Research Article ISSN 2833-0390

International Journal of Research in Oncology

Accuracy and Cost-Effectiveness of AI-Assisted Cervical Cancer Screening: A Scoping Review and Comparative Evaluation of Pap Testing, Colposcopy, and Mobile Applications

Sanzida Shams[2], Raquel Rodrigues Peres[1], Poonam Poonam[1], Ruby Chalana[1], Karen Lobo[1], Agafya Krivova[1], Rifat Farzan Nipun[1], Tamjida Hanfi[4], Chitra Ghosh[3], Atmodipta Monisha Hira[5] and Shafi Bhuiyan[1]*

[1]*University of Toronto, Ontario, Canada.*

[2]*McMaster University, Hamilton, ON, Canada.*

[3]*University of Memphis, TN, USA.*

[4]*Mahidol University, Thailand.*

[5]*Faridpur Medical College Diabetic Association, Bangladesh.*

***Correspondence:**
Prof. Dr. Shafi Bhuiyan, PhD, MBBS, MPH, MBA, University of Toronto, Canada. E-mail: shafi.bhuiyan@utoronto.ca.

Received: 17 Dec 2025; **Accepted:** 31 Jan 2026; **Published:** 10 Feb 2026

Citation: Shafi Bhuiyan, Sanzida Shams, Raquel Rodrigues Peres, et al. Accuracy and Cost-Effectiveness of AI-Assisted Cervical Cancer Screening: A Scoping Review and Comparative Evaluation of Pap Testing, Colposcopy, and Mobile Applications Int J Res Oncol. 2026; 5(1): 1-11.

ABSTRACT

Background: *Cervical cancer remains a leading cause of death among women and is largely preventable; the burden is most significant in low- and middle-income countries despite the availability of effective HPV vaccination and screening. Artificial intelligence (AI) could strengthen screening programs by improving accuracy, throughput, and access across cytology, colposcopy, and mobile app–based platforms.*

Methods: *Following PRISMA-ScR, we searched PubMed, EMBASE, MEDLINE, and CINAHL (2014–2025) for peer-reviewed studies applying Artificial Intelligence (AI)/ Machine Learning (ML)/ Deep Learning (DL) to cervical cancer screening or diagnosis using Pap/liquid-based cytology (LBC), colposcopy, or mobile tools, and reporting diagnostic or economic outcomes. Two reviewers independently screened, extracted, and adjudicated data.*

Results: *Of 724 records, 19 studies met inclusion criteria. AI-assisted Pap/LBC cytology for CIN2+ showed sensitivity ~86–90% and specificity ~51–95%; for CIN3+, sensitivity ~87–92% with specificity ~51–61%, often matching or exceeding conventional workflows while increasing slide throughput. Colposcopy AI achieved sensitivity/specificity trade-offs depending on thresholds (e.g., high-grade-or-worse: ~72%/94%; more sensitive setting: ~91%/52%), supporting targeted biopsy decisions and reducing operator variability. Mobile/point-of-care tools reported strong performance in validation settings (e.g., smartphone VIA ~99% sensitivity/~97% specificity; AI-supported HPV assays ~100%/~100%), with minimal hardware requirements and potential for task-shifting. Across modalities, AI commonly reduced time-to-diagnosis and expert workload. Eight studies assessed economics; AI-LBC at 5-year intervals was cost-effective (example ICER ≈ US$8,790/Quality-Adjusted Life Year(QALY)), and decision-support systems reduced unnecessary colposcopies, suggesting favorable value in resource-constrained settings.*

Conclusions: *AI-enabled screening can deliver clinically meaningful accuracy with operational gains that are likely to improve coverage and equity, particularly where cytology expertise and colposcopy capacity are limited. The greatest value appears where moderate-to-high diagnostic performance coincides with low deployment costs (e.g., AI-assisted cytology and mobile platforms). Evidence is promising but heterogeneous; real-world, prospective studies with standardized endpoints and full economic evaluations are still needed to confirm effectiveness, scalability, and affordability within national screening programs aligned to World Health Organization (WHO) elimination targets.*

Keywords
Cervical Cancer, AI-Assisted Screening, Cost-Effectiveness, Comparative Evaluation, Pap Testing & Colposcopy, Mobile Applications.

Introduction

Cervical cancer is the fourth most common cancer among women globally and remains a leading cause of cancer-related mortality [1]. Human papillomavirus (HPV) causes cervical cancer, but it is largely preventable through human papillomavirus (HPV) vaccination, timely screening, and effective treatment of precancerous lesions. In 2022, the World Health Organization (WHO) reported over 350,000 cervical cancer deaths, with projections indicating an 80.7% increase by 2050, reaching nearly 630,000 annual deaths if current trends persist [2,3].

The burden of cervical cancer is disproportionately higher among older women and in low- and middle-income countries (LMICs). According to GLOBOCAN 2022, women aged 40 years and above account for the majority of cases, with incidence and mortality rates 10-fold and 20-fold higher, respectively, than in younger women. About 80% of cases occur in Asia and Africa, with incidence in low-Human Development Index (HDI) countries being twice as high as in very high-HDI countries. Mortality varies widely from 1.1 to 64.3 per 100,000 women, highlighting persistent structural health disparities.

Scaling up effective screening is essential to reducing these disparities. Vaccination among adults aged 30–45 has shown limited cost-effectiveness, reinforcing the need to vaccinate girls before puberty [4]. The cervical cancer survival rate decreases with age; the risk of death among women aged ≥50 years is 1.82 times higher than in younger age groups [5]. Modeling indicates that screening women aged 30–50 could reduce incidence by 51–56% and mortality by 63–66% [6]. Early-stage detection is critical; stage I diagnosis reduces the risk of mortality by a factor of 6.77 compared to stage IV [5]. However, in LMICs, treatment cost is up to USD 70,968 per patient, combined with limited insurance coverage, often resulting in catastrophic out-of-pocket expenses [7].

The WHO's global elimination strategy aims to reduce cervical cancer incidence to fewer than 4 per 100,000 women via the "90–70–90 by 2030" targets. The aims are the following: 90% of girls fully vaccinated by age 15, 70% of women screened at least twice by ages 35 and 45, and 90% of women with disease receiving appropriate care [2]. Despite global commitment, progress remains uneven. In 2023, HPV vaccine coverage among girls was only 27% globally, and screening coverage among women aged 30–49 exceeded 10% in just 149 countries [8,9]. In resource-constrained settings, there is an urgent need for strategies that can enhance diagnostic reach while minimizing cost and clinical burden.

