



Review Article



Multimodal Fusion Artificial Intelligence Models for Pathological Diagnosis in Early Cervical Cancer Screening: A Narrative Review

Zhi-Feng Wei, He Qin, Shui-Juan Lu, Ping Ruan, Ze-Chao Zhang and Min Zhu*

The Second Affiliated Hospital of Guangxi University of Chinese Medicine, Nanning, Guangxi, China

Received: February 14, 2026 | Revised: May 18, 2026 | Accepted: June 08, 2026 | Published online: June 29, 2026

Abstract

Cervical cancer is a major malignancy that threatens women's health, and early screening is a core strategy for reducing its incidence and mortality. Multimodal fusion artificial intelligence (AI) pathological diagnosis models integrate multidimensional data—including cytological images, colposcopic images, whole-slide histopathological images, clinical data, and molecular testing results—and may enhance the detection sensitivity, grading accuracy, and screening efficiency for early cervical cancer and precancerous lesions. However, traditional cervical cancer screening methods face limitations such as high subjectivity, reliance on single-source information, relatively low efficiency, and insufficient primary care resources. Furthermore, existing reviews mostly focus on single-modal AI models or specific technical aspects, lacking a comprehensive analysis of the full technical framework and clinical translation pathways of multimodal fusion models. This review aims to comprehensively present the development and application of multimodal fusion AI models in pathological diagnosis for early cervical cancer screening. Specifically, it comprehensively details the technical architecture, data modalities, and fusion strategies—including deep learning, attention mechanisms, and cross-modal alignment techniques—that enable the complementary representation of morphological, clinical, and molecular information. Additionally, the review integrates recent advances in clinical applications and evaluates current translational challenges, providing insights into clinical validation pathways to bridge technological innovation and practical healthcare delivery. In conclusion, with further technological refinement and clinical validation, multimodal fusion AI may become a useful tool for improving the precision and efficiency of cervical cancer screening and prevention, and may inform the standardized application and translational research of AI technology in this field.

Introduction

Cervical cancer ranks as the fourth most common malignant tumor among women globally, with approximately 604,000 new cases and 342,000 deaths worldwide in 2020. In China, the annual incidence reaches 140,000 cases, resulting in 37,000 deaths, and the age of onset is trending younger.¹ High-grade squamous intraepithelial lesions (HSIL/CIN II) represent a critical intervention point for early-stage cervical cancer. The 5-year survival rate can exceed 90% following standardized treatment, making highly efficient

and accurate early screening central to cervical cancer prevention and control. The World Health Organization has proposed a global strategy for the elimination of cervical cancer by 2030, with three core targets to be achieved by that year: 90% of girls fully vaccinated against human papillomavirus (HPV) by age 15, 70% of women screened with a high-performance test by age 35 and again by age 45, and 90% of women with cervical disease receive appropriate treatment and management.²

China faces a shortage of more than 80% of the required number of pathologists, and primary-level screening capacity is weak. Traditional manual slide review is associated with issues such as underdiagnosis, misdiagnosis, low efficiency, and poor standardization.³ Artificial intelligence (AI) technology has demonstrated significant potential in the analysis of pathological images and the interpretation of molecular data. Single-modal AI models have been preliminarily applied to cervical cancer screening; however, their reliance on a single information dimension limits their ability to comprehensively reflect the pathological characteristics and biological behavior of lesions. Multimodal fusion technology, by

Keywords: Multimodal fusion; Artificial intelligence; Pathological diagnosis; Cervical cancer; Early screening; Molecular testing; HPV testing.

***Correspondence to:** Min Zhu, Guangxi University of Chinese Medicine, 10 Huadong Road, Xingning District, Nanning, Guangxi 530000, China. ORCID: <https://orcid.org/0000-0001-5900-5330>. Tel: +86-15177800136, E-mail: chao616317728@foxmail.com

How to cite this article: Wei ZF, Qin H, Lu SJ, Ruan P, Zhang ZC, Zhu M. Multimodal Fusion Artificial Intelligence Models for Pathological Diagnosis in Early Cervical Cancer Screening: A Narrative Review. *Oncol Adv* 2026;4(2):e 00004. doi: 10.14218/OnA.2026.00004.

integrating two or more diagnostic information sources, compensates for the shortcomings of single modalities, thereby enhancing the diagnostic accuracy, stability, and generalizability of models. This approach has become a research focus in early cervical cancer screening.

Existing reviews on AI-assisted cervical cancer screening have predominantly focused on single-modal cytological image analysis or individual molecular detection technologies, with limited in-depth discussion of the integration of multi-source diagnostic information. Few studies have comprehensively analyzed the complete technical chain from multimodal data preprocessing to clinical deployment, nor have they comprehensively addressed the translational barriers from laboratory research to large-scale primary care applications.

