multicenter, clinical-based, observational study. Gynecol Oncol. 2020;159(1):171–8. https://doi.org/10.1016/j.ygyno.2020.07.099
40 Bao H, Sun X, Zhang Y, Pang B, Li H, Zhou L, et al. The artificial intelligence-assisted cytology diagnostic system in large-scale cervical cancer screening: a population-based cohort study of 0.7 million women. Cancer Med. 2020;9(18):6896–906. https://doi.org/10.1002/cam4.3296
41 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. https://doi.org/10.1016/j.modpat.2024.100486
42 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. https://doi.org/10.1016/j.modpat.2023.100186
43 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. https://doi.org/10.1038/s41467-024-48705-3
44 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. https://doi.org/10.1186/s12885-024-12532-y
45 Zhu X, Li X, Ong K, Zhang W, Li W, Li L, et al. Hybrid AI-assistive diagnostic model permits rapid TBS classification of cervical liquid-based thin-layer cell smears. Nat Commun. 2021;12(1):3541. https://doi.org/10.1038/s41467-021-23913-3
46 Du H, Dai W, Zhou Q, Li C, Li SC, Wang C, et al. AI-assisted system improves the work efficiency of cytologists via excluding cytology-negative slides and accelerating the slide interpretation. Front Oncol. 2023;13:1290112. https://doi.org/10.3389/fonc.2023.1290112
47 Zeng X, Starr D, Li J, Bi X, Wang C, Bai X, et al. AICyte-alone capabilities as an independent screener for triaging cervical cytology using a 50% negative cutoff value. Cancer Cytopathol. 2024;132(11):723–30. https://doi.org/10.1002/cncy.22896
48 Kim D, Sundling KE, Virk R, Thrall MJ, Alperstein S, Bui MM, et al. Digital cytology part 2: artificial intelligence in cytology: a concept paper with review and recommendations from the American Society of Cytopathology Digital Cytology Task Force. J Am Soc Cytopathol. 2024;13(2):97–110. https://doi.org/10.1016/j.jasc.2023.11.005
49 Pantanowitz L, Bui MM. Computer-assisted pap test screening. In: Bui MM, Pantanowitz L, editors. Modern techniques in cytopathology. Monogr Clin Cytol. Basel, Switzerland: Karger; 2020. p. 67–74.
50 Li Z, Quick M, Pantanowitz L. Whole-slide imaging in cytopathology. In: Bui MM, Pantanowitz L, editors. Modern techniques in cytopathology. Monogr Clin Cytol. Basel, Switzerland: Karger; 2020. p. 84–90.
51 Hologic genius digital diagnostics system with the genius cervical AI algorithm, instructions for use. Available from: https://www.hologic.com/package-inserts/diagnostic-products/genius-digital-diagnostics-system (accessed January 15, 2025).
52 Nayar R, Wilbur DC. The Bethesda system for reporting cervical cytology: a historical perspective. Acta Cytol. 2017;61(4–5):359–72. https://doi.org/10.1159/000477556
53 College of American Pathologists. Cytopathology checklist. CAP accreditation program. 2023.
54 Harinath L, Elishaev E, Ye Y, Matsko J, Colaizzi A, Wharton S, et al. Analysis of the sensitivity of high-grade squamous intraepithelial lesion Pap diagnosis and interobserver variability with the hologic genius digital diagnostics system. Cancer Cytopathol. 2025;133(1):e22918. https://doi.org/10.1002/cncy.22918
55 Harinath L, Elishaev E, Ye Y, Matsko J, Colaizzi A, Wharton S, et al. Diagnostic performance of the hologic genius digital diagnostics system for low-grade squamous intraepithelial lesion (LSIL) ThinPrep papanicolaou tests. J Am Soc Cytopathol. 2025;14(3):199–207. https://doi.org/10.1016/j.jasc.2025.01.003
56 Hammond MEH, Stehlik J, Drakos SG, Kfoury AG. Bias in medicine: lessons learned and mitigation strategies. JACC Basic Transl Sci. 2021;6(1):78–85. https://doi.org/10.1016/j.jacbts.2020.07.012
57 Berner ES, Graber ML. Overconfidence as a cause of diagnostic error in medicine. Am J Med. 2008;121(5 Suppl):S2–23. https://doi.org/10.1016/j.amjmed.2008.01.001
58 Larson AR, Cibas ES, Granter SR, Laga AC. Uncovering bias in the cytologic evaluation of cervical squamous lesions. Am J Clin Pathol. 2015;143(1):143–8. https://doi.org/10.1309/AJCP0HFEY0WFBNAP
59 Evans KK, Tambouret RH, Evered A, Wilbur DC, Wolfe JM. Prevalence of abnormalities influences cytologists' error rates in screening for cervical cancer. Arch Pathol Lab Med. 2011;135(12):1557–60. https://doi.org/10.5858/arpa.2010-0739-OA
60 Doxtader EE, Brainard JA, Underwood D, Chute DJ. Knowledge of the HPV status biases cytotechnologists' interpretation of Pap tests originally diagnosed as negative for intraepithelial lesion or malignancy. Cancer Cytopathol. 2017;125(1):60–9. https://doi.org/10.1002/cncy.21783
61 Bolger N, Heffron C, Regan I, Sweeney M, Kinsella S, McKeown M, et al. Implementation and evaluation of a new automated interactive image analysis system. Acta Cytol. 2006;50(5):483–91. https://doi.org/10.1159/000326001
62 Acanfora G, Carillo AM, Dello Iacovo F, Salatiello M, Pisapia P, Bellevicine C, et al. Interobserver variability in cytopathology: how much do we agree? Cytopathology. 2024;35(4):444–53. https://doi.org/10.1111/cyt.13378

63 Stoler MH, Schiffman M; Atypical Squamous Cells of Undetermined Significance-Low-grade Squamous Intraepithelial Lesion Triage Study ALTS Group. Interobserver reproducibility of cervical cytologic and histologic interpretations: realistic estimates from the ASCUS-LSIL Triage Study. JAMA. 2001;285(11):1500–5. https://doi.org/10.1001/jama.285.11.1500

64 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):102–31. https://doi.org/10.1097/LGT.0000000000000525

65 Abels E, Pantanowitz L. Current state of the regulatory trajectory for whole slide imaging devices in the USA. J Pathol Inform. 2017;8:23. https://doi.org/10.4103/jpi.jpi_11_17

66 Sullivan HR, Schweikart SJ. Are current tort liability doctrines adequate for addressing injury caused by AI? AMA J Ethics. 2019;21(2):E160–6. https://doi.org/10.1001/amajethics.2019.160