Conventional screening methods such as the Papanicolaou (Pap) test and its liquid-based cytology (LBC) counterpart, although effective, suffer from variability in sensitivity, particularly for glandular lesions, and require skilled personnel. Abnormal results often necessitate confirmatory colposcopy and biopsy. Colposcopy itself is limited by equipment costs and operator dependence, although portable and smartphone-based digital colposcopes have improved accessibility and documentation. High-risk HPV DNA testing, usually conducted via Polymerase Chain Reaction (PCR), offers superior negative predictive value over cytology, and a negative result is more predictive of low cancer risk than a negative Pap test [10].

Pap smears remain a mainstay of cervical cancer screening in most countries due to their low cost and relatively high specificity (65–97%) [11,12]. However, their sensitivity is limited, particularly for glandular lesions, and false-negative rates have been reported as high as 25–50% [13]. Causes include sampling errors, interpretive variability, and the limitations of detecting early-stage or glandular pathology [14]. The discomfort associated with Pap testing may also deter participation, contributing to suboptimal coverage [15].

Artificial intelligence (AI) offers an opportunity to enhance cervical cancer screening, particularly in low-resource environments. Machine learning algorithms allow rapid and reproducible interpretation of complex datasets, including cytological images, colposcopy visuals, and radiological scans. AI-assisted cytology has demonstrated high accuracy (sensitivity: 0.991; specificity: 0.996; accuracy: 0.995) [16]. Convolutional neural networks (CNNs) are especially suited for image-based tasks, while support vector machines and ensemble models enable risk stratification using structured clinical data. Commercial AI systems such as BestCyte, Cytoprocessor, and Genius Digital Diagnostics integrate imaging, remote consultation, and algorithmic triage. Though still emerging, deep learning tools also show promise in histopathological analysis for identifying dysplasia and invasive lesions [17,18].

To address these challenges, there is a need for diagnostic tools that offer high accuracy, throughput, and sensitivity at a cost that is feasible for resource-constrained health systems. This scoping review analyzed 19 studies published between 2014 and 2025 to evaluate the accuracy, feasibility, and cost-effectiveness of AI-assisted diagnostic tools, particularly those leveraging Pap smear analysis, in cervical cancer screening.

Although AI has demonstrated diagnostic accuracy and operational efficiency, comparative data on its cost-effectiveness compared to traditional methods (e.g., Pap, HPV testing, colposcopy) remain limited. Few studies have explored AI's real-world utility across different screening platforms and healthcare contexts.

This scoping review aims to evaluate the diagnostic performance and cost-effectiveness of AI-assisted cervical cancer screening across Pap testing, colposcopy, and mobile app–based platforms. We compare these AI-integrated modalities against conventional approaches in terms of sensitivity, specificity, area under the curve (AUC), time to diagnosis, resource utilization, and economic outcomes such as incremental cost-effectiveness ratios (ICERs).

Emphasis is placed on feasibility and value in resource-constrained environments. By identifying where AI offers the greatest clinical and economic benefit, this review seeks to inform implementation strategies and policies aligned with global cervical cancer elimination goals.

Methods

This scoping review was conducted in accordance with the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews) guidelines. We included peer-reviewed, English-language articles published between 2014 and 2025 that applied artificial intelligence (AI)—including machine learning (ML) and deep learning (DL) to the screening or diagnosis of cervical cancer using Pap smears, colposcopy, or mobile applications. Eligible studies utilized human data, employed quantitative, qualitative, or mixed methods, and reported outcomes related to diagnostic accuracy, cost-effectiveness, or socio-economic impact.

Studies were excluded if they: (1) did not incorporate AI, (2) focused solely on clinical outcomes without economic analysis, (3) lacked methodological detail on AI implementation, (4) involved non-cervical diseases, (5) were not original research (e.g., reviews, editorials, commentaries), or (6) were animal studies.

Information Sources and Search Strategy

A comprehensive search was conducted in four electronic databases: PubMed, Embase, MEDLINE and CINAHL. The search strategy combined Medical Subject Headings (MeSH), Emtree terms, and free-text keywords such as "artificial intelligence", "machine learning", "deep learning", "cervical cancer", "screening" and "diagnosis". Boolean operators were used to refine results, and the strategies were customized for each database. The search covered studies published between 2014 and 2025, without geographical restrictions. All retrieved citations were exported to Covidence (Veritas Health Innovation, Melbourne, Australia) for deduplication and screening.

Study Selection

The selection process involved two stages: (1) title and abstract screening, and (2) full-text review. In both phases, two independent reviewers screened all records in Covidence. The reviewers resolved disagreements by discussion or by consulting a third reviewer. Reviewers recorded reasons for exclusion at the full-text stage. The study selection process is illustrated in a PRISMA-ScR flow diagram (Figure 1).

