

High Grade Squamous Intraepithelial Lesion (HSIL) Positive Disease Risk Nomogram Model: A Single Centre Retrospective Analysis

Qiu-Zhen Wu¹, Mao-Hua Lin¹, Jian-Rui Zheng¹, Xiu-Qing Weng¹, Li-Li Zheng^{2,*}, Ying-Yu Mao^{1,*}

¹Department of Pathology, Mindong Hospital Affiliated to Fujian Medical University, Fuan, 355000, People's Republic of China; ²Department of Gynecology, Mindong Hospital Affiliated to Fujian Medical University, Fuan, 355000, People's Republic of China

*These authors contributed equally to this work

Correspondence: Ying-Yu Mao, Department of Pathology, Mindong Hospital Affiliated to Fujian Medical University, Fuan, 355000, People's Republic of China, Tel +86-18596659660, Fax +86-593-8981236, Email myy905@qq.com; Li-Li Zheng, Department of Gynecology, Mindong Hospital Affiliated to Fujian Medical University, Fuan, 355000, People's Republic of China, Tel +86-15080352165, Fax +86-593-8981915, Email 470326165@qq.com

Purpose: To develop and validate a nomogram for predicting high-grade squamous intraepithelial lesions or worse (HSIL+), incorporating results from ThinPrep cytologic test (TCT) and Aptima HPV E6/E7 mRNA (AHPV) testing.

Patients and Methods: This diagnostic study consecutively enrolled 3,202 patients referred for colposcopy due to abnormal cervical screening results. All participants underwent colposcopy with biopsy (targeted and/or endocervical) to obtain a definitive histopathological result, which served as the reference standard. The cohort was randomly split into training (70%) and validation (30%) sets. A binary logistic regression model was developed, and a nomogram was constructed. Model performance was evaluated using receiver operating characteristic (ROC) curves, calibration plots, and decision curve analysis (DCA).

Results: The final multivariate model was defined by the equation: $\text{Logit}(P) = -4.014 + 1.677 \times \text{OHR} + 2.917 \times \text{HPV16} + 1.938 \times \text{HPV18/45} + 2.343 \times \text{HPV}(16+18/45) + 0.326 \times \text{ASC-US} + 1.676 \times \text{ASC-H} + 1.161 \times \text{LSIL} + 1.593 \times \text{AGC} + 4.939 \times \geq \text{HSIL}$. A nomogram was developed using the R rms package. The model demonstrated excellent discrimination in internal validation, with areas under the ROC curves (AUCs) of 0.843 (95% CI: 0.824–0.863) in the training set and 0.833 (95% CI: 0.813–0.873) in the validation set, along with good calibration. DCA confirmed its clinical utility across a risk threshold of 2%–50%.

Conclusion: The developed logistic-nomogram provides an accurate and practical tool for predicting HSIL+, potentially aiding in individualized clinical management.

Keywords: cervical cancer screening, TCT, AHPV, HSIL+, nomogram

Introduction

Cervical cancer remains one of the most significant gynecological malignancies globally, with over 500,000 new cases and approximately 270,000 deaths annually worldwide.^{1,2} The global disease burden is unevenly distributed, with about 80% of cases concentrated in developing countries.³ The situation is particularly concerning in China, which reports more than 130,000 new cases each year and shows a disturbing trend toward younger onset ages.⁴

In 2021, the WHO released the latest edition of its Guidelines for Screening and Treatment of Cervical Cancer and Precancerous Lesions, which for the first time explicitly recommended HPV DNA testing as the preferred primary screening method for cervical cancer.⁵ The guidelines also specifically highlighted that mRNA testing targeting the HPV E6/E7 regions could serve as an effective alternative screening strategy for the general population.⁶ Among the five HPV molecular testing technologies approved by the US FDA, AHPV is the only assay designed to detect high-risk HPV E6/E7 mRNA.⁷ Extensive studies have confirmed the clinical value of mRNA testing: Oliveira et al found that mRNA testing offered higher specificity for CIN2+ (72.6% vs 42.8%); Benevolo et al demonstrated its ability to significantly reduce unnecessary colposcopy referrals in the triage of ASC-US/LSIL cases; and Ge et al reported that AHPV testing



achieved higher specificity (85.2% vs 72.3%) and positive predictive value (76.8% vs 65.4%) for HSIL detection compared to Cobas HPV testing.^{8–10}