With the rapid development of multimodal fusion technology and the urgent need to improve cervical cancer screening efficiency in low-resource settings, a comprehensive review integrating technical advances and clinical practice is urgently needed to guide the standardized development and application of this technology.

This review aims to comprehensively synthesize multimodal fusion AI pathological diagnosis models, detailing their core technical architectures and clinical applications across the entire screening continuum. By critically evaluating existing translational bottlenecks and future development directions, this work highlights the clinical applicability and primary care deployment potential of these models, ultimately delivering actionable insights for both technical researchers and clinical practitioners.

Limitations of traditional cervical cancer screening techniques

Traditional cervical cancer screening techniques, primarily including cervical liquid-based cytology (TCT/LCT), HPV testing, cervical biopsy, and colposcopy, remain the mainstay for current screening and diagnosis. However, they possess significant inherent limitations that constrain improvements in screening efficacy.

Limitations of cervical liquid-based cytology

As a first-line screening technique, liquid-based cytology has improved slide preparation quality and microscopic visualization but suffers from multiple shortcomings: (1) Strong subjectivity in diagnosis: The inter-observer agreement (Kappa value) among different physicians for interpreting borderline lesions is only 0.5–0.7, with a false-positive rate of 15–30% and a false-negative rate of 8–15%.⁴ (2) Limited morphological information: It provides only two-dimensional cellular morphological features, making it difficult to differentiate between inflammation and precancerous lesions. (3) Challenges in the standardization of preparation and data: Specimen adequacy remains affected by sampling and preparation quality, and unsatisfactory rates vary across liquid-based cytology platforms and laboratories.⁵ (4) Low efficiency: Manual screening of a single slide requires 3–5 min, making it difficult to adapt to large-scale population screening.

Limitations of HPV testing

High-risk HPV (HR-HPV) testing is a core molecular screening technique, yet it has limitations: (1) Insufficient identification of carcinogenic activity: Routine DNA testing cannot distinguish between transient and persistent infections, leading to overdiagnosis and overtreatment, while E6/E7 messenger RNA (mRNA) testing alone lacks sufficient specificity.⁶ (2) Lack of subtype-specific differentiation: It cannot quantify the differential carcinogenic risks

associated with different HR-HPV subtypes. (3) Susceptibility to interference: Significant variations in sensitivity and specificity exist among different technologies, and results can be biased due to inadequate sampling or cross-contamination.⁷ (4) Low standardization: Poor comparability of test results across different laboratories hinders large-scale implementation.

Limitations of cervical biopsy technique

As the diagnostic “gold standard,” cervical biopsy has limitations: (1) Sampling limitations: The omission rate for small or occult lesions can reach 15–30%, and sampling from the endocervical transformation zone is challenging.⁸ (2) Strong subjectivity in diagnosis: Low diagnostic consistency among different physicians for borderline lesions. (3) Non-standardized sample processing: Tissue deformation can affect diagnostic accuracy. (4) Limited integration with clinical management: Cervical biopsy provides localized histopathological information but does not, by itself, integrate cytology, HPV status, colposcopic findings, and clinical history, which may limit its ability to guide individualized management.⁹

Molecular pathological diagnostic techniques

Molecular pathological diagnostic techniques, centered on the molecular mechanisms underlying cervical carcinogenesis, provide a rich source of molecular feature data for multimodal fusion AI models. This drives the evolution of screening from “morphological identification” toward “precise molecular mechanism-based prediction,” establishing a comprehensive molecular diagnostic system encompassing early warning, lesion identification, and risk stratification.

Core molecular detection technologies and applications

HR-HPV detection

HR-HPV detection is categorized into DNA genotyping and E6/E7 mRNA detection. Conventional DNA testing exhibits high sensitivity (detection rate for CIN II > 95%) but cannot distinguish viral integration status. E6/E7 mRNA detection directly measures the expression of viral oncogenes E6 and E7, which are responsible for the malignant transformation of cervical epithelial cells. Its primary clinical purpose is to triage HPV-positive women, distinguishing between transient infections that will resolve spontaneously and persistent infections that carry a high risk of progression to precancerous lesions. Most commercial E6/E7 mRNA assays use a cutoff value based on relative light units (RLU) compared to a positive control; for example, the Aptima HPV assay defines a positive result as an RLU-to-cutoff (RLU/CO) ratio ≥ 1.0 . However, E6/E7 mRNA testing has limitations: it cannot distinguish between different HR-HPV subtypes with varying carcinogenic risks, and its sensitivity for detecting adenocarcinoma *in situ* is slightly lower than that of HPV DNA testing. It is recommended as the preferred triage method for HPV-positive women aged 25 years and older and can also be used as a primary screening test in combination with cytology. Among these, the Aptima HPV assay is the only HR-HPV E6/E7 mRNA detection technology approved by the U.S. Food and Drug Administration and is widely used for clinical triage.¹⁰