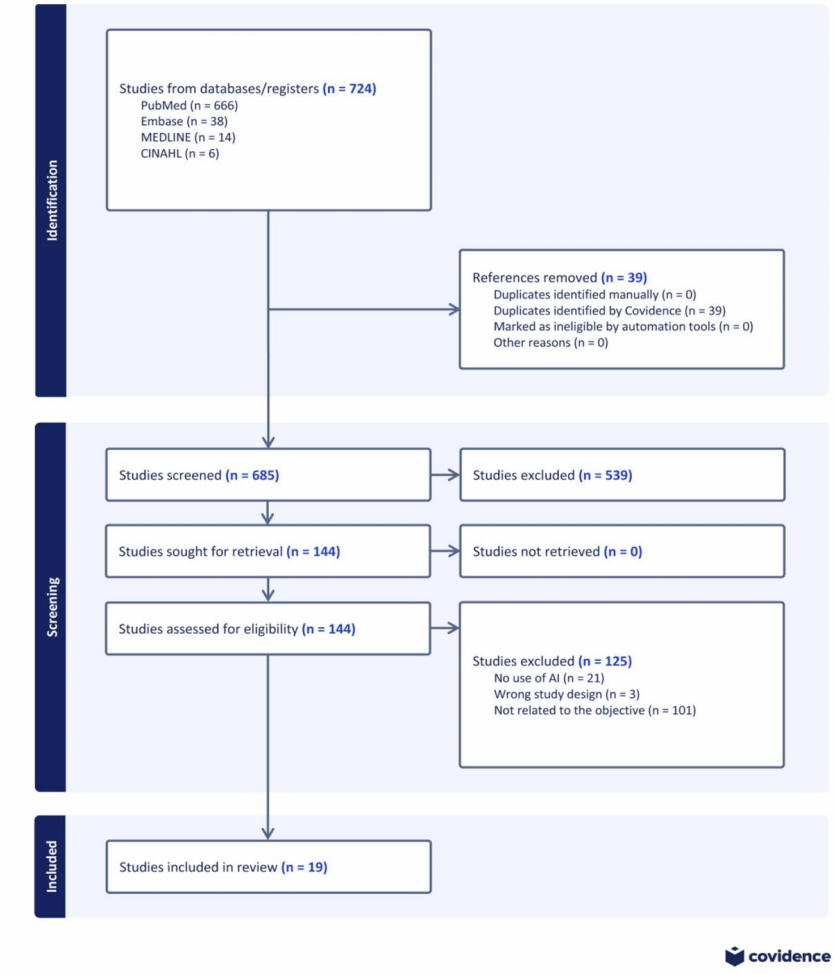

Figure 1: PRISMA-ScR flow diagram of study selection for the scoping review, including identification, screening, eligibility, and inclusion.

Data Extraction

Data were extracted independently by two reviewers using Covidence, with arbitration by a third reviewer when needed. A total of 724 studies were retrieved (PubMed 666; Embase 38; MEDLINE 14; and CINAHL 6). After removing duplicates, 685 remained for screening. Following title/abstract screening, 144 full texts were assessed; 125 were excluded, leaving 19 studies for inclusion.

Key data items from each study included: article title, lead author, country of study, study aim, design, AI technologies or digital tools used, data collection period, funding source, conflict of interest disclosures, participant characteristics (e.g., inclusion/exclusion criteria, recruitment method, total sample size, demographics), interventions and comparators, outcome measures, statistical analysis, diagnostic accuracy, accessibility across populations, economic impact, and reported strengths or limitations.

Synthesis of Results

The 19 included studies primarily employed AI-based techniques such as deep Learning(DL)/Machine Learning(ML), and digital image analysis for cervical cancer screening and diagnosis. Screening tools included Pap smear cytology, visual inspection with acetic acid (VIA), HPV DNA testing (via qPCR and genotyping), p16/Ki67 immunostaining, and nucleic acid sequence-based amplification (NASBA).

AI methods encompassed cytology automation with liquid-based cytology (LBC), probabilistic neural networks (PNN), convolutional neural networks (CNN), artificial neural networks (ANN), and multilayer perceptron (MLP). Algorithms, such as Support Vector Machine (SVM), CNN-4, ResNet-50, ResNet-101, ResNet-152, AlexNet, InceptionV3, and DP-Net + RRS, were evaluated against traditional cytological analysis by experienced pathologists. Several studies also utilized smartphone-enabled remote image analysis, along with image processing and data augmentation techniques.

Results

Literature Search and Study Selection

During the literature search, 724 records were identified. After removing duplication, title/abstract screening and full-text review, 19 studies were included for further analysis. (Figure 1).

Study Characteristics and Summary

Table 1 summarizes the core characteristics of the included studies, covering intervention types, AI models, diagnostic performance, and cost-effectiveness outcomes. The interventions ranged from mobile-based screening platforms and deep learning classifiers to AI-enabled cytology and PCR automation systems. Studies varied in scale, with some large population-based screenings (e.g., Bao, 2020) [19] and others focusing on device validation or algorithm comparison.

Participant Characteristics and Recruitment Methods

Participants represented diverse populations, typically women aged 25–65 years undergoing routine cervical cancer screening. Many studies focused on high-risk groups, including immunocompromised individuals (e.g., HIV-positive women). Common inclusion criteria included abnormal Pap smears or positive high-risk HPV tests, while exclusion criteria often involved pregnancy, prior cervical surgery, comorbid infections, or poor-quality cytological samples.

Recruitment strategies varied: clinic-based referrals, population registry invitations, hospital-based prospective recruitment, and biobank utilization were all reported. This diversity enhanced the generalizability and validity of the study findings.

Types of Interventions and AI Models

The majority of studies employed convolutional neural networks (CNNs), including variants like ResNet-50, InceptionV3, and hybrid models with attention mechanisms or pooling strategies [20-22]. AI-assisted cytology systems, such as Landing CytoScanner and CytoProcessor™, demonstrated high throughput and scalability for large-screening settings.

Classical machine learning approaches, including support vector machines (SVM), probabilistic neural networks (PNN), and multilayer perceptrons (MLP), were used in clinical decision support systems [23]. Mobile and cloud-based solutions [24,25] targeted low-resource environments with promising sensitivity and usability outcomes.

Intervention Characteristics

Interventions varied by complexity, clinical setting, and resource requirements. For example, Kudva [24] demonstrated that mobile-based screening achieved high sensitivity (99.05%) and specificity (97.16%) with minimal hardware costs. While performance exceeded 99% with ResNet ensembles, these models were highly compute-intensive [26].

AI platforms also supported automation of HPV genotyping [25], cytological classification [27], and colposcopic biopsy guidance [28], offering integrated solutions to streamline workflows and reduce diagnostic delays.

Diagnostic Accuracy Outcomes

To avoid cross-comparisons across heterogeneous tasks, we summarize performance by modality, clinical endpoint, and unit of analysis.

1) Pap/LBC cytology (slide- or patient-level; histology reference where reported)

Endpoint: CIN2+ detection (triage/screening)

Sensitivity ≈86–90%; Specificity ≈51–95% across studies using AI-assisted liquid-based cytology (e.g., AI-LBC vs cytologists vs HPV16/18 typing) and population-based AI cytology systems. Some examples are: AI-LBC CIN2+ Se 86.5%, Sp 51.3% (HPV-positive triage); large-scale AI cytology CIN2+ Se 90.1%, Sp 94.8%.

Endpoint: CIN3+ detection

Sensitivity ≈87–92%; Specificity ≈51–61% in automated

dual-stain or AI-LBC workflows, depending on thresholds and cohorts. Some examples are: AI-LBC CIN3+ Se 92.3%, Sp ≈50.8%; automated DS strategies with higher sensitivity at lower cell-count cutoffs and moderate specificity.

