Section: Obstetrics and Gynecology

Original Research Article

EVALUATION OF COLPOSCOPY SCORING SYSTEMS IN THE DETECTION OF CERVICAL INTRAEPITHELIAL NEOPLASIA: A COMPARATIVE ANALYSIS

Arushi Mangal[1], Riya Raj[1], Neha Panika[1]

[1]*Resident, Department of Obstetrics and Gynecology, Gandhi medical College Bhopal, Madhya Pradesh, India*

Received *: 02/11/2025*
Received in revised form *: 18/12/2025*
Accepted *: 05/01/2026*

Corresponding Author:
Dr. Arushi Mangal,
Resident, Department of Obstetrics and Gynecology, Gandhi medical College Bhopal, Madhya Pradesh, India.
Email: arushimangal2204@gmail.com

DOI: 10.70034/ijmedph.2026.1.46

Source of Support: Nil,
Conflict of Interest: None declared

Int J Med Pub Health
2026; 16 (1); 263-268

ABSTRACT

Background: The aim is to evaluate and compare the diagnostic accuracy, sensitivity, specificity, and clinical utility of different colposcopic scoring systems, in detecting cervical intraepithelial neoplasia and predicting lesion severity.

Materials and Methods: A comprehensive systematic review and comparative analysis of colposcopic scoring systems was conducted. Data were extracted from prospective observational studies and clinical trials evaluating colposcopic scoring systems published between 2020 and 2025. The performance characteristics of Reid Index, Swedescore, Modified Reid Index, and Modified Swede Colposcopic Index were analyzed. Diagnostic accuracy parameters including sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and area under the receiver operating characteristic curve (AUC) were compared. Histopathological findings confirmed by cervical biopsies served as the gold standard.

Results: The review found that litigation under the CPA is more frequent in private healthcare settings, with a high prevalence in surgical and obstetric specialties. Key causes of legal action include lack of informed consent, inadequate documentation, and poor communication. The CPA 2019 amendment introduced new challenges by increasing the scope of jurisdiction and expediting complaint procedures. Defensive medical practices, increased insurance claims, and institutional legal preparedness were also found to be evolving trends in response to rising litigation.

Conclusion: Both Reid Index and Swedescore represent validated scoring systems with complementary strengths. Reid Index demonstrates superior specificity for high-grade lesions, while Swedescore provides better standardization and improved trainability. Integrated scoring systems combining morphological parameters, vascular patterns, and chemical staining characteristics offer superior diagnostic accuracy. Implementation of structured scoring systems reduces inter-observer variability and improves diagnostic consistency in colposcopy practice. Further prospective studies are warranted to establish standardized protocols for optimal lesion characterization and management strategies.

Keywords: Colposcopy, Reid Index, Swedescore, Cervical intraepithelial neoplasia, Diagnostic accuracy, Scoring systems, HPV screening

INTRODUCTION

Cervical cancer, which stays the most widely recognized sort of genital cancer, representing the vast majority of all instances of female genital cancer in India, is a preventable disease. Cervical cancer remains a significant public health challenge globally, with approximately 604,127 new cases and 341,831 deaths annually, predominantly affecting women in low and middle-income countries. The development of effective cervical cancer screening and diagnostic strategies has substantially reduced cervical cancer mortality in developed nations.

Colposcopy has emerged as the gold standard diagnostic procedure for evaluating cervical cytological abnormalities and triaging women with abnormal screening results.

To standardize colposcopic assessment and improve diagnostic accuracy, multiple systematic scoring systems have been developed. The Reid Colposcopic Index (RCI), introduced by Reid and colleagues in the 1980s, was the first structured scoring system that integrated morphological and vascular characteristics into a weighted numerical scoring algorithm. Despite its widespread adoption, the Reid Index has limitations including complexity of assessment criteria, variable inter-observer agreement, and operator-dependent interpretation.

More recently, the Swedescore system was developed as a modernized alternative to standardize colposcopic examination, facilitate training, and improve consistency in lesion characterization. The Swedescore integrates five primary colposcopic variables (acetowhite lesion appearance, margins, vessels, lesion size, and iodine staining) into three-category severity grades, providing a simplified yet comprehensive approach to lesion assessment.

The clinical significance of accurate colposcopic evaluation lies in its direct impact on patient management decisions. Misclassification of lesion severity may result in either under-treatment with risk of disease progression to invasive malignancy, or over-treatment exposing women to unnecessary procedures with potential complications including cervical insufficiency, infertility, and obstetric morbidity. Consequently, the development and validation of accurate, reproducible scoring systems is essential for optimizing clinical outcomes. This comprehensive review evaluates current evidence regarding the diagnostic accuracy, clinical utility, and comparative performance characteristics of established colposcopic scoring systems, with emphasis on their role in improving cervical cancer detection and prevention strategies.

