

PERFORMANCE OF DIFFERENT CERVICAL CANCER PRIMARY SCREENING METHODS: AN UMBRELLA REVIEW AND UPDATED META-ANALYSIS

Phung Thi TRAN¹, Kunlawat THADANIPON^{1,2*}, Ammarin THAKKINSTIAN^{1,2} and Montarat THAVORNCHAROENSAP^{1,3}

1 Health Technology Assessment Graduate Program, Mahidol University, Thailand

2 Department of Clinical Epidemiology and Biostatistics, Faculty of Medicine

Ramathibodi Hospital, Mahidol University, Thailand; kunlawat.tha@mahidol.ac.th
(Corresponding Author)

3 Division of Social and Administrative Pharmacy, Department of Pharmacy, Faculty of Pharmacy, Mahidol University, Thailand

ARTICLE HISTORY

Received: 10 April 2026

Revised: 24 April 2026

Published: 16 May 2026

ABSTRACT

Previous systematic reviews and meta-analyses (SRMA) have assessed individual cervical cancer screening tests, but none has comprehensively assessed all major modalities using harmonized accuracy metrics. The objective of this research was to pool the diagnostic performance of visual inspection with acetic acid (VIA), visual inspection with Lugol's iodine (VILI), cytology, human papillomavirus (HPV) DNA (clinician-collected and self-sampling), and HPV mRNA testing for detecting cervical intraepithelial neoplasia grade 2 or worse (CIN2+) in primary screening. We combined an umbrella review of 27 SRMAs (570 primary studies, >1.15 million women) with an updated meta-analysis of 34 individual studies using colposcopy-directed biopsy as the reference standard. Pooled sensitivity (Sn), specificity (Sp), positive and negative likelihood ratios (LR+ and LR-, respectively), and diagnostic odds ratios (DOR) were estimated using bivariate hierarchical summary receiver operating characteristic models. The results show that clinician-collected HPV DNA and mRNA testing had strong performance (HPV DNA: Sn 0.88, Sp 0.80, LR- 0.15, DOR 30; HPV mRNA: Sn 0.86, Sp 0.72, LR- 0.19, DOR 16), but the latter was based on only four studies. VILI appeared comparable (Sn 0.86, Sp 0.89, LR- 0.16, DOR 47) but VIA did not (Sn 0.71, Sp 0.87, LR- 0.33, DOR 17). Weaker accuracy was observed with self-sampled HPV DNA testing. Performance of cytology varied by threshold. In conclusion, clinician-collected HPV DNA testing showed the strongest overall accuracy for CIN2+ detection, supporting WHO recommendations. VILI is a viable alternative in resource-limited settings. HPV mRNA and self-sampling approaches need further evaluation.

Keywords: Cervical Cancer Screening, Diagnostic Accuracy, Umbrella Review, Meta-Analysis, HPV Testing, VIA, VILI, Cytology, CIN2+

CITATION INFORMATION: Tran, T. P., Thadanipon, K., Thakkestian, A., & Thavorncharoensap, M. (2026). Performance of Different Cervical Cancer Primary Screening Methods: An Umbrella Review and Updated Meta-Analysis. *Procedia of Multidisciplinary Research*, 4(5), 44.

INTRODUCTION

Cervical cancer remains a leading cause of cancer mortality among women worldwide, with approximately 341,831 deaths in 2020. Of these, nearly 90% occurred in low- and middle-income countries (LMICs), which accounted for 84% of new cases (Arbyn et al., 2020; Sung et al., 2021). Several primary screening modalities are available, including visual inspection with acetic acid (VIA), visual inspection with Lugol's iodine (VILI), conventional cytology, liquid-based cytology (LBC), and human papillomavirus (HPV) testing (DNA- and mRNA-based, clinician- or self-collected). They differ in accuracy, cost, and infrastructure requirements (WHO, 2021).

Although several systematic reviews and meta-analyses (SRMA) have evaluated these tests individually (Arbyn et al., 2018; Mustafa et al., 2016; Smith et al., 2023), none has provided a consolidated assessment of all major modalities under a harmonized analytic framework. This study addresses that gap through a two-step approach: an umbrella review of existing SRMAs and an updated diagnostic test accuracy meta-analysis on the performance of all main screening modalities for detecting cervical intraepithelial neoplasia grade 2 or worse (CIN2+).