Endpoint: Cytology category thresholds (ASC-US+/LSIL+/HSIL+)

Compared with conventional imaging systems, AI platforms reported higher or comparable sensitivity with similar specificity for thresholds such as ASC-US+ and HSIL+, depending on operating points and adjudication methods.

2) Colposcopy (visit- or patient-level; histology reference where reported)

Endpoint: HSIL+ / "high-grade-or-worse" triage

Sensitivity ≈72%; Specificity ≈94% when classifying high-grade or-worse at stricter thresholds.

Endpoint: "low-grade-or-worse" (more sensitive operating point):

Sensitivity ≈91%; Specificity ≈52%.

Note: Threshold choice trades sensitivity vs specificity and affects downstream colposcopy volume.

3) Mobile/POC tools (device- or image-level; VIA/HPV endpoints)

Smartphone/Android VIA image classification (VIA+ vs VIA−):

Sensitivity ≈99.1%; Specificity ≈97.2% in low-resource, device constrained settings (endpoint based on VIA, not histology).

POC HPV microholography + DL (assay-level)

Sensitivity ≈100%; Specificity ≈100% for HPV16/18 DNA detection against a laboratory comparator in validation datasets.

4) Algorithmic image classification (cell/patch-level; non-clinical endpoints)

Tasks such as cell-level Bethesda class or patch-level abnormal vs normal reported accuracy from ~67% to ~100%, including very high values (e.g., hybrid CNNs and ResNet ensembles). Note: These metrics are not directly comparable to patient-level CIN2+/CIN3+ endpoints and are therefore reported separately.

Cost-Effectiveness and Economic Impact

Cost-effectiveness was explicitly evaluated in eight studies. Metrics such as the Incremental Cost-Effectiveness Ratio (ICER) and Quality-Adjusted Life Years (QALY) gained were used to quantify economic impact. For instance:

- Shen ([9] estimated that AI-LBC screening every five years resulted in an ICER of $8,790 per QALY gained, which was cost-effective under standard thresholds.
- Kudva [24] and Pereira [25] highlighted low-cost deployment, with reduced training and processing time, making them suitable for scale-up in LMICs.
- Clinical decision support systems [23] demonstrated cost savings by reducing unnecessary colposcopies and early identification of high-risk cases.

Interventions combining diagnostic accuracy with low resource requirements were consistently deemed the most cost-effective, particularly in under-resourced settings.

Discussion

Achieving the WHO's target of screening 70% of women by age 35 for cervical cancer by 2030 appears implausible given current global screening rates, which remain below 10% [9]. While high-Human Development Index countries have experienced declining cervical cancer incidence and mortality through integrated screen–triage–treat approaches, low-resource settings continue to struggle with inadequate coverage. In such settings, conventional methods alone are insufficient due to the high burden of disease, logistical challenges, and limited healthcare infrastructure. Moreover, the implementation of combined screening strategies is often constrained by financial considerations, particularly when incremental costs are evaluated in relation to national Gross Domestic Product (GDP). These costs vary depending on screening frequency (e.g., 3-, 5-, or 10-year intervals) and the lifetime risk of cervical cancer among women in each country.

The findings demonstrate substantial progress in the integration of AI and machine learning into cervical cancer diagnostics. Most studies utilized AI algorithms, primarily convolutional neural networks (CNNs) to analyze cytological images for cervical cell abnormalities, employing classification systems such as the Bethesda (TBS) criteria. The TBS standardizes cervicovaginal cytology reporting clinically relevant categories: specimen type, adequacy, general category, interpretation, and adjunctive testing [30]. Reported diagnostic accuracies ranged from 66.8% to 99.98%, consistent with prior research [31], underscoring the reliability of AI in distinguishing between low-grade lesions (e.g., CIN 1) and high-grade precancerous or cancerous lesions (e.g., CIN 2/3+), a key clinical priority in cervical screening [32].

For example, Mishra et al. [20] reported a hybrid CNN approach that achieved high classification accuracy across various categories: carcinoma in situ (98.54%), moderate dysplasia (98.91%), and severe dysplasia (97.45%), with similarly strong performance in normal cells. These results highlight the capacity of CNNs to detect subtle morphological distinctions, enabling precise risk stratification and reducing unnecessary clinical interventions. Similarly, Crowell et al. [33] demonstrated that automated cytoprocessors can substantially increase slide throughput without adding workforce demands, facilitating remote and large-scale screening.

AI-assisted cytology systems, especially those employing deep learning offers significantly improved performance. For instance, studies by Pathania and Ke et al. report sensitivity and specificity nearing 100%. AI has also shown marked improvements in colposcopy interpretation. CNNs, SVMs, and hybrid models have demonstrated strong performance in image classification, reducing misdiagnosis and supporting more targeted biopsies. Mathivanan [34] found that ResNet152 achieved diagnostic accuracies above 98%, validating the use of deep learning fusion models in clinical

workflows.

From an economic and implementation perspective, AI-based tools offer notable advantages. Studies such as Kudva [24] and Pathania [25] reported successful deployment of low-cost, portable, mobile-based screening platforms, making early detection feasible in low-resource settings. Other innovations like FastFinder [25] and AI-assisted cytology platforms [28,33] significantly reduced workload and diagnosis time while improving consistency. Automation minimizes inter-observer variability and supports remote diagnostics, improving equity in healthcare delivery. Economic evaluations, such as that by Shen [29], demonstrated favorable incremental cost-effectiveness ratios (ICERs), confirming that AI-assisted screening can yield better health outcomes without proportionally increasing system costs.

The methodological rigor of this review was enhanced by a comprehensive search strategy with no geographic limitations and the use of dual independent reviewers for data extraction. Most included studies demonstrated sound internal validity through the use of randomization, blinding, and controlled comparisons. However, several limitations must be acknowledged. First, the small number of studies assessing socio-economic outcomes limits the generalizability of cost-effectiveness findings. Second, many studies relied on archived images from registries rather than clinical trials or real-world implementation studies. Consequently, limited evidence is available on the deployment and scalability of these tools in diverse clinical and population contexts. Lastly, significant heterogeneity exists in outcome definitions, economic endpoints, and diagnostic thresholds, making cross-study comparisons difficult.

To advance the field, future research should focus on prospective, diagnostic accuracy studies and implementation research in diverse population groups. More studies are needed to examine the real-world cost-effectiveness, feasibility, and social impact of scaling AI-assisted diagnostics within national screening programs. While current evidence supports the accuracy and feasibility of AI-based tools, more robust data are necessary to substantiate their superiority over conventional methods and justify their broad implementation in low-resource settings.