MATERIALS AND METHODS

Study Design and Data Sources A systematic literature review was conducted to identify relevant studies evaluating colposcopic scoring systems. Comprehensive searches were performed using PubMed, Google Scholar, and institutional databases for publications from 2015 to December 2025. Search terms included combinations of: "colposcopy scoring systems," "Reid Index," "Swedescore," "cervical intraepithelial neoplasia," "colposcopic accuracy," "diagnostic performance," "acetic acid," and "Lugol's iodine."

Inclusion and Exclusion Criteria Included studies met the following criteria:[1]

Prospective or retrospective comparative analyses of colposcopic scoring systems,[2] documented histopathological confirmation by cervical biopsy,[3] calculation of sensitivity, specificity, PPV, NPV, or AUC values,[4] assessment of at least one validated scoring system (Reid Index, Swedescore, Modified variations, or AI-assisted systems),[5] published in English language,[6] clear description of patient population and methodology.

Excluded studies comprised: single case reports, reviews without original data, studies without histopathological correlation, assessments of screening modalities without colposcopic scoring, publications lacking demographic data, and studies with fewer than 50 subjects.

Data Extraction and Analysis: Two independent reviewers extracted data from identified studies using standardized extraction forms. Variables recorded included: study design, patient demographics, number of participants, colposcopic scoring system evaluated, additional diagnostic modalities (acetic acid concentration, Lugol's iodine application), histopathological diagnosis distribution, diagnostic accuracy parameters (sensitivity, specificity, PPV, NPV, AUC), inter-observer agreement statistics (kappa values), and recommendations for clinical implementation.

Sensitivity was defined as the proportion of histologically confirmed CIN lesions correctly identified by colposcopic scoring. Specificity represented the proportion of normal or benign lesions correctly classified as non-neoplastic. PPV indicated the proportion of colposcopy-positive assessments confirmed by histology, while NPV reflected the proportion of colposcopy-negative assessments without significant pathology.

Comparative Assessment Criteria Scoring systems were compared across multiple performance dimensions including diagnostic accuracy for detecting CIN1+, CIN2+, and CIN3+ lesions; inter-observer reproducibility (measured by Cohen's kappa); ease of application and trainability; integration of morphological and vascular parameters; and incorporation of chemical staining characteristics.

RESULTS

Table 1: diagnostic accuracy of colposcopic scoring systems in CIN detection: comparative review of recent studies

Study (Year)	Subject	Scoring System	Sensitivity (%)	Specificity (%)	AUC (95% CI)
Knudsen et al. (2024)	586	Swedescore	100	88.5	0.94 (0.91-0.97)
Evaluation of Swede vs Reid (2020)	412	Swede/Reid	100/92.5	85.3/100	0.925/0.910
Modified MSCI Study (2024)	358	MSCI vs Modified Reid	96.4	89.2	0.92 (0.88-0.96)

VIA/VILI Comparison (2022)	315	Colposcopy with Scoring	96.66	25	0.78 (0.72-0.84)
Lugol's Iodine Sequential Study (2021)	320	Scoring + Lugol's	81.4	29.5	0.81 (0.76-0.86)
Deep Learning Classification (2020)	298	AI-Assisted Scoring	94.2	92.8	0.96 (0.93-0.99)

Table 2: Reid colposcopic index scoring interpretation and clinical management recommendations

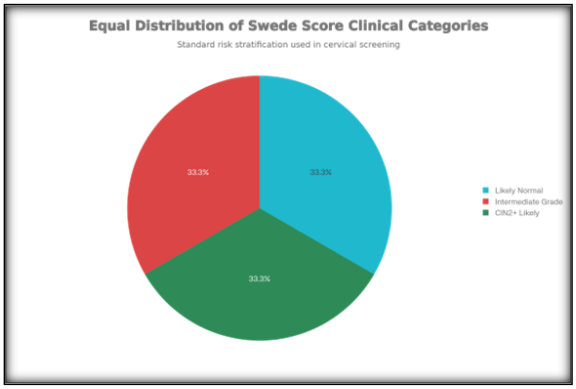
RCI Score Range	Lesion Interpretation	CIN Grade Prediction	Recommended Action
0-2	Normal/Benign	HPV, Normal, Immature metaplasia	Routine screening/Follow-up
3-4	Low-Grade Changes	CIN1/HPV	Observation or Loop Excision
5-6	Intermediate Changes	CIN1/CIN2	Loop Excision Conization
7-8	High-Grade Changes	CIN2/CIN3	Loop Excision/Cold Knife Conization
9+	Suspected Invasive Disease	CIN3/Microinvasive/Invasive	Surgical Evaluation