LITERATURE REVIEWS

Several SRMAs have assessed individual cervical cancer screening tests but each has covered only a limited subset. Arbyn et al. (2014, 2018) compared HPV testing on self- versus clinician-collected samples but did not evaluate visual inspection tests. Mustafa et al. (2016) included HPV testing, cytology and VIA but accepted colposcopy with or without biopsy as the reference standard, risking verification bias. Smith et al. (2023) evaluated VIA, VILI, HPV testing, and conventional cytology but not LBC. More recent reviews (Downham et al., 2024; Karisani et al., 2024; Terasawa et al., 2022) covered additional methods, but no single SRMA has consolidated all major modalities (i.e., VIA, VILI, conventional cytology, LBC, HPV DNA [clinician- and self-collected], and HPV mRNA) under uniform inclusion criteria and a consistent analytic framework. This gap limits the availability of comprehensive evidence to inform screening choices in different settings.

RESEARCH METHODOLOGY

This study combined an umbrella review with an updated diagnostic test accuracy meta-analysis, following the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) guidelines and the umbrella review methodology of Aromataris et al. (2015). PubMed and Scopus were searched from inception to September 2024, combining terms for cervical screening, diagnostic accuracy, CIN outcomes, and SRMA filters. For primary studies, the same base search strategy was used but without the SRMA filter, using publication-year filters based on the most recent corresponding SRMA for each primary screening modality.

Eligible SRMAs were those that evaluated diagnostic performance of VIA, VILI, cytology, or HPV testing (i.e., index test) in apparently healthy women with intact cervixes (i.e., population), used colposcopy-directed or random biopsy applied to all women as the reference standard, and reported CIN2+ as an outcome. Eligible primary studies were cross-sectional diagnostic accuracy studies that met the same criteria for population, index test, reference standard, and outcome as the above, with index-to-reference interval less than 12 weeks. SRMAs or studies restricted to HIV-positive or pregnant populations, or reporting only co-testing/triage strategies, were excluded.

Overlap among SRMAs was quantified using the corrected covered area (Pieper et al., 2014). Methodological quality of SRMAs and primary studies was appraised using AMSTAR-2 (Shea et al., 2017) and QUADAS-2 (Whiting et al., 2011), respectively. Pooled sensitivity, specificity, positive and negative likelihood ratios (LR+ and LR-, respectively), and diagnostic odds ratios (DOR) were estimated with the MIDAS module in Stata (Dwamena, 2007), which

fits hierarchical bivariate random-effects hierarchical summary receiver operating characteristic models. For index tests with insufficient data to support stable bivariate estimation, study-level Sn, Sp, LR+, LR-, and DOR are reported instead of pooled estimates. Heterogeneity was quantified using the I^2 statistic, interpreted as low (< 25%), moderate (25-50%), substantial (50-75%), or considerable (> 75%) following Higgins et al. (2003).

RESEARCH RESULTS

From 1,388 records identified, 27 SRMAs met the eligibility criteria after screening (Figure 1), encompassing 570 primary studies and over 1.15 million women, published between 2000 and 2024.