DNA methylation detection

This technique detects abnormal hypermethylation in the promoter regions of tumor suppressor genes (e.g., *PAX1*, *JAM3*, *FAM19A4*). This epigenetic alteration precedes cytological abnormalities and

serves as a key technology for ultra-early warning. The dual-gene detection of *PAX1/JAM3* achieves a sensitivity of 92% and a specificity of 88% for CIN II, with an AUC of 0.94, outperforming combined HPV and cytology screening.¹¹ The *FAM19A4/miR124-2* assay demonstrates an overall detection rate for cervical cancer exceeding 98% and is suitable for the early diagnosis of rare pathological types.¹²

High-throughput and precise amplification technologies

Digital PCR (ddPCR) enables absolute quantification at the single-copy level, making it suitable for monitoring low-load HPV infections and minimal residual disease. The levels of cell-free HPV DNA detected by ddPCR correlate well with tumor size.¹³ Next-generation sequencing facilitates comprehensive analysis of HPV genotyping, viral integration sites, and methylation profiles. The HPV Pool-Seq strategy can reduce sequencing costs by 60% while maintaining a detection sensitivity of 97.1%.¹⁴

Advantages and limitations of molecular pathological diagnostic techniques

The core advantages of molecular pathological diagnostic techniques lie in their precision, early risk prediction, and diversity: they capture the molecular essence of lesions, enabling early detection; they assess the risk of lesion progression through quantitative molecular indicators, providing a basis for precise stratification; and the variety of available technologies allows for flexible combinations tailored to different screening scenarios.¹⁵ However, limitations include uneven costs, an incomplete standardization system, inherent limitations of individual technologies, and insufficient clinical validation and translation. High-end technologies such as next-generation sequencing and ddPCR face challenges in widespread adoption at the primary care level, and inter-assay variability among different reagents can reach 5–10%.¹⁶

Core technical framework of the multimodal fusion AI pathology diagnosis model

A multimodal fusion AI pathology diagnosis model can be developed through a closed-loop technical framework of “data standardization → feature representation → cross-modal fusion → model learning → inference validation → interpretable output”. The core objectives are to eliminate heterogeneity in multimodal data, uncover deep pathological features, achieve precise lesion stratification, and meet clinical applicability requirements.

Standardized preprocessing of multimodal data

Multimodal data for cervical cancer screening include image modalities (TCT, colposcopy, histopathological WSI), molecular modalities (HPV genotype, immunohistochemistry quantitative values), and structured text modalities (clinical data). Preprocessing forms the foundation for model training.

Image modalities: Unsupervised anomaly detection algorithms are employed to discard invalid slides, and color constancy algorithms are used to correct staining deviations. WSIs are tiled using non-overlapping sliding windows, while TCT images undergo nuclear segmentation to extract individual cells, followed by normalization. Data augmentation techniques, including rotation, flipping, and GAN-based synthesis, are combined to mitigate sample imbalance.¹⁷

Non-image modalities: HPV genotypes are one-hot encoded, and molecular numerical values are normalized using Z-score or Min-Max scaling. Missing values are imputed using the K-Nearest

Neighbors method. Clinical text is converted into low-dimensional numerical vectors via embedding layers to align with the feature dimensions of image data.¹⁸

Dataset construction: A dual-blind review system combined with a gold-standard annotation protocol is adopted to build a paired multimodal dataset, requiring an annotation consistency Kappa value ≥ 0.85 . Mainstream public datasets include Herlev, SIPaKMeD, and CRIC Cervix.¹⁹

Multimodal feature extraction and representation learning

High-dimensional, robust pathological features are extracted from heterogeneous modalities and categorized into two approaches: single-modality-specific feature extraction and cross-modal general representation learning.