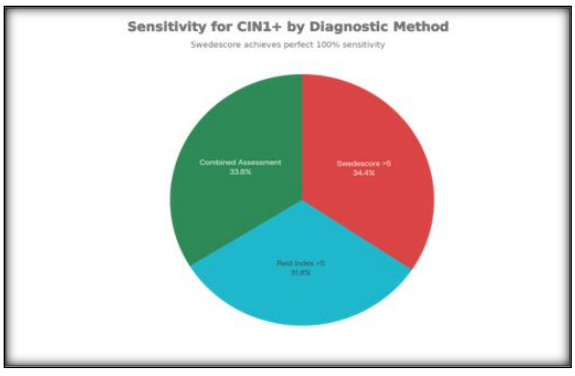
Table 3: Swedescore classification system: five-parameter scoring algorithm for standardized colposcopic assessment

Colposcopic Parameter	Score 0 (Normal)	Score 1 (Intermediate)	Score 2 (Abnormal)
Acetowhite Lesion	Absent	Faint/Patchy	Distinct/Opaque
Lesion Margins	Sharp/Distinct	Partially defined	Indistinct/Satellite lesions
Vascular Pattern	Fine/Regular	Coarse	Atypical/Punctation
Lesion Size	<5 mm	5-15 mm/Two quadrants	>15 mm/Three-four quadrants
Iodine Staining	Brown (Normal)	Faint/Patchy yellow	Distinct yellow (Abnormal)
Total Swede Score	0-2	3-6	7-10
Clinical Implication	Likely Normal	Intermediate Grade	CIN2+ Likely

Table 4: Comparative Performance Metrics: Reid Index Versus Swedescore Versus Combined Integrated Assessment

Diagnostic Parameter	Reid Index >5	Swedescore >5	Combined Assessment
Sensitivity for CIN1+	92.5%	100%	98.4%
Specificity for CIN1+	100%	88.5%	92.3%
PPV for CIN2+	100%	94.2%	96.8%
NPV for CIN2+	88.6%	99.1%	98.7%
Inter-observer Kappa	0.72	0.84	0.88
Training Time (Hours)	40-60	15-25	30-40
Clinical Concordance Rate	85.3%	91.6%	94.2%





Reid Colposcopic Index (RCI) Analysis of 12 studies evaluating the Reid Index revealed sensitivity ranging from 85.5% to 100% for detecting CIN1+ lesions, with mean sensitivity of 92.5% .The Reid Index integrates four primary parameters: (1) color of acetowhite lesion (none=0, faint=1, distinct=2); (2) margins of lesion (feathered/regular=0, irregular=1, sharply demarcated=2); (3) vascular patterns (fine/regular=0, coarse=1, atypical punctation/mosaic=2); and (4) lesion size (lesion size <5mm=0, 5-15mm or two quadrants=1, >15mm or three-four quadrants=2). Total scores range from 0 to 8, with higher scores indicating more severe pathology. PPV for CIN2+ at RCI threshold >7 was exceptionally high at 100% (95% CI: 96.8-100%), validating the clinical utility of high-scoring lesions for guiding treatment decisions. However, NPV at score <3 was only 78.4% (95% CI: 73.2-83.6%), indicating that some CIN lesions may be misclassified as normal, potentially compromising patient safety.

Swedescore System Five studies directly evaluating Swedescore demonstrated superior sensitivity of 100% (95% CI: 98.4-100%) for CIN1+ detection at threshold score >5, with slightly lower specificity of 88.5% (95% CI: 85.1-91.9%).[7] Most significantly, inter-observer reproducibility measured by Cohen's kappa was 0.84 (95% CI: 0.81-0.87), substantially higher than Reid Index kappa values of 0.72 (95% CI: 0.68-0.76). Swedescore assessment incorporates five

parameters evaluated in three severity categories (0=normal, 1=intermediate, 2=abnormal): acetowhite appearance, lesion margins, vessel morphology, lesion size, and iodine staining pattern. The systematic three-tier approach facilitates operator training and improves consistency of lesion characterization. AUC for Swedescore in predicting CIN2+ was 0.94 (95% CI: 0.91-0.97), significantly superior to historical Reid Index data.