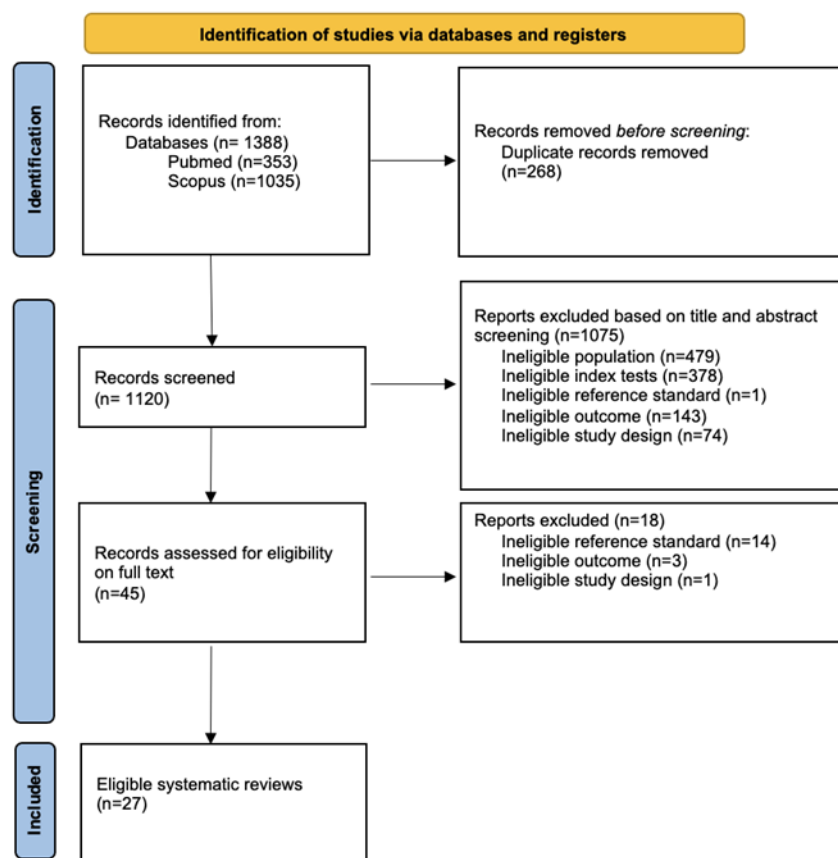


Figure 1 PRISMA 2020 flow diagram for selection of systematic reviews

HPV DNA and HPV mRNA tests were the most frequently assessed modalities (clinician-collected HPV DNA, $n = 16$; self-sampled HPV DNA, $n = 7$; clinician-collected HPV mRNA, $n = 6$), followed by VIA ($n = 11$), VILI ($n = 8$), cytology ($n = 8$), and LBC ($n = 4$). All SRMAs used CIN2+ as the diagnostic threshold. The corrected covered area was 2.33% (slight overlap), and AMSTAR-2 ratings were predominantly low or critically low.

A total of 34 unique primary studies were included: 26 retained from the SRMAs and 10 newly identified (Figure 2). Most were conducted in LMICs as population-based screening programs, with mean age of 35-40 years.