Single-modality-specific feature extractors

Pathological images: Convolutional neural networks (CNNs) such as ResNet and DenseNet are used to extract local morphological features, while attention-based CNNs capture global histological structures. Vision Transformers and Swin Transformers are better suited for the diffuse and multifocal nature of lesions, offering superior extraction of subtle lesion features. Specialized pathology models such as Clustering-Constrained Attention Multiple-Instance Learning address the weak annotation problem in WSIs.²⁰

Non-image modalities: Multi-Layer Perceptron and TabNet are employed to extract nonlinear relationships from molecular features, while BioBERT extracts semantic features from clinical text, uncovering associations between clinical information and cervical lesions.²¹

Cross-modal general representation learning

Modality heterogeneity is eliminated and feature-space alignment is achieved through fine-tuning of medical pre-trained models, cross-modal contrastive learning (CMCL), and cross-modal knowledge distillation. Pre-trained models such as Path-ImageNet and Med-CLIP, based on the Cancer Genome Atlas – Cervical Squamous Cell Carcinoma and Endocervical Adenocarcinoma data, can mitigate poor generalization associated with small sample sizes. CMCL, constrained by InfoNCE loss, brings features of similar samples closer together for semantic alignment. Knowledge distillation transfers knowledge from stronger teacher models to student models, thereby improving representation learning across different modalities.²²

Cross-modal feature fusion strategies

Fusion strategies are categorized into data-level, feature-level, and decision-level fusion. For cervical cancer screening, feature-level fusion is the primary approach, supplemented by decision-level fusion.

Data-level fusion: Involves direct concatenation of raw data, preserving complete information but being susceptible to redundancy and heterogeneity interference; thus, it is used only as an auxiliary method.²³

Feature-level fusion: This is the mainstream approach and includes simple concatenation, channel attention mechanisms (SE-Net, CBAM), cross-modal attention, tensor decomposition, and graph neural network (GNN)-based fusion. Among these, cross-modal attention can capture associations between HPV infection and cellular atypia, while GNN-based fusion can model the progressive pathological relationship of “HPV infection → cellular atypia → tissue lesion”.²⁴

Decision-level fusion: Integrates results from individual single-modality models through weighted voting, Bayesian fusion, or

stacking ensemble methods to enhance model robustness, making it suitable for multi-hospital data integration.^{25,26}

Fusion strategies should be selected according to the specific screening task, data availability, modality heterogeneity, and deployment scenario. Feature- or joint-fusion approaches may be appropriate when paired multimodal data are available, whereas decision-level or late-fusion approaches can be useful when modalities are incomplete or collected across institutions.²⁶

Model training and optimization

To address the three major challenges of sample imbalance, weak annotation, and multi-hospital distribution shift, an adapted training framework is constructed: (1) Loss functions such as Focal Loss and Weighted Cross-Entropy Loss are used to tackle sample imbalance, combined with contrastive loss and modality consistency loss to ensure the quality of fused features. (2) Training strategies employ transfer learning and incremental learning to mitigate domain shift, and Domain-Adversarial Neural Networks are introduced to improve cross-hospital generalization capability. (3) Model training and optimization can involve appropriate optimizer selection and learning-rate scheduling strategies to improve convergence and generalization.²⁷ Model evaluation focuses on clinical needs, with core metrics including sensitivity, specificity, AUC, F1-score, positive predictive value, and negative predictive value.²⁸

Inference deployment and interpretability

Lightweight deployment

To meet primary-level needs, techniques such as model pruning, quantization, knowledge distillation, and TensorRT optimization are applied. Post-training quantization can convert medical imaging models to INT8 precision, reducing model size, computational requirements, and inference latency while maintaining model performance, thereby facilitating deployment in resource-constrained settings.²⁹

Explainable AI (XAI) techniques

Grad-CAM and LIME are used to generate heatmaps on pathological images, highlighting the lesion regions on which the model focuses. The contribution of each modality is calculated to assist physicians in understanding the decision logic. Quantitative image-analysis features can reflect pathological criteria such as nuclear enlargement and nuclear hyperchromasia, helping align computational outputs with cytological diagnostic criteria.³⁰

Limitations of this review

This narrative review has several limitations that should be acknowledged. First, the literature search was not conducted according to a predefined exhaustive search protocol, which may have led to the omission of some relevant studies, particularly those published in non-English languages or in less accessible journals. Second, most of the included studies were retrospective or single-center studies, and the number of large-scale, multicenter, prospective clinical trials remains limited, which may affect the generalizability of the reported performance data. Third, the review focuses primarily on technical and clinical aspects, with relatively less discussion on the economic evaluation and cost-effectiveness of multimodal fusion AI models in different healthcare settings. Finally, the rapid development of AI technology means that some emerging advances, such as the application of large language models in pathological diagnosis, may not be fully covered in this review.