Conclusion

This scoping review shows that AI can materially strengthen cervical cancer screening across three key touchpoints like Pap/LBC cytology, colposcopy, and mobile/point-of-care platforms by delivering clinically meaningful accuracy while easing workload and speeding time to diagnosis. In cytology, AI assistance typically achieves high sensitivity for CIN2+/CIN3+ and improves throughput, offering a pragmatic path to expand screening where expert readers are scarce. In colposcopy, algorithmic support can stabilize performance across operators and sharpen biopsy targeting, provided thresholds are tuned to program goals. Despite validation primarily on surrogate endpoints, mobile/cloud tools pair strong technical performance with low hardware and training demands, making them well-suited to resource-scarce rural and remote LMIC settings.

Economically, the most compelling value emerges where moderate-to-high diagnostic accuracy aligns with low deployment and operating costs, namely AI-assisted cytology and mobile screening yielding favorable ICERs and the potential to reduce unnecessary referrals. These gains map directly onto the WHO 90–70–90 strategy by improving triage efficiency, expanding reach, and enabling task-shifting without sacrificing quality.

At the same time, evidence remains heterogeneous. Many studies rely on archival images, surrogate endpoints, or non-standardized thresholds; few report full economic evaluations or long-term outcomes. Real-world, prospective studies that use harmonized CIN2+/CIN3+ endpoints, linkage to treatment, and transparent cost/QALY reporting are essential to confirm effectiveness, equity, and affordability at scale.

Study Implications

- Prioritize AI-assisted cytology or mobile VIA/HPV workflows in resource-constrained settings, integrating them as front-end triage to preserve specialist capacity for high-risk cases.
- Adopt clear operating thresholds that reflect local goals (maximize sensitivity for first-pass screening; rebalance toward specificity in triage).
- Plan for implementation fundamentals: workforce training, quality assurance dashboards, data governance and privacy, calibration and drift monitoring, and interoperability with existing laboratory and registry systems.
- Pair deployment with economic monitoring (program costs, ICERs, avoided referrals) and equity metrics (coverage gains among underserved groups).

Research priorities

- Multi-site pragmatic trials comparing AI-integrated pathways to standard care on patient-level outcomes (CIN2+/CIN3+ detection, treatment completion, stage shift, and time-to-diagnosis).
- Country-specific cost-effectiveness analyses that include capital, maintenance, connectivity, and supervision costs, and report budget impact.
- Fairness, robustness, and safety studies (domain shift, image quality variation, HPV genotype mix), with open protocols and external validation.
- Human factors and implementation science to optimize workflow fit, task-shifting, and patient acceptability.

Overall, AI is not a replacement for proven screening strategies but a force multiplier. When embedded thoughtfully starting with AI-assisted cytology and scalable mobile solutions, AI can help countries extend coverage, and improve diagnostic yield and cost effectiveness. With rigorous prospective evidence and careful program design, AI-enabled screening can accelerate progress toward the WHO target of fewer than 4 cases per 100,000 women, narrowing persistent global disparities in cervical cancer outcomes.

Recommendations

1. **Implementation Research**
   Future large-scale, prospective studies should assess AI-assisted screening in real-world clinical and community settings to evaluate feasibility, cost-effectiveness, and patient outcomes across diverse populations.
2. **Integration into National Screening Programs**
   Health ministries should pilot AI-enabled cytology and colposcopy tools within existing screening frameworks to optimize early detection, triage, and treatment pathways.
3. **Capacity Building and Training**
   Investment in training healthcare workers to operate AI-based systems and interpret outputs is essential for sustainable deployment, especially in low-resource environments.
4. **Policy and Cost Frameworks**
   Governments and donors should establish regulatory and reimbursement policies that incentivize the adoption of validated AI technologies, ensuring affordability and equitable access.
5. **Ethical and Data Governance Standards**
   Standardized guidelines for data privacy, algorithm transparency, and bias mitigation are necessary to maintain trust and ensure ethical AI integration in women's health diagnostics.