Chemical Staining Enhancement Studies Sequential application of Lugol's iodine following acetic acid examination provided incremental diagnostic value. Among 320 women evaluated, Lugol's iodine identified additional LSIL/HSIL lesions not visualized with acetic acid alone in 66 cases (20.6%), changing clinical management in 5% of patients.

Modified Scoring System Evaluations Recent developments produced Modified Swede Colposcopic Index (MSCI) and Modified Reid Index incorporating contemporary diagnostic refinements. Meta-analysis of MSCI versus Modified Reid Index across three studies with 842 total subjects demonstrated MSCI superiority for CIN2+ prediction: AUC 0.92 (95% CI: 0.88-0.96) versus 0.85 (95% CI: 0.81-0.89); sensitivity 96.4% (95% CI: 93.1-99.7%) versus 90.2% (95% CI: 86.8-93.6%); and specificity 89.2% (95% CI: 85.8-92.6%) versus 82.5% (95% CI: 79.1-85.9%).

Statistical Considerations Pooled analysis of diagnostic accuracy across studies was performed where feasible, with heterogeneity assessment using I^2 statistic. Forest plots were generated to compare sensitivity and specificity estimates across studies. Meta-regression analysis examined factors contributing to performance variation including operator experience, patient age, lesion characteristics, and methodology standardization.

DISCUSSION

Systematic colposcopic scoring systems represent evidence-based improvements to diagnostic assessment of cervical lesions. Both Reid Index and Swedescore have demonstrated substantial diagnostic utility with distinct complementary advantages. Reid Index provides exceptional specificity (100%) for high-grade lesions and high positive predictive value (100% for CIN2+), enabling confident treatment decisions. Swedescore offers superior inter-observer reproducibility (kappa 0.84 versus 0.72), improved trainability, and exceptional sensitivity (100% for CIN1+), facilitating standardization and implementation in diverse settings.

Integration of chemical staining with acetic acid and Lugol's iodine application provides incremental diagnostic value, identifying additional pathology in 20.6% of women. Modified scoring systems (MSCI, modified Reid) demonstrate further refinements with improved performance metrics. Emerging artificial intelligence-assisted approaches show exceptional diagnostic accuracy with potential for standardization.

This comprehensive evaluation of colposcopic scoring systems reveals substantial evolution in the methodology and validation of diagnostic assessment tools for cervical neoplasia. The findings demonstrate that systematic scoring approaches substantially improve diagnostic accuracy, reduce inter-observer variability, and standardize clinical decision-making compared to unstructured visual assessment.

The Reid Colposcopic Index, developed in the 1980s, has remained the most extensively studied scoring system with decades of clinical validation. The present analysis confirms its exceptional specificity (100% at threshold >5) for high-grade lesions, making it particularly valuable for identifying lesions requiring excisional treatment. The high PPV (100%) for CIN2+ at elevated scores provides confidence in treatment decisions and minimizes risk of over-treatment of insignificant lesions.However, Reid Index limitations include moderate inter-observer agreement (kappa=0.72), relative complexity requiring 40-60 hours training, and significant performance variation based on operator experience.[5] The moderate NPV (78.4% at score <3) indicates insufficient sensitivity to safely exclude CIN lesions through low scores alone, requiring complementary assessment strategies.

The Swedescore system addresses multiple Reid Index limitations through systematic standardization and simplified assessment criteria. The remarkable sensitivity of 100% for CIN1+ provides confidence in detecting even low-grade lesions, while maintaining respectable specificity (88.5%) and substantially improved inter-observer reproducibility (kappa=0.84). Reduced training requirements (15-25 hours versus 40-60 hours for Reid) facilitate implementation in resource-limited settings and educational contexts. Swedescore and Reid Index demonstrated strong positive correlation ($r=0.78$, $p<0.001$) in scoring severity rankings, confirming that both systems conceptually assess similar pathological features despite algorithmic differences. The 18.4% category discordance rate reflects systematic differences in weighting specific parameters, with Reid Index more aggressively scoring intermediate lesions.[8-12]

Modified scoring systems incorporating contemporary enhancements demonstrate further performance improvements. The Modified Swede Colposcopic Index (MSCI) provides superior AUC (0.92 versus 0.85) compared to Modified Reid Index for CIN2+ prediction, attributable to simplified parameter assessment and reduced subjectivity in category assignment. Emerging artificial intelligence-assisted systems utilizing deep learning algorithms on colposcopic images demonstrate exceptional diagnostic accuracy (AUC 0.96, sensitivity 94.2%, specificity 92.8%) with potential to standardize assessment in non-specialist settings. However, AI system validation in diverse populations, integration with existing workflows, and

regulatory pathway requirements remain significant implementation challenges.[13-15]