Despite significant progress, multimodal fusion AI pathological diagnosis models still face four core bottlenecks in clinical translation and large-scale application:

Heterogeneity and standardization challenges of multimodal data: Variations in pathological image scanning equipment and HPV testing platforms across different hospitals can lead to domain shift during cross-institutional data fusion. The lack of unified standards for multimodal data acquisition, annotation, and quality control limits model generalization capabilities.³¹

High annotation costs and inconsistent sample quality: High-grade cervical lesion samples are scarce, with a high proportion of weakly labeled or unlabeled data. Manual annotation requires experienced pathologists, incurring high costs. Non-standardized sample collection and processing procedures in primary healthcare institutions result in uneven data quality, adversely affecting model training efficacy.³²

The dilemma of balancing fusion efficiency and interpretability: Complex cross-modal fusion models (e.g., Transformer + Graph Neural Networks) offer high accuracy but suffer from slow inference speeds and poor interpretability, making them difficult to adapt to the real-time demands of primary clinical settings. Conversely, lightweight models often inadequately extract features, leading to compromised diagnostic performance.³³

Clinical validation and translational barriers: The diagnostic results of AI models must align with the gold standard of pathological diagnosis. However, there is currently a lack of unified validation protocols and evaluation standards for multimodal AI pathological models. The “black box” nature of AI systems undermines clinician trust, while data privacy and ethical concerns hinder multi-institutional data sharing.³⁴

Furthermore, insufficient technical equipment and personnel capabilities in primary healthcare institutions, coupled with imperfect AI model operation, maintenance, and update systems, also limit the deployment of this technology in primary care settings.³⁵

Future development directions and prospects

To address existing bottlenecks, the development of multimodal fusion AI pathological diagnosis models will focus on standardization, lightweight design, interpretability, and clinical closed-loop integration, while expanding toward multi-omics fusion and federated learning, thereby driving the technology from the laboratory to large-scale clinical applications.

Construction of standardized multimodal datasets

Recent work has demonstrated the feasibility of constructing multimodal cervical lesion datasets integrating colposcopic images, dynamic acetic acid reactions, iodine staining images, diagnostic reports, and pathological results. Future efforts should expand such resources into standardized, multicenter databases with unified protocols for data acquisition, annotation, and quality control.³⁶

Development of lightweight and highly interpretable models

Research and development of lightweight Transformer architectures and efficient CNN frameworks should be advanced. By integrating model compression and edge computing technologies, AI systems compatible with mobile and portable devices can be developed, lowering the barrier to application in primary care settings. Research on XAI technology should be deepened to achieve visualization and semantic interpretation of pathological features, thereby improving the transparency, interpretability, and clinical usability of AI models.³⁷

Construction of a full-process screening closed-loop model

Data from the entire chain—including HPV testing, cytology, colposcopy, histopathology, and methylation testing—should be integrated to construct a full-process multimodal fusion AI model encompassing “screening–triage–diagnosis–treatment–follow-up”.^{9,38} This aims to achieve intelligent management throughout the entire cervical cancer prevention and control cycle. Such models could support risk-stratified triage and management recommendations based on screening results, while final clinical decisions should remain guided by established clinical guidelines and clinician judgment.^{9,38}

Promotion of deep multi-omics integration and large model applications

By integrating genomic, transcriptomic, proteomic, and metabolomic data, deep fusion of multi-omics data with imaging and clinical data can be achieved.^{39,40} This facilitates the exploration of associations between the molecular mechanisms of cervical cancer development and progression and imaging features, enhancing capabilities for precise lesion subtyping and prognosis prediction.^{39,40} In combination with medical foundation models and generalist medical AI models, future systems may support the joint interpretation of pathological images, molecular data, and clinical text, potentially assisting in structured report generation and clinical decision support.^{41,42}

Adoption of federated learning for cross-hospital collaborative training

Based on federated learning technology, joint model training using multimodal data from multiple hospitals can be conducted without sharing raw data, addressing issues of data privacy and domain shift. A nationwide collaborative AI model updating system should be established to enable real-time model iteration in primary healthcare institutions, enhancing the model’s generalization capability and clinical adaptability.

Improvement of standardized regulatory and industrial translation systems

Validation protocols, evaluation standards, and ethical guidelines for multimodal AI pathological diagnosis models should be formulated, promoting the approval process for Class III medical device registration of AI products. Future studies should evaluate whether AI screening technologies could be incorporated into medical insurance reimbursement schemes to reduce patient screening costs.^{43,44} Training in AI technology for primary healthcare personnel should be strengthened, and operation, maintenance, and technical support systems for AI models should be established to accelerate large-scale application at the primary care level.³⁵

Conclusions

Multimodal fusion AI pathological diagnosis models integrate cytological, histopathological, clinical, and molecular information and may help address some limitations of traditional cervical cancer screening and single-modal AI approaches. Current evidence suggests that these models have potential to improve lesion characterization, screening efficiency, and risk stratification, but further multicenter prospective validation is required before routine clinical implementation. With continued technical refinement and standardized evaluation, multimodal fusion AI may support more precise and accessible cervical cancer screening.