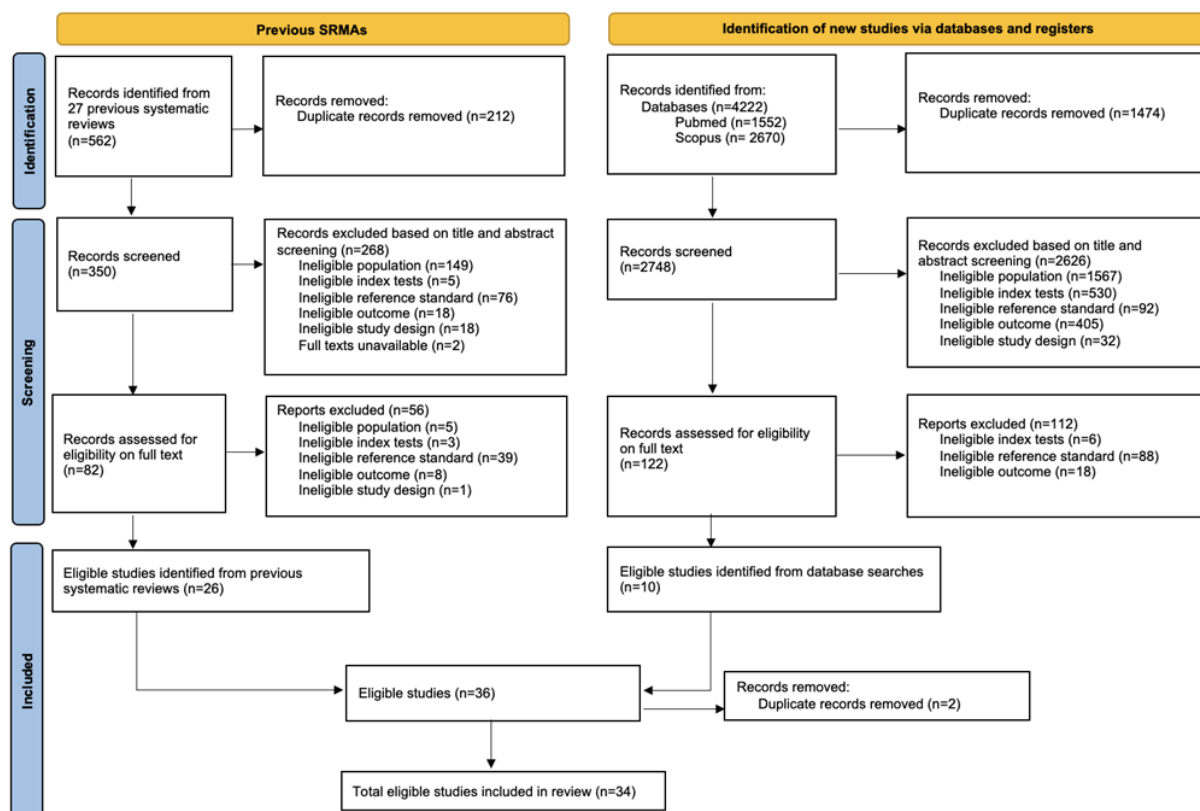


Figure 2 PRISMA 2020 flow diagram for individual primary study selection

Thirty-three studies used colposcopy-directed biopsy and one study (Thapa et al., 2018) used random biopsy, both applied to all screened women, as the reference standard. QUADAS-2 showed low risk of bias in patient selection (79.4%), index test (67.6%), and reference standard (73.5%) domains, but high risk in flow-and-timing (50%), mainly due to concerns regarding incomplete patient flow, including unexplained exclusions from the analysis and unclear blinding of assessors.

Pooled diagnostic accuracy estimates by screening method are shown in Table 1. Clinician-collected HPV DNA achieved high sensitivity (0.88) and a low LR- (0.15), with specificity of 0.80 and a DOR of 30. VILI also showed high sensitivity (0.86), specificity (0.89), and DOR (47), along with a low LR- (0.16). VIA had a moderate LR- (0.33) and sensitivity (0.71). Conventional cytology and LBC showed the expected threshold trade-off, with high-grade squamous intraepithelial lesion or worse (HSIL+) reaching specificity of 0.98-0.99 and very high DORs but low sensitivity. Low performance was observed with self-sampling HPV DNA (sensitivity, 0.62; LR-, 0.45; and DOR, 9), while clinician-collected HPV mRNA achieved high sensitivity (0.86) as well as a low LR- (0.19), but lower specificity (0.72). Substantial to considerable between-study heterogeneity was observed for most modalities ($I^2 > 75\%$ for VIA, conventional cytology subgroups, clinician-collected HPV DNA, and clinician-collected HPV mRNA; $I^2 = 51\%$ for self-sampled HPV DNA; see Table 1). Only three studies were available for the low-grade squamous intraepithelial lesion or worse (LSIL+) and HSIL+ thresholds of LBC (Belinson et al., 2001; Belinson et al., 2002; Hovland et al., 2010); thus, pooled bivariate estimation could not be performed due to model non-convergence, and the corresponding study-level estimates are reported in Table 1

Table 1 Pooled diagnostic accuracy estimates by screening method for CIN2+ detection