In China, cervical cancer screening typically employs strategies combining HPV DNA testing with cytology (TCT) or using cytology for triage.^{11,12} Although AHPV testing has demonstrated high sensitivity and specificity,^{13,14} its application in routine clinical practice is not yet widespread and its potential for integration with other screening indicators to enhance risk stratification remains underexplored. However, a single screening test (be it TCT or high-risk HPV) still has room for improvement in terms of specificity and positive predictive value, potentially leading to unnecessary colposcopy referrals and imposing a burden on both the healthcare system and patients.¹⁵ When confronted with co-testing results, clinicians often rely on empirical judgment due to the lack of a standardized, quantitative tool that integrates TCT findings with AHPV genotyping information to precisely estimate an individual's absolute risk of developing high-grade lesions. This challenge is particularly pronounced for the substantial intermediate-risk population characterized by "positive for non-16/18/45 high-risk HPV types but normal cytology." Current management strategies for this group remain relatively uniform (typically recommending follow-up after 12 months), resulting in inefficient allocation of medical resources, as well as patient anxiety and loss to follow-up. Moreover, while existing guidelines provide categorical management pathways, they lack tools for quantitative individual risk assessment. Understanding the natural history of cervical lesions, such as the non-negligible regression rates of cervical intraepithelial neoplasia grade 2 (CIN2) and lower-grade lesions, further underscores the need for precise risk stratification to avoid overtreatment.^{16,17} The 2021 WHO guidelines, despite comprehensively outlining management categories, explicitly acknowledge this limitation and call for the development of better risk stratification tools to guide clinical decision-making.⁵

Based on the above context, this study aims to develop and validate a clinical prediction model integrating TCT results with AHPV. The model is presented as an intuitive nomogram, serving as a clinical decision-support tool to facilitate personalized, risk-stratified management. Ultimately, by implementing this model, we seek to move beyond the simplistic "test-and-triage" paradigm and achieve precise risk assessment for high-grade cervical lesions, thereby supporting evidence-based and individualized prevention and treatment strategies.

Materials and Methods

Subjects of the Study

This study initially enrolled 3,566 consecutive patients referred to the colposcopy clinic of Mindong Hospital, Ningde City, Fujian Province, between January 2016 and December 2024 based on abnormal cervical screening results (AHPV testing and/or TCT). A total of 364 patients were excluded because they did not meet the study's inclusion criteria, resulting in 3,202 patients for the final analysis. All included patients underwent colposcopy with successful biopsy (targeted biopsies and/or endocervical curettage), providing a definitive histopathological outcome. Inclusion criteria: (1) Sexually active women aged 18 years or older who were referred following abnormal cervical screening results and met one of the following criteria: age 25–65 years, or being 18–24 years old with high-risk factors, or being over 65 years without adequate prior screening; (2) complete clinical-pathological documentation; (3) non-menstruating status during examination. Exclusion criteria: (1) prior cervical excision/conization; (2) hysterectomy history; (3) other malignancies. This study was reviewed and approved by the Ethics Committee of Mindong Hospital, Ningde City, affiliated to Fujian Medical University (No: K2025032403), and written informed consent was obtained from all participants.

ThinPrep Cytologic Test (TCT)

Cervical exfoliated cell samples were collected using sterile cervical brushes through 8–10 rotations within the transformation zone, immediately preserved in ThinPrep Preservative Solution (Hologic, USA), and processed using the ThinPrep 2000 Processor (Hologic, USA) followed by Papanicolaou staining. Cytological diagnoses were established according to the 2001 Bethesda System (TBS) and categorized as: negative for intraepithelial lesion or malignancy (NILM); atypical squamous cells (ASC), subdivided into ASC of undetermined significance (ASC-US) and ASC cannot exclude high-grade SIL (ASC-H); low-grade squamous intraepithelial lesion (LSIL); high-grade squamous intraepithelial lesion (HSIL); squamous cell carcinoma (SCC); atypical glandular cells (AGC), subdivided into AGC not otherwise

specified (AGC-NOS) and AGC favor neoplastic (AGC-FN); adenocarcinoma in situ (AIS); and adenocarcinoma (ADC). For analytical purposes, HSIL, SCC, AIS, and ADC were collectively classified as the \geq HSIL category.