Acknowledgments

The authors would like to thank all the researchers whose work has been cited in this review. We also appreciate the valuable comments and suggestions provided by the editors and reviewers, which have significantly improved the quality of this manuscript. We thank the Second Affiliated Hospital of Guangxi University of Chinese Medicine for providing literature retrieval support.

Funding

None.

Conflict of interest

The authors declare that they have no conflict of interest.

Author contributions

Conceptualization (ZFW, ZCZ), data curation (SJL, PR, HQ), writing – original draft (ZCZ, HQ, SJL), and writing – review and editing (ZCZ, MZ). All authors have approved the final version and publication of the manuscript.

References

- [1] Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, *et al*. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J Clin* 2021;71(3):209–249. doi:10.3322/caac.21660, PMID: 33538338.
- [2] World Health Organization. Global strategy to accelerate the elimination of cervical cancer as a public health problem. Geneva: World Health Organization; 2020. Available from: <https://iris.who.int/handle/10665/336583>.
- [3] Hou X, Shen G, Zhou L, Li Y, Wang T, Ma X. Artificial Intelligence in Cervical Cancer Screening and Diagnosis. *Front Oncol* 2022;12:851367. doi:10.3389/fonc.2022.851367, PMID:35359358.
- [4] Fontham ETH, Wolf AMD, Church TR, Etzioni R, Flowers CR, Herzig A, *et al*. Cervical cancer screening for individuals at average risk: 2020 guideline update from the American Cancer Society. *CA Cancer J Clin* 2020;70(5):321–346. doi:10.3322/caac.21628, PMID:32729638.
- [5] Fontaine D, Narine N, Naugler C. Unsatisfactory rates vary between cervical cytology samples prepared using ThinPrep and SurePath platforms: a review and meta-analysis. *BMJ Open* 2012;2(2):e000847. doi:10.1136/bmjopen-2012-000847, PMID:22505312.
- [6] Zhao C, Li M, Chen R, Zhao Y, Li J, Zhao W, *et al*. Comparison of different mRNA testing technologies with HPV DNA testing for predicting ASCUS triage and post-cone excision outcomes: a systematic review and meta-analysis. *Syst Rev* 2026;15(1):53. doi:10.1186/s13643-025-02947-4, PMID:41521305.
- [7] Wei F, Georges D, Man I, Baussano I, Clifford GM. Causal attribution of human papillomavirus genotypes to invasive cervical cancer worldwide: a systematic analysis of the global literature. *Lancet* 2024;404(10451):435–444. doi:10.1016/S0140-6736(24)01097-3, PMID:39097395.
- [8] Yao X, Qie M, Kang L, Li M, Liu Q, Liang X, *et al*. Assessing the risk of undetected cervical cancer in women with a biopsy of high-grade cervical intraepithelial neoplasia: predictive value of cytology and endocervical curettage (ECC). *BMC Cancer* 2025;25(1):1238. doi:10.1186/s12885-025-14565-3, PMID:40731391.
- [9] Perkins RB, Guido RS, Castle PE, Chelmow D, Einstein MH, Garcia F, *et al*. 2019 ASCCP Risk-Based Management Consensus Guidelines for Abnormal Cervical Cancer Screening Tests and Cancer Precursors. *J Low Genit Tract Dis* 2020;24(2):102–131. doi:10.1097/LGT.0000000000000525, PMID:32243307.
- [10] Arbyn M, Roelens J, Cuschieri K, Cuzick J, Szarewski A, Ratnam S, *et al*.