References

1. Li Z, Liu P, Yin A, et al. Global landscape of cervical cancer incidence and mortality in 2022 and predictions to 2030: The urgent need to address inequalities in cervical cancer. International Journal of Cancer. 2025; 157: 288-297.
2. https://www.who.int/campaigns/cervical-cancer-elimination-day-of-action/2024
3. Wu J, Jin Q, Zhang Y, et al. Global burden of cervical cancer: Current estimates, temporal trend and future projections based on the GLOBOCAN 2022. Journal of the National Cancer Center. 2025; 5: 322-329.
4. Kim JJ, Simms KT, Killen J, et al. Human papillomavirus vaccination for adults aged 30 to 45 years in the United States: A cost-effectiveness analysis. PLOS Medicine. 2021; 18: 1003534.
5. Muneeha S, Chaudhary RK, Shetty VV, et al. Exploring and detecting predictors associated with survival and mortality of cervical cancer patients: A 10-year retrospective study. Beni-Suef University Journal of Basic and Applied Sciences. 2025; 14: 13.
6. Simms KT, Keane A, Nguyen DTN, et al. Benefits, harms and cost-effectiveness of cervical screening, triage and treatment strategies for women in the general population. Nat Med. 2023; 29: 3050-3058.
7. Fuady A, Setiawan D, Man I, et al. Toward a Framework to Assess the Financial and Economic Burden of Cervical Cancer in Low- and Middle-Income Countries: A Systematic Review. JCO Global Oncology. 2024; 10: 2400066.
8. Brotherton JML, Vajdic CM, Nightingale C. The socio-economic burden of cervical cancer and its implications for strategies required to achieve the WHO elimination targets. Expert Rev Pharmacoecon Outcomes Res. 2025; 25: 487-506.
9. Zhou L, Li Y, Wang H, et al. Global cervical cancer elimination: Quantifying the status, progress, and gaps. BMC Medicine. 2025; 23: 67.
10. Goldstein A, Gersh M, Skovronsky G, et al. The Future of Cervical Cancer Screening. Int J Womens Health. 2024; 16: 1715-1731.
11. Peirson L, Fitzpatrick-Lewis D, Ciliska D, et al. Screening for cervical cancer: A systematic review and meta-analysis. Systematic Reviews. 2013; 2: 35.
12. Arbyn M, Buntinx F, Ranst MV, et al. Virologic Versus Cytologic Triage of Women With Equivocal Pap Smears: A Meta-analysis of the Accuracy To Detect High-Grade Intraepithelial Neoplasia. J Natl Cancer Inst. 2004; 96: 280-293.
13. Macios A, Nowakowski A. False Negative Results in Cervical Cancer Screening-Risks, Reasons and Implications for Clinical Practice and Public Health. Diagnostics (Basel, Switzerland). 2022; 12: 1508.
14. Miller RA, Mody DR, Tams KC, et al. Glandular Lesions of the Cervix in Clinical Practice: A Cytology, Histology, and Human Papillomavirus Correlation Study From 2 Institutions. Arch Pathol Lab Med. 2015; 139: 1431-1436.
15. Rask M, Oscarsson M, Lindell G, et al. Women with abnormal Pap smear result: A qualitative study of Swedish healthcare professionals' experiences. Eur J Cancer Care. 2016; 25: 980-991.
16. Wang J, Yu Y, Tan Y, et al. Artificial intelligence enables precision diagnosis of cervical cytology grades and cervical cancer. Nature Communications. 2024; 15: 4369.
17. Dellino M, Cerbone M, d'Amati A, et al. Artificial Intelligence in Cervical Cancer Screening: Opportunities and Challenges. AI. 2024; 5: 2984-3000.
18. Wang R, Gunesli GN, Skingen VE, et al. Deep learning for predicting prognostic consensus molecular subtypes in cervical cancer from histology images. NPJ Precis Oncol. 2025; 9: 11.
19. Bao H, Sun X, Zhang Y, 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 Medicine. 2020; 9: 6896-6906.
20. Mishra L, Dasgupta R, Chowdhury YS, et al. Cervical Cancer Detection Using Hybrid Pooling-Based Convolutional Neural Network Approach. Indian Journal of Gynecologic Oncology. 2023; 21: 37.
21. Lin CC, Kuo CH, Chiang HT. CNN-Based Classification for Point Cloud Object With Bearing Angle Image. IEEE Sensors Journal. 2022; 22: 1003-1011.
22. Mahto R, Sood K. Predicting Cervical Cancer Based on Behavioral Risk Factors. International Journal of Advanced Computer Science and Applications. 2024; 15.
23. Kyrgiou M, Athanasiou A, Paraskevaidi M, et al. Adverse obstetric outcomes after local treatment for cervical

preinvasive and early invasive disease according to cone depth: Systematic review and meta-analysis. BMJ. 2016; 354: 3633.

24. Kudva V, Prasad K, Guruvare S. Andriod Device-Based Cervical Cancer Screening for Resource-Poor Settings. J Digit Imaging. 2018; 31: 646-654.
25. Pereira AR, Redzic N, Van Vooren S, et al. Development, Validation, and Implementation of an Augmented Multiwell, Multitarget Quantitative PCR for the Analysis of Human Papillomavirus Genotyping through Software Automation, Data Science, and Artificial Intelligence. J Mol Diagn. 2024; 26: 781-791.
26. Karasu Benyes Y, Welch EC, Singhal A, et al. A Comparative Analysis of Deep Learning Models for Automated Cross-Preparation Diagnosis of Multi-Cell Liquid Pap Smear Images. Diagnostics. 2022; 12: 1838.
27. Wentzensen N, Lahrmann B, Clarke MA, et al. Accuracy and Efficiency of Deep-Learning–Based Automation of Dual Stain Cytology in Cervical Cancer Screening. J Natl Cancer Inst. 2021; 113: 72-79.
28. Xue P, Tang C, Li Q, et al. Development and validation of an artificial intelligence system for grading colposcopic impressions and guiding biopsies. BMC Medicine. 2020; 18: 406.
29. Shen M, Zou Z, Bao H, et al. Cost-effectiveness of artificial intelligence-assisted liquid-based cytology testing for cervical cancer screening in China. Lancet Reg Health West Pac. 2023; 34: 100726.
30. Pangarkar MA. The Bethesda System for reporting cervical cytology. Cyto journal. 2022; 19: 28.
31. Allahqoli L, Ghiasvand MM, Mazidimoradi A, et al. Diagnostic and Management Performance of ChatGPT in Obstetrics and Gynecology. Gynecol Obstet Invest. 2023; 88: 310-313.
32. Hou D, Yang B, Li Y, et al. Utility of Scoring System for Screening and Early Warning of Cervical Cancer Based on Big Data Analysis. Front Public Health. 2022; 10: 920956.
33. Crowell EF, Bazin C, Saunier F, et al. CytoProcessorTM: A New Cervical Cancer Screening System for Remote Diagnosis. Acta Cytologica. 2019; 63: 215-223.
34. Mathivanan SK, Francis D, Srinivasan S, et al. Enhancing cervical cancer detection and robust classification through a fusion of deep learning models. Scientific Reports. 2024; 14: 10812.

Appendices

Table 1: Summarizes the main characteristics of the articles included in this scoping review. Cost-effectiveness was also evaluated for studies that mentioned this criterion, highlighting the advantages and disadvantages in the application of AI and other tools in the screening and diagnosis of cervical cancer.