AHPV Testing

Cervical cell samples collected for TCT were utilized for AHPV testing. The residual samples preserved in ThinPrep preservative solution were tested using the Aptima HPV Assay (Hologic, USA) on the Panther fully automated molecular detection system (Hologic, USA) to qualitatively detect E6/E7 mRNA from 14 high-risk HPV genotypes (16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, 66, 68). Positive samples were subsequently genotyped for HPV 16, 18, and 45 using the Aptima HPV Genotyping Assay (Hologic, USA) on the same platform, with all procedures strictly following manufacturer protocols. HPV results were categorized as: Negative (no detectable high-risk HPV), HPV16 (HPV type 16 detected), HPV18/45 (HPV types 18 or 45 detected), HPV(16+18/45) (co-detection of HPV types 16 and 18/45), and OHR (other high-risk types excluding 16, 18, and 45 detected).

Colposcopic Biopsy

Colposcopic biopsy was performed according to the 2012 ASCCP guidelines. Referral to colposcopy was indicated for: women with HPV 16/18 positivity regardless of cytology results; those with ASC-US cytology and positive HPV testing; and those with cytological findings of ASC-H, LSIL, or HSIL, irrespective of their HPV status. All biopsies adhered to these clinical indications, with systematic four-quadrant sampling in non-lesion areas and LEEP/conization for HSIL+ cases. Histopathological diagnoses followed WHO (2020) criteria (normal/inflammatory, LSIL, HSIL, SCC, AIS, ADC, adenosquamous carcinoma), using the highest grade for discordant cases. Outcomes were dichotomized into HSIL- (normal/inflammatory, LSIL) and HSIL+ (HSIL, SCC, AIS, ADC, adenosquamous carcinoma) for analysis.

Statistical Methods

Statistical analyses were performed using SPSS 26.0 and R 4.4.3. The study cohort was randomly divided into training ($n = 2,241$) and validation ($n = 961$) sets at a 7:3 ratio. Categorical variables were presented as frequencies and percentages [n (%)] and compared using Chi-square or Fisher's exact tests as appropriate. Continuous variables with normal distribution were expressed as mean \pm standard deviation.

Univariate logistic regression was first conducted to identify potential predictors of HSIL+, followed by multivariate analysis using forward likelihood ratio method to determine independent risk factors. Regression coefficients from the final model were used to calculate risk scores and construct a predictive nomogram via the "rms" package (version 6.8.1) in R.

Model performance was evaluated using ROC curves with area under the curve (AUC), Hosmer-Lemeshow goodness-of-fit test, calibration curves, and decision curve analysis (DCA). All statistical tests were two-sided, with $P < 0.05$ considered statistically significant.

Results

Baseline Information

Based on the inclusion and exclusion criteria, a total of 3,202 cases were included in this study. The distribution of cases by histopathological diagnosis was as follows: Inflammation: 1,584 cases, aged 18–89 years, with a mean age of 46.8 ± 11.5 years; LSIL: 770 cases, aged 18–79 years, with a mean age of 42.9 ± 12.0 years; HSIL: 692 cases, aged 20–80 years, with a mean age of 45.0 ± 12.0 years; CC: 156 cases, aged 27–93 years, with a mean age of 55.4 ± 13.1 years.

The analysis of different factors influencing cervical lesions revealed that age group, TCT results, and AHPV status were all significantly associated with the severity of lesions ($P < 0.001$, Table 1). Age Distribution: The highest proportion of histologically confirmed LSIL (27.3%) and HSIL (27.5%) cases was observed in the 30–39 years age group. The ≥ 60 years age group had the highest proportion of cervical cancer cases (38.5%), followed by the 50–59 years group (26.9%). Younger women (< 30 years) showed a higher prevalence of histologic LSIL (13.9%) compared to histologic HSIL (8.2%) and CC (0.6%). TCT Results: The majority of inflammatory cases (47.6%) showed NILM results. Cytologic ASC-US was most prevalent in cases with histologic LSIL (39.5%). Cases with histologic HSIL and

Table 1 Association of Clinicopathological Characteristics with Cervical Histopathological Outcomes