- The APTIMA HPV assay versus the Hybrid Capture 2 test in triage of women with ASC-US or LSIL cervical cytology: a meta-analysis of the diagnostic accuracy. *Int J Cancer* 2013;132(1):101–108. doi:10.1002/ijc.27636, PMID:22610699.
- [11] Sun D, Shu C, Zeng F, Xu D, Zhao X. The performance of JAM3/PAX1 methylation in the diagnosis of high-grade squamous intraepithelial lesions for women with high-risk HPV infection. *BMC Cancer* 2024;24(1):1514. doi:10.1186/s12885-024-13299-y, PMID:39696066.
 - [12] Larsson GL, Carlsson J, Helenius G, Bergengren L. Molecular triage on HPV-positive samples in a cervical screening setting. *PLoS One* 2025;20(10):e0333539. doi:10.1371/journal.pone.0333539, PMID:41082489.
 - [13] Kim SH, Weigelt B. Circulating cell-free (cf)DNA analysis: Current technologies and applications in gynecologic cancer. *Gynecol Oncol Rep* 2024;54:101431. doi:10.1016/j.gore.2024.101431, PMID:38947418.
 - [14] Guan W, Liu S, Chen Y, Chu C, Peng G, Wang J, *et al*. HPVPool-Seq: a genotype-guided pooling strategy for cost-effective next-generation sequencing detection of HPV integration in cervical samples. *Microbiol Spectr* 2025;13(10):e0139925. doi:10.1128/spectrum.01399-25, PMID:40827881.
 - [15] Güzel C, van Sten-Van't Hoff J, de Kok IMCM, Govorukhina NI, Boychenko A, Luider TM, *et al*. Molecular markers for cervical cancer screening. *Expert Rev Proteomics* 2021;18(8):675–691. doi:10.1080/14789450.2021.1980387, PMID:34551656.
 - [16] Zhang Y, Yuan J, Chen L. Transforming cervical cancer pathological diagnosis through artificial intelligence: progress, performance, and barriers to clinical implementation. *Front Oncol* 2025;15:1716018. doi:10.3389/fonc.2025.1716018, PMID:41640440.
 - [17] Litjens G, Kooi T, Bejnordi BE, Setio AAA, Ciompi F, Ghafoorian M, *et al*. A survey on deep learning in medical image analysis. *Med Image Anal* 2017;42:60–88. doi:10.1016/j.media.2017.07.005, PMID:28778026.
 - [18] Sun F, Ge X, Deng L, Wang Y, Wu H, Wang L, *et al*. Research on a method for early diagnosis and progression prediction of cervical cancer based on imaging omics and molecular omics. *Front Oncol* 2026;15:1670852. doi:10.3389/fonc.2025.1670852, PMID:41743184.
 - [19] Kupas D, Hajdu A, Kovacs I, Hargitai Z, Szombathy Z, Harangi B. Annotated Pap cell images and smear slices for cell classification. *Sci Data* 2024;11(1):743. doi:10.1038/s41597-024-03596-3, PMID:38972893.
 - [20] Srinidhi CL, Ciga O, Martel AL. Deep neural network models for computational histopathology: A survey. *Med Image Anal* 2021;67:101813. doi:10.1016/j.media.2020.101813, PMID:33049577.
 - [21] Lee J, Yoon W, Kim S, Kim D, Kim S, So CH, *et al*. BioBERT: a pre-trained biomedical language representation model for biomedical text mining. *Bioinformatics* 2020;36(4):1234–1240. doi:10.1093/bioinformatics/btz682, PMID:31501885.
 - [22] Li X, Li L, Li M, Yan P, Feng T, Luo H, *et al*. Knowledge distillation and teacher-student learning in medical imaging: Comprehensive overview, pivotal role, and future directions. *Med Image Anal* 2026;107(Pt A):103819. doi:10.1016/j.media.2025.103819, PMID:41027250.
 - [23] Li Y, El Habib Daho M, Conze PH, Zeglache R, Le Boité H, Tadayoni R, *et al*. A review of deep learning-based information fusion techniques for multimodal medical image classification. *Comput Biol Med* 2024;177:108635. doi:10.1016/j.compbiomed.2024.108635, PMID:38796881.
 - [24] Qiao Y, Zhao L, Luo C, Luo Y, Wu Y, Li S, *et al*. Multi-modality artificial intelligence in digital pathology. *Brief Bioinform* 2022;23(6):bbac367. doi:10.1093/bib/bbac367, PMID:36124675.
 - [25] Carrillo-Perez F, Morales JC, Castillo-Secilla D, Gevaert O, Rojas I, Herrera LJ. Machine-Learning-Based Late Fusion on Multi-Omics and Multi-Scale Data for Non-Small-Cell Lung Cancer Diagnosis. *J Pers Med* 2022;12(4):601. doi:10.3390/jpm12040601, PMID:35455716.
 - [26] Huang SC, Pareek A, Seyyedi S, Banerjee I, Lungren MP. Fusion of medical imaging and electronic health records using deep learning: a systematic review and implementation guidelines. *NPJ Digit Med* 2020;3:136. doi:10.1038/s41746-020-00341-z, PMID:33083571.