Screening method	N	Sn (95% CI)	Sp (95% CI)	LR+ (95% CI)	LR- (95% CI)	DOR (95% CI)	I ² (%)
VIA	22	0.71 (0.62-0.78)	0.87 (0.80-0.92)	5.6 (3.6-8.8)	0.33 (0.25-0.44)	17 (9-30)	100
VILI	7	0.86 (0.70-0.94)	0.89 (0.80-0.94)	7.6 (3.9-14.8)	0.16 (0.07-0.37)	47 (13-172)	97
ConvCytol	11	0.70 (0.55-0.81)	0.91 (0.85-0.94)	7.4 (4.9-11.2)	0.34 (0.23-0.50)	22 (13-37)	100
ASCUS+							
LBC ASCUS+	5	0.85 (0.69-0.94)	0.80 (0.62-0.91)	4.3 (2.0-9.3)	0.19 (0.08-0.44)	23 (6-94)	92
ConvCytol LSIL+	13	0.61 (0.50-0.72)	0.95 (0.93-0.97)	12.7 (8.6-18.7)	0.41 (0.31-0.54)	31 (18-55)	98
LBC LSIL+	3	0.60-0.87*	0.94-0.98*	13.5-24.5*	0.14-0.41*	60-106*	—
ConvCytol HSIL+	7	0.57 (0.42-0.71)	0.99 (0.97-0.99)	39.7 (25.0-63.2)	0.44 (0.31-0.61)	91 (62-133)	98
LBC HSIL+	3	0.33-0.77*	0.98-0.99*	34.9-47.7*	0.24-0.67*	71-163*	—
HPV DNA CC	12	0.88 (0.74-0.95)	0.80 (0.67-0.89)	4.4 (2.7-7.3)	0.15 (0.06-0.33)	30 (13-74)	100
HPV DNA SS	4	0.62 (0.41-0.80)	0.84 (0.78-0.88)	3.9 (2.3-6.6)	0.45 (0.25-0.80)	9 (3-25)	51
HPV mRNA CC	4	0.86 (0.79-0.91)	0.72 (0.47-0.88)	3.1 (1.4-6.6)	0.19 (0.11-0.32)	16 (5-51)	76

N = number of studies; ASCUS+ = atypical squamous cells of undetermined significance or worse; CC = clinician-collected; ConvCytol = conventional cytology; DOR = diagnostic odds ratio; HSIL+ = high-grade squamous intraepithelial lesion or worse; LBC = liquid-based cytology; LR+ = positive likelihood ratio; LR- = negative likelihood ratio; LBC = liquid-based cytology; LSIL+ = low-grade squamous intraepithelial lesion or worse; Sn = sensitivity; Sp = specificity; SS = self-sampling; VIA = visual inspection with acetic acid; VILI = visual inspection with Lugol's iodine.

*Range across 3 individual studies (Belinson et al., 2001; Belinson et al., 2002; Hovland et al., 2010).

DISCUSSION & CONCLUSION

This umbrella review integrated 27 SRMAs (570 primary studies, >1.15 million women) and 34 individual primary studies to evaluate eight cervical cancer screening modalities under a harmonized framework. A consistent performance gradient emerged: clinician-collected HPV DNA and VILI provided the best balance of sensitivity and specificity, VIA performed moderately, and cytology showed the expected threshold-related trade-off, with HSIL+ delivering very high specificity but lower sensitivity. Self-sampling HPV DNA had relatively weak overall performance, while clinician-collected HPV mRNA achieved high sensitivity with strong rule-out capability but lower specificity.

Our findings are broadly consistent with earlier SRMAs (Arbyn et al., 2018, 2022; Smith et al., 2023) and reinforce clinician-collected HPV DNA as the preferred primary screening test where resources allow. The high accuracy of VILI corroborates earlier LMIC-focused work (Driscoll et al., 2018; Fokom-Domgue et al., 2015). The pooled HPV mRNA estimates represent the first hierarchical summary receiver operating characteristic meta-analysis combining multiple mRNA studies under a uniform framework, though the wide specificity confidence intervals reflect the limited number of studies on this relatively new test.

From a programmatic perspective, HPV DNA testing requires laboratory infrastructure and trained personnel, with per-test costs of approximately US\$5-15 in LMICs versus US\$1-5 for the visual inspection methods. Among the latter, VILI represents an attractive option for resource-limited settings. It requires minimal infrastructure and enables screen-and-treat strategies, yet in our analysis it achieved favorable sensitivity (0.86) and LR- (0.16), indicative of strong rule-out capability approaching that of clinician-collected HPV DNA. VIA remains a fallback with lower and more variable performance. Policymakers should balance diagnostic accuracy with cost, infrastructure, and population coverage (WHO, 2021).