Author	Title	Participants/Samples	Type of Intervention	Comparator/Control	Outcome(s)	Cost-effectiveness
Mishra 2023	Cervical Cancer Detection Using Hybrid Pooling-Based Convolutional Neural Network Approach	Two publicly available benchmark datasets: Herlev Dataset: n = 917 single-cell Pap smear images, labeled into 7 classes. IARC Colposcopy Dataset: n = 367 colposcopy images, labeled into 9 classes.	AlexNet-SVM network with hybrid pooling strategy.	Standard CNN pooling.	Classification accuracy of cervical cell/lesion types.	Reports computational time savings (test report in 2 min) with 95.45% accuracy.
Kudva 2018	Android Device-Based Cervical Cancer Screening for Resource-Poor Settings.	n = 102 cervix images obtained from VIA screening programmes at local health centres (42 VIA-positive and 60 VIA-negative) from married women aged ≥ 25.	Mobile-based automated cervical cancer screening system using an Android device.	Manual evaluation by a medical expert as the gold standard for VIA classification.	Accuracy, sensitivity, and specificity of the system in classifying VIA-positive and VIA-negative images. Performance of SVM and DT classifiers under k-fold CV and leave-one-out CV.	Low-cost Android device platform, tailored for low-resource settings.
Lin 2021	Dual-path network with synergistic grouping loss and evidence driven risk stratification for whole slide cervical image analysis.	n = 19,303 whole-slide cervical Pap smear WSIs from 4 medical centers (13,486 for training, 2,486 for validation, and 3,331 for independent testing.	Artificial intelligence–based image analysis (DP-Net + RRS)	Conventional classifiers (Random Forest, SVM, AdaBoost).	Smear-level classification into five Bethesda classes and risk stratification. Cell-level classification performance.	Achieves fast processing of large-scale WSIs with potential to reduce 80% cytologists' workload.
Pirovano 2021	Computer-aided diagnosis tool for cervical cancer screening with weakly supervised localization and detection of abnormalities using adaptable and explainable classifier.	n = 1,808 simulated Pap-smear tiles built from 917 single-cell images Herlev public dataset) and 90 WSIs collected in-house and annotated by an expert cytopathologist.	Deep learning classifier for cervical cancer screening	Plain soft-max classifier; ordinal-regression and "soft-label" baselines.	Binary normal vs abnormal classification accuracy on normal/abnormal (Herlev cell level). Simulated tiles (regression-constraint model).	Reduces the amount of data to analyze, making slide review faster for the cytopathologist.
Kyrgiou 2016	Personalised management of women with cervical abnormalities using a clinical decision support scoring system.	n = 2,267 women attending opportunistic LBC screening with any age/ethnicity/menopausal status, but no HIV- or hepatitis B/C-positive or autoimmune disorders (3565 visits analysed with 3561 complete data).	Clinical Decision Support Scoring System (DSSS) built with a multilayer perceptron ANN.	Conventional cytology thresholds (ASCUS+, LSIL+, HSIL+) ± high-risk HPV DNA test.	MLP ANN classifier. Stratified performance by cytology category.	Potential for saving resources by diagnosing high-risk cases early and avoiding unnecessary treatments. Suitable for software or mobile apps, making it easier to use.
Bao 2020	The artificial intelligence-assisted cytology diagnostic system in large-scale cervical cancer screening: A population-based cohort study of 0.7 million women.	n = 703,103 women (aged 20–65) attending maternal child hospitals in 83 counties covered under cervical screening program, screened by AI (15,494 women with abnormal cytology attended colposcopy. 98,549 women were independently screened by both AI and manual reading (34,738 abnormal + 63,811 normal)).	AI-assisted cervical cancer screening based on deep learning models.	Manual, randomly selected subsample, double examined by cytologists and reviewed by pathologists for supervision.	Detection of histologically confirmed cervical intraepithelial neoplasia grade 2 or worse (CIN2+).	Reduces human resource costs. Potential to treat a greater number of cases quickly. Identifies more cases earlier, resulting in cheaper treatment. Requires initial investment in equipment and software, but saves money over time.
Pereira 2024	Development, Validation, and Implementation of an Augmented Multiwell, Multitarget Quantitative PCR for the Analysis of Human Papillomavirus Genotyping through Software Automation, Data Science, and Artificial Intelligence.	n = 1,040 liquid-based cervical cytology samples from women attending the national cervical cancer screening program (Belgium), selected on HPV type-specific positivity ratios (validation), n = 1,043 samples analyzed for qualitative comparison.	Software automation via the FastFinder semi-automatic AI-based PCR curve analysis plugin for RIATOL qPCR.	Manual analysis by LightCycler 480 Software acted as the reference comparator.	Qualitative concordance (k coefficient) and quantitative agreement in Cq (threshold) and viral load between FastFinder and manual analysis; improved accuracy and reduced analysis time. Improved viral load quantification; reduced training time; streamlined routine analysis; minimized user variability.	Training time reduced by 33%; analysis time decreased 90% per plate.