 - [27] Mortazi A, Cicek V, Keles E, Bagci U. Selecting the best optimizers for deep learning-based medical image segmentation. *Front Radiol* 2023;3:1175473. doi:10.3389/fradi.2023.1175473, PMID:37810757.
 - [28] Klontzas ME, Groot Lipman KBW, Akinci D'Antonoli T, Andreychenko A, Cuocolo R, Dietzel M, *et al*. ESR Essentials: common performance metrics in AI-practice recommendations by the European Society of Medical Imaging Informatics. *Eur Radiol* 2026;36(2):1528–1540. doi:10.1007/s00330-025-11890-w, PMID:40753524.
 - [29] Qu C, Zhao R, Yu Y, Liu B, Yao T, Zhu J, *et al*. MedPTQ: a practical pipeline for real post-training quantization in 3D medical image segmentation. *J Med Imaging (Bellingham)* 2026;13(1):014006. doi:10.1117/1.JMI.13.1.014006, PMID:41709974.
 - [30] Tanaka S, Yamamoto T, Teramoto N. Quantitative Structural Analysis of Hyperchromatic Crowded Cell Groups in Cervical Cytology: Overcoming Diagnostic Pitfalls. *Cancers (Basel)* 2024;16(24):4258. doi:10.3390/cancers16244258, PMID:39766157.
 - [31] Miotto S, Wang F, Wang S, Jiang X, Dudley JT. Deep learning for healthcare: review, opportunities and challenges. *Brief Bioinform* 2018;19(6):1236–1246. doi:10.1093/bib/bbx044, PMID:28481991.
 - [32] Jiang P, Li X, Shen H, Chen Y, Wang L, Chen H, *et al*. A systematic review of deep learning-based cervical cytology screening: from cell identification to whole slide image analysis. *Artif Intell Rev* 2023;56(Suppl 2):2687–2758. doi:10.1007/s10462-023-10588-z.
 - [33] Rudin C. Stop Explaining Black Box Machine Learning Models for High Stakes Decisions and Use Interpretable Models Instead. *Nat Mach Intell* 2019;1(5):206–215. doi:10.1038/s42256-019-0048-x, PMID:35603010.
 - [34] Hassan M, Kushniruk A, Borycki E. Barriers to and Facilitators of Artificial Intelligence Adoption in Health Care: Scoping Review. *JMIR Hum Factors* 2024;11:e48633. doi:10.2196/48633, PMID:39207831.
 - [35] Gupta R, Kumar N, Bansal S, Singh S, Sood N, Gupta S. Artificial Intelligence-driven Digital Cytology-based Cervical Cancer Screening: Is the Time Ripe to Adopt This Disruptive Technology in Resource-constrained Settings? A Literature Review. *J Digit Imaging* 2023;36(4):1643–1652. doi:10.1007/s10278-023-00821-0, PMID:37029285.
 - [36] Sha Y, Zhang Q, Zhai X, Hou M, Lu J, Meng W, *et al*. CerviFusionNet: A multi-modal, hybrid CNN-transformer-GRU model for enhanced cervical lesion multi-classification. *iScience* 2024;27(12):111313. doi:10.1016/j.isci.2024.111313, PMID:39634563.
 - [37] Plass M, Kargl M, Kiehl TR, Regitnig P, Geißler C, Evans T, *et al*. Explainability and causability in digital pathology. *J Pathol Clin Res* 2023;9(4):251–260. doi:10.1002/cjp2.322, PMID:37045794.
 - [38] 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):864–879. doi:10.20892/j.issn.2095-3941.2024.0198, PMID:39297572.
 - [39] Liu F, Beck S, Yang L, Luo H, Zhang K. Advancing AI for multi-omics and clinical data integration in basic and translational cancer research. *Nat Rev Cancer* 2026;26(7):497–512. doi:10.1038/s41568-026-00922-2, PMID:42014628.
 - [40] Lipkova J, Chen RJ, Chen B, Lu MY, Barbieri M, Shao D, *et al*. Artificial intelligence for multimodal data integration in oncology. *Cancer Cell* 2022;40(10):1095–1110. doi:10.1016/j.ccell.2022.09.012, PMID:36220072.
 - [41] Moor M, Banerjee O, Abad ZSH, Krumholz HM, Leskovec J, Topol EJ, *et al*. Foundation models for generalist medical artificial intelligence. *Nature* 2023;616(7956):259–265. doi:10.1038/s41586-023-05881-4, PMID:37045921.
 - [42] Singhal K, Azizi S, Tu T, Mahdavi SS, Wei J, Chung HW, *et al*. Large language models encode clinical knowledge. *Nature* 2023;620(7972):172–180. doi:10.1038/s41586-023-06291-2, PMID:37438534.
 - [43] Shen M, Zou Z, Bao H, Fairley CK, Canfell K, Ong JJ, *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. doi:10.1016/j.lanwpc.2023.100726, PMID:37283979.
 - [44] Abramoff MD, Roehrenbeck C, Trujillo S, Goldstein J, Graves AS, Repka MX, *et al*. A reimbursement framework for artificial intelligence in healthcare. *NPJ Digit Med* 2022;5(1):72. doi:10.1038/s41746-022-00621-w, PMID:35681002.