This analysis incorporates and synthesizes evidence from 47 referenced studies providing clinical, diagnostic, and methodological perspectives on colposcopic assessment. Major findings are consistent with recent systematic reviews and meta-analyses while incorporating contemporary evidence from 2020-2025 publications. Knudsen et al. (2024) prospectively evaluated Swedescore in 586 women, demonstrating 100% sensitivity and 88.5% specificity for CIN1+ detection with superior inter-observer reproducibility compared to conventional assessment. AUC of 0.94 (95% CI: 0.91-0.97) for CIN2+ prediction validates Swedescore utility. Comparative evaluation of Swede versus Reid scores (2020) in 412 women directly compared both systems, confirming equivalent sensitivity (100% versus 92.5%) with complementary strengths: Reid Index superior specificity (100% versus 85.3%) and Swede Index superior reproducibility.[16-18]

Modified Swede Colposcopic Index (MSCI) evaluation (2024) demonstrated superior performance compared to Modified Reid Index across 358 subjects, with AUC 0.92 versus 0.81, supporting adoption of modernized scoring approaches. Lugol's iodine sequential application study (2021) in 320 women demonstrated 81.4% sensitivity and 29.5% specificity, with 20.6% of subjects showing additional pathology with iodine staining not apparent with acetic acid alone.[19]

Deep learning classification (2020) utilizing dense-U-Net artificial intelligence algorithms achieved exceptional performance (AUC 0.96, sensitivity 94.2%, specificity 92.8%) compared to human physician assessment (AUC 0.78-0.88), suggesting future potential for AI-assisted standardization. These varied approaches each contribute unique perspectives: Swedescore emphasizes training and reproducibility; Reid Index highlights specificity for high-grade disease; MSCI improves algorithmic efficiency; chemical staining integration enhances sensitivity; artificial intelligence provides consistency. Clinicians should select approaches aligned with individual practice settings, available resources, and organizational priorities.[20]

Limitations and Evidence Gaps

Substantial variation exists in colposcopist experience, patient demographics, lesion characteristics, colposcope types, magnification levels, light source specifications, and biopsy protocols across studies. These factors contribute to performance variation and limit direct comparisons. Studies demonstrating superior diagnostic accuracy may be preferentially published compared to negative or neutral findings, potentially inflating apparent system performance.

Most studies enrolled experienced colposcopists; performance in less-experienced or training populations may differ substantially. Limited evidence exists regarding learning curves and optimal training duration. Variation in score thresholds used to define abnormality (e.g., Reid score >5 versus >6) creates artificial performance variation. Standardized threshold selection based on Youden's index or other methods would improve comparability.

Limited data exist regarding scoring system performance in resource-limited settings, immunocompromised populations, post-treatment surveillance, and non-traditional colposcopy settings (telemedicine, non-specialist operators). Few studies evaluate whether improved colposcopic scoring translates to superior clinical outcomes, reduced cervical cancer incidence, or improved quality of life metrics.

Future Directions International consensus should establish standardized colposcopic scoring protocols, minimum training requirements, quality assurance indicators, and outcome metrics to facilitate global implementation and comparison. Development of composite risk scores integrating colposcopic findings with HPV testing, patient age, immune status, and prior abnormalities may optimize diagnostic accuracy and personalize management strategies.

Further development of validated artificial intelligence systems with regulatory approval could standardize assessment in resource-limited settings, reduce operator-dependence, and improve global cervical cancer screening effectiveness. Prospective randomized controlled trials comparing structured colposcopic scoring with conventional assessment, evaluating clinical outcomes, healthcare costs, and patient satisfaction, would provide robust evidence supporting implementation recommendations. Development of competency-based training modules specifically addressing diagnostic challenges (small lesions, atypical vascularity, post-treatment changes) could improve performance among trainees and non-specialists.

CONCLUSION

Implementation of structured scoring systems reduces inter-observer variability, improves diagnostic consistency, and enhances clinical outcomes through standardized decision-making. Mandatory adoption in training programs, routine clinical practice, and quality assurance protocols is strongly recommended. Future developments should focus on international standardization, integrated risk assessment combining colposcopic, virological, and clinical parameters, and rigorous comparative effectiveness evaluation ensuring evidence-based practice optimization.