Strengths of this study include the combination of an umbrella review with an updated meta-analysis, systematic quality appraisal using AMSTAR-2 and QUADAS-2, and the wide range of modalities and settings examined. Limitations include substantial heterogeneity for visual and cytology tests, the limited evidence base for LBC and self-sampling HPV DNA and HPV

mRNA assays, the rarity of within-population multi-test comparisons, and the absence of stratified reporting by age, HIV status, or screening history.

In conclusion, clinician-collected HPV DNA testing showed strong overall diagnostic accuracy for CIN2+ detection, supporting current WHO recommendations for HPV-based primary screening. VILI represents a viable visual alternative in resource-limited settings, while cytology at the ASCUS+ cut-off offers the most balanced performance for triage where cytology infrastructure is available. Further studies on HPV mRNA assays and self-sampling approaches are still needed, alongside head-to-head studies and stratified analyses to guide context-appropriate screening strategies.

REFERENCES

- Arbyn, M., Simon, M., de Sanjosé, S., Clarke, M. A., Poljak, M., Rezhake, R., Berkhof, J., Nyaga, V., Gultekin, M., Canfell, K., & Wentzensen, N. (2022). Accuracy and effectiveness of HPV mRNA testing in cervical cancer screening: A systematic review and meta-analysis. *The Lancet Oncology*, 23(7), 950-960.
- Arbyn, M., Smith, S. B., Temin, S., Sultana, F., Castle, P., & Collaboration on Self-Sampling and HPV Testing. (2018). Detecting cervical precancer and reaching underscreened women by using HPV testing on self samples: Updated meta-analyses. *BMJ*, 363, k4823.
- Arbyn, M., Verdoodt, F., Snijders, P. J., Verhoef, V. M., Suonio, E., Dillner, L., Minozzi, S., Bellisario, C., Banzi, R., Zhao, F. H., Hillemanns, P., & Anttila, A. (2014). Accuracy of human papillomavirus testing on self-collected versus clinician-collected samples: A meta-analysis. *The Lancet Oncology*, 15(2), 172-183.
- Arbyn, M., Weiderpass, E., Bruni, L., de Sanjosé, S., Saraiya, M., Ferlay, J., & Bray, F. (2020). Estimates of incidence and mortality of cervical cancer in 2018: A worldwide analysis. *The Lancet Global Health*, 8(2), e191-e203.
- Aromataris, E., Fernandez, R., Godfrey, C. M., Holly, C., Khalil, H., & Tungpunkom, P. (2015). Summarizing systematic reviews: Methodological development, conduct and reporting of an umbrella review approach. *International Journal of Evidence-Based Healthcare*, 13(3), 132-140.
- Belinson, J. L., Pan, Q.-J., Biscotti, C., Wu, L.-Y., Pretorius, R. G., Li, L., Elson, P., Rong, S.-D., Zhang, W.-H., & Qiao, Y.-L. (2002). Primary screening with liquid-based cytology in an unscreened population in rural China, with an emphasis on reprocessing unsatisfactory samples. *Acta Cytologica*, 46(3), 470-474.
- Belinson, J. L., Pretorius, R. G., Zhang, W. H., Wu, L. Y., Qiao, Y. L., & Elson, P. (2001). Cervical cancer screening by simple visual inspection after acetic acid. *Obstetrics and Gynecology*, 98(3), 441-444.
- Downham, L., Jaafar, I., Rol, M. L., Nyawira Nyaga, V., Valls, J., Baena, A., Zhang, L., Gunter, M. J., Arbyn, M., & Almonte, M. (2024). Accuracy of HPV E6/E7 oncoprotein tests to detect high-grade cervical lesions: A systematic literature review and meta-analysis. *British Journal of Cancer*, 130(4), 517-525.
- Driscoll, S. D., Tappen, R. M., Newman, D., & Voege-Harvey, K. (2018). Accuracy of visual inspection performed by community health workers in cervical cancer screening. *International Journal of Gynecology and Obstetrics*, 142(3), 260-269.
- Dwamena, B. (2007). *MIDAS: Stata module for meta-analytical integration of diagnostic test accuracy studies*. Statistical Software Components, Boston College Department of Economics.
- Fokom-Domgue, J., Combescure, C., Fokom-Defo, V., Tebeu, P. M., Vassilakos, P., Kengne, A. P., & Petignat, P. (2015). Performance of alternative strategies for primary cervical