Sood 2024	Enhancing pap smear image classification: integrating transfer learning and attention mechanisms for improved detection of cervical abnormalities.	n = 963 digitized LBC pap-smear images from the Mendeley LBC dataset (open-access), annotated by expert cytopathologists	Development of CNN-based image-classification pipeline with transfer learning (TL) + self-attention + preprocessing (normalisation, resizing, histogram of oriented gradients (HOG)).	Same TL backbones without attention or without preprocessing (reported side-by-side).	Image-level classification accuracy and related metrics on held-out test set. Comparison with other published approaches.	Pre-trained AI models mean savings on local datasets. More accurate results, fewer errors, and fewer extra tests. Simple image cleaning steps improve results at no additional cost. Cloud-based AI is cheaper than robust local servers.
Karasu-Benyes 2022	A Comparative Analysis of Deep Learning Models for Automated Cross-Preparation Diagnosis of Multi-Cell Liquid Pap Smear Images.	n = 1,193 LBC images. Training set: 963 SurePath multi-cell LBC images (public database). Test set: 230 ThinPrep images from a cytopathology teaching-slide collection at a Women & Infants Hospital, re-reviewed and re-classified by two pathologists.	Development/comparison of AI classifiers for Pap-smear image diagnosis.	Other TL models (DenseNet121, ResNet101, ResNet152, EfficientNetB0) and an ensemble of the top five models.	Image-level diagnostic performance on SurePath dataset. Generalisability to ThinPrep and impact of Deep CORAL domain adaptation.	Small, light AI models preserve functionality but need less power, memory, and resources. Models can be adapted to different lab methods, saving retraining costs. Easier to run in labs with limited computers.
Crowell 2019	CytoProcessorTM: A New Cervical Cancer Screening System for Remote Diagnosis.	n = 1,352 cervical screening patients aged ≥ 18, diagnosed by both TIS and CytoProcessor™. Random sample taken from routine TIS screens at a private lab.	Diagnostic accuracy comparison of two automated screening systems.	ThinPrep Imaging System (HOLOGIC).	Diagnostic sensitivity and specificity comparison between CytoProcessor™ and ThinPrep Imaging System.	CytoProcessor™ offers better sensitivity detection than TIS with equal specificity, streamlines workflow, reduces costs and improves access to faster cervical cancer diagnosis, especially in low-resource settings.
Kurita 2023	Accurate deep learning model using semi-supervised learning and Noisy Student for cervical cancer screening in low magnification images.	n = 140 WSIs from LBC specimens (56,996 images from the slides) diagnosed by two cytologists (≥20 y & 10 y experience) and two cytopathologists (≥10 y experience). Cervical specimens from patients who had not undergone hysterectomy or cervical conization; only the first specimen per patient in study period.	Semi-supervised deep-learning model using Noisy Student training with EfficientNet-B3 CNN	Performance compared across model stages (Stage 1 vs 2 vs 3).	Measured AUC, accuracy, F1 Score across 40 slides.	Reduces cytologist workload and is suitable for low-resource settings.
Xue 2020	Development and validation of an artificial intelligence system for grading colposcopic impressions and guiding biopsies.	n = 9,435 patients aged 24–65, with indications for colposcopy who underwent cytology/ HPV, and pathology screening, with definitive pathology results and with colposcopic images, from six hospitals. Retrospective collection from archived hospital databases.	Colposcopic Artificial Intelligence Auxiliary Diagnostic System (CAIADS) to grade colposcopic impressions and guide biopsies.	Original colposcopic interpretation by colposcopists; pathology as gold standard.	Agreement between CAIADS-graded colposcopic impressions and pathology; diagnostic performance for detecting HSIL+ at different biopsy thresholds. Biopsy site prediction accuracy.	Cost-effectiveness improved by diagnosis quality (fewer false negatives means fewer unnecessary procedures).
Crowell 2019	Adaptation of CytoProcessor for cervical cancer screening of challenging slides.	n = 293 patients selected from routine workflow in a public hospital (patients aged ≥ 18); diagnoses compared between CytoProcessor and conventional microscopy (309 slides total; 216 slides used for diagnostic performance; 93 slides used for diagnostic duration (timing)).	Comparison of automated AI-based CytoProcessor diagnosis vs conventional microscope screening; discordances adjudicated by consensus committee (truth diagnosis).	Conventional microscope screening (no additional comparator reported).	Diagnostic accuracy (sensitivity & specificity) of CytoProcessor vs microscope using a truth diagnosis (concordance or consensus) as reference. Time to diagnose.	Supports large-scale remote screening, enabling more slides to be processed without increasing workforce.
Pathania 2019	Point-of-care cervical cancer screening using deep learning-based microholography.	n = 28 adult women referred from previously abnormal Pap smears, yielding 35 cervical specimens (28 brushings, 7 biopsies); in vitro validation performed on CaSki, HeLa, and C33A cervical cancer cell lines	Point-of-care (POC) HPV detection using microholography and deep learning.	Roche Cobas HPV test (FDA-approved) used as the gold standard comparator for HPV16 and 18 positivity.	Sensitivity and specificity of AIM-HPV assay for HPV16/18 DNA detection; concordance with Cobas assay. Processing time: <2 min per full-sized image on single CPU Brushing vs biopsy correlation.	Low-cost deployment, portable POC device.

Ke 2021	Quantitative analysis of abnormalities in gynecologic cytopathology with deep learning.	n = 130 LBC digital slides scanned into WSIs at a Tumor Hospital.	Computer-aided diagnostic pipeline (segmentation + classification + spatial correlation + nuclear-area analysis + aggregation)	Single-arm diagnostic accuracy study.	Diagnostic performance of the AI system on slide-level abnormal detection. Prediction of ASC-US, CIN 1+, CIN 2+, SCC. Pixel accuracy ± 0.001; mean pixel accuracy ± 0.007; mean IoU ± 0.007	Automation detection leads to faster and more objective screening.
Xue 2023	Assessing artificial intelligence enabled liquid-based cytology for triaging HPV-positive women: a population-based cross-sectional study.	n = 3,514 women aged ≥ 21, attending opportunistic cervical cancer screening in outpatient clinics (489 HPV-positive underwent triage by AI-LBC, human cytologists, and HPV16/18 genotyping followed by colposcopic assessment).	AI-enabled liquid-based cytology (AI-LBC), human cytologists' interpretation, and HPV16/18 genotyping for triage.	Human cytology and HPV16/18 genotyping.	Diagnostic performance of AI-LBC, human cytologists, and HPV16/18 typing for detecting CIN2+ and CIN3+ among HPV-positive women. Efficiency of triage strategies: colposcopy referral rates and number of colposcopies needed per CIN2+ or CIN3+ detected.	Digitized scanning and cloud-based AI-LBC enable high-throughput, objective triage; deployable via AI cloud services in low-resource settings.
Elakkiya 2022	Cervical Cancer Diagnostics Healthcare System Using Hybrid Object Detection Adversarial Networks.	n = 1,993 patients from heterogeneous open-source cervical colposcopy datasets (3,105 colposcopy images included)	FSOD-GAN: hybrid small-object detection (Faster RCNN) integrated within a GAN for cervical spot localization and multiclass lesion classification.	Standalone Faster RCNN and standalone GAN models.	Diagnostic performance of FSOD-GAN for normal vs abnormal classification and staging.	Reduces manual workload and improves access in low-resource settings.
Shen 2023	Cost-effectiveness of artificial intelligence-assisted liquid-based cytology testing for cervical cancer screening in China.	Simulated cohort of n = 100,000 HPV-negative, unvaccinated, unscreened women aged = 30, followed for lifetime (80 y)	AI-assisted LBC screening using Landing CytoScanner versus manual LBC and HPV-DNA testing.	HPV DNA primary screening strategies.	Incremental cost-effectiveness ratio (ICER) in US$/QALY gained AI-LBC every 5 y: ICER $8 790/QALY vs next cheapest non-dominated strategy.	Probabilistic sensitivity: AI-LBC every 3 year optimal if sensitivity & specificity ↓10 %.
Wentzensen 2021	Accuracy and Efficiency of Deep-Learning-Based Automation of Dual Stain Cytology in Cervical Cancer Screening.	Total patients across studies: n = 4,253. Three datasets: (1) Biopsy Study—women referred to colposcopy (n = 602 slides); (2) ACSS—HIV-positive men who have sex with men (n = 318 slides); (3) KPNC screening—HPV-positive women undergoing routine screening (n = 3,333 slides).	Automated deep-learning evaluation of dual-stain (p16/Ki-67) cytology slides using CYTOREADER.	Conventional Pap smear cytology and manual DS.	Detection of CIN3+ (Biopsy Study; KPNC screening) and AIN2+ (ACSS). Clinical efficiency vs Pap cytology and manual DS in HPV-positive women (KPNC).	Automated DS evaluation offers objective, efficient, and cost-effective cervical cancer screening, reducing false positives.