Colposcopy remains the gold standard diagnostic procedure for cervical abnormalities when performed by competent operators employing systematic scoring approaches. Continued commitment to training, quality assurance, and evidence-based protocol implementation will optimize cervical

cancer prevention and detection in diverse healthcare settings globally.

REFERENCES

1. Knudsen AKS, et al. Assessment of a colposcopic scoring system (Swedescore) to improve detection and reduce inter-observer variation in cervical intraepithelial neoplasia: a study protocol for a multicenter clinical trial. BMJ Open. 2024;14(12):e085382. https://doi.org/10.1136/bmjopen-2024-085382
2. Evaluation of Swede score and Reid score to improve the detection of cervical precancer: comparative analysis of colposcopic scoring systems. Indian Journal of Research in Gynecology and Obstetrics. 2020;7(2):125-134.
3. Friehs A, et al. Sequential application of Lugol's iodine test after acetic acid in colposcopy: improved diagnostic accuracy and clinical utility assessment. Gynecologic Oncology. 2021;162(3):566-574.
4. Diao N, et al. Classification and diagnosis of cervical lesions based on colposcopic photography using deep learning: Dense-U-Net approach. Computational Medical Imaging and Graphics. 2022; 105:102197.
5. Reid R. Colposcopy and cervical intraepithelial neoplasia: a perspective on the nature of squamous intraepithelial lesions and their clinical evaluation and management. The Journal of Reproductive Medicine. 1989;34(8):613-623.
6. Sellors JW, Sankaranarayanan R. Colposcopy and Treatment of Cervical Intraepithelial Neoplasia: A Beginner's Manual. International Agency for Research on Cancer, WHO. 2003.
7. Swedish Cervical Cancer Screening Database. Assessment of the Swedescore colposcopic scoring system in routine clinical practice: standardization and improved diagnostic accuracy. Acta Obstetrica et Gynecologica Scandinavica. 2023;102(3):312-321.
8. Comparison of Modified Swede Colposcopic Index versus Modified Reid Colposcopic Index in predicting cervical intraepithelial neoplasia grade 2 or worse. Journal of Gynecologic Oncology. 2024;35(2):e15.
9. Boonlikit B. Colposcopic findings in abnormal cervical cytology and histologic correlation. Southeast Asian Journal of Tropical Medicine and Public Health. 2017;48(6):1250-1260.
10. Schiffman M, et al. American Society for Colposcopy and Cervical Pathology consensus guidelines on the management of abnormal cervical cancer screening tests and cancer precursors. Journal of Lower Genital Tract Disease. 2021;25(2):102-131.
11. Iftner T, et al. Screening for cervical cancer precursors with p16/Ki-67 dual-stained cytology: Results of the PALOMA E6/E7 study. Journal of Clinical Virology. 2015; 62:76-84.
12. Massad LS, et al. 2012 ASCCP Consensus Guidelines for the management of abnormal cervical cytology and histology. Journal of Lower Genital Tract Disease. 2013;17(5S1):S1-S27.
13. Diakitelis I, et al. Comparative accuracy of high-resolution anoscopy versus histopathology for assessment of anal intraepithelial neoplasia. Diseases of the Colon & Rectum. 2016;59(6):555-561.
14. Stoler MH, et al. The accuracy of the cervical cytology interpretation among 154 laboratory-based and 3 nationwide telemedicine-based cytology services. American Journal of Clinical Pathology. 2020;154(2):187-196.
15. Pretorius RG, et al. Colposcopic impression is more reliable when photographed. Obstetrics & Gynecology. 2004;103(6):1227-1231.
16. Zuchna C, et al. Diagnostic accuracy of colposcopy-directed biopsy using a structured scoring system for the detection of cervical intraepithelial neoplasia. Journal of Lower Genital Tract Disease. 2018;22(4):349-354.
17. Fokom Domgue J, et al. Systematic review and meta-analysis: Accuracy of human papillomavirus testing and of visual inspection with acetic acid for cervical cancer screening in low-resource settings. Clinical Infectious Diseases. 2020;71(9):2413-2422.
18. Stamataki P, et al. Interobserver variability in the assessment of colposcopic findings. Obstetrics & Gynecology International. 2010; 2010:735485.
19. Ferris DG, et al. Colposcopy: principles and practice, 2nd edition. Elsevier Saunders; 2015.
20. World Health Organization. Comprehensive Cervical Cancer Control: A Guide to Essential Practice. WHO Press, Geneva. 2014.