- cancer screening in sub-Saharan Africa: Systematic review and meta-analysis of diagnostic test accuracy studies. *BMJ*, 351, h3084.
- Higgins, J. P. T., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *BMJ*, 327(7414), 557-560.
- Hovland, S., Arbyn, M., Lie, A. K., Ryd, W., Borge, B., Berle, E. J., Skomedal, H., Kadima, T. M., Kyembwa, L., Billay, E. M., Mukwege, D., Chirimwami, R. B., Mvula, T. M., Snijders, P. J., Meijer, C. J. L. M., & Karlsen, F. (2010). A comprehensive evaluation of the accuracy of cervical pre-cancer detection methods in a high-risk area in East Congo. *British Journal of Cancer*, 102(6), 957-965.
- Karisani, N., Aminimoghaddam, S., Kashanian, M., Baradaran, H. R., & Moradi, Y. (2024). Diagnostic accuracy for alternative cervical cancer screening strategies: A systematic review and meta-analysis. *Health Care for Women International*, 45(3), 323-362.
- Mustafa, R. A., Santesso, N., Khatib, R., Mustafa, A. A., Wiercioch, W., Kehar, R., Gandhi, S., Chen, Y., Cheung, A., Hopkins, J., Ma, B., Lloyd, N., Wu, D., Broutet, N., & Schünemann, H. J. (2016). Systematic reviews and meta-analyses of the accuracy of HPV tests, visual inspection with acetic acid, cytology, and colposcopy. *International Journal of Gynaecology and Obstetrics*, 132(3), 259-265.
- Pieper, D., Antoine, S.-L., Mathes, T., Neugebauer, E. A. M., & Eikermann, M. (2014). Systematic review finds overlapping reviews were not mentioned in every other overview. *Journal of Clinical Epidemiology*, 67(4), 368-375.
- Shea, B. J., Reeves, B. C., Wells, G., Thuku, M., Hamel, C., Moran, J., Moher, D., Tugwell, P., Welch, V., Kristjansson, E., & Henry, D. A. (2017). AMSTAR 2: A critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. *BMJ*, 358, j4008.
- Smith, S. K., Nwosu, O., Edwards, A., Zerihun, M., Chung, M. H., Suvada, K., & Ali, M. K. (2023). Performance of screening tools for cervical neoplasia among women in low- and middle-income countries: A systematic review and meta-analysis. *PLOS Global Public Health*, 3(2), e0001598.
- Sung, H., Ferlay, J., Siegel, R. L., Laversanne, M., Soerjomataram, I., Jemal, A., & Bray, F. (2021). Global Cancer Statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: A Cancer Journal for Clinicians*, 71(3), 209-249.
- Terasawa, T., Hosono, S., Sasaki, S., Hoshi, K., Hamashima, Y., Katayama, T., & Hamashima, C. (2022). Comparative accuracy of cervical cancer screening strategies in healthy asymptomatic women: A systematic review and network meta-analysis. *Scientific Reports*, 12(1), 94.
- Thapa, N., Maharjan, M., Shrestha, G., Maharjan, N., Lindell, D., Zuo, N., Yang, J., Maskey, N., & Cai, H. (2018). Accuracy of Visual Tests for Primary Cervical Cancer Screening in Rural Nepal. *JNMA: Journal of the Nepal Medical Association*, 56(214), 917-923.
- Whiting, P. F., Rutjes, A. W. S., Westwood, M. E., Mallett, S., Deeks, J. J., Reitsma, J. B., Leeflang, M. M. G., Sterne, J. A. C., Bossuyt, P. M. M., & QUADAS-2 Group. (2011). QUADAS-2: A revised tool for the quality assessment of diagnostic accuracy studies. *Annals of Internal Medicine*, 155(8), 529-536.
- WHO. (2021). *WHO guideline for screening and treatment of cervical pre-cancer lesions for cervical cancer prevention*. World Health Organization, Geneva.

Data Availability Statement: The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Conflicts of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.



Copyright: © 2026 by the authors. This is a fully open-access article distributed under the terms of the Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0).