Variable	Inflammatory (n=1584)	LSIL (n=770)	HSIL (n=692)	CC (n=156)	χ^2	P-value
Age (years)					188.3	<0.001
<30	89(5.6%)	107(13.9%)	57(8.2%)	1(0.6%)		
30-39	34(21.7%)	210(27.3%)	190(27.5%)	16(10.3%)		
40-49	51(32.5%)	241(31.3%)	217(31.4%)	37(23.7%)		
50-59	432(27.3%)	137(17.8%)	136(19.7%)	42(26.9%)		
≥60	204(12.9%)	75(9.7%)	92(13.3%)	60(38.5%)		
TCT					1455.9	<0.001
NILM	754(47.6%)	161(20.9%)	87(12.6%)	10(6.4%)		
ASC-US	515(32.5%)	304(39.5%)	127(18.4%)	9(5.8%)		
ASC-H	148(9.3%)	126(16.4%)	135(19.5%)	31(19.9%)		
LSIL	101(6.4%)	157(20.4%)	79(11.4%)	3(1.9%)		
AGC	58(3.7%)	12(1.6%)	13(1.9%)	12(7.7%)		
≥HSIL	8(0.5%)	10(1.3%)	251(36.3%)	91(58.3%)		
AHPV					567.0	<0.001
Negative	376 (23.7%)	44(5.7%)	12(1.7%)	8(5.1%)		
OHR	895 (56.5%)	544(70.6%)	381(55.1%)	45(28.8%)		
HPV16	170 (10.7%)	92(11.9%)	243(35.1%)	83(53.2%)		
HPV18/45	131 (8.3%)	80(10.4%)	45(6.5%)	19(12.2%)		
HPV(16+18/45)	12 (0.8%)	10(1.3%)	11(1.6%)	1(0.6%)		

Notes: 1. Data are presented as n (%).

Abbreviations: NILM, no intraepithelial lesion or malignancy; ASC-US, atypical squamous cells of undetermined significance; ASC-H, atypical squamous cells, cannot exclude high-grade squamous intraepithelial lesion; LSIL, low-grade squamous intraepithelial lesion; ≥HSIL, high-grade squamous intraepithelial lesion or worse; AGC, atypical glandular cells; OHR, other high-risk HPV types (excluding 16, 18, and 45); CC, Cervical Cancer.

CC predominantly exhibited ≥HSIL cytology results (36.3% and 58.3%, respectively), followed by ASC-H cytology (19.5% and 19.9%, respectively). Cytologic AGC was more common in CC cases (7.7%) compared to other groups. AHPV Status: OHR types were the most prevalent in inflammatory (56.5%) and histologic LSIL (70.6%) cases. In histologic HSIL cases, OHR positivity (55.1%) was followed by HPV16 positivity (35.1%). HPV16 positivity was highest in CC cases (53.2%), followed by OHR (28.8%). HPV18/45 positivity was relatively low across all groups but slightly higher in CC cases (12.2%). Results are detailed in [Table 1](#).

Construction and Validation of the HSIL+ Prediction Model

Comparison of Training Set and Validation Set Results

The comparison between the training set (n = 2,241) and the validation set (n = 961) revealed no significant differences in baseline characteristics ($P > 0.05$, [Table 2](#)), indicating balanced distribution across the two datasets. Results are detailed in [Table 2](#).

Univariate Analysis of Factors Associated with HSIL+

Univariate analysis was conducted on 2,241 patients in the training set. Analysis demonstrated that older age (≥60 years), AHPV positivity (especially HPV16 and OHR types), and abnormal TCT results (particularly ≥HSIL and ASC-H) were significant risk factors for HSIL+ ($P < 0.05$, [Table 3](#)).

Multivariate Logistic Regression for HSIL+ Predictors

Multivariate logistic regression analysis identified AHPV status and TCT findings as independent predictors of HSIL+ ($P < 0.05$, [Table 4](#)). Among all HPV genotypes, HPV16 positivity demonstrated the strongest association with HSIL+, followed by co-infection with HPV (16+18/45). Among cytological findings, ≥HSIL results showed markedly elevated risk, while ASC-H and AGC also demonstrated substantial risk increases. Results are detailed in [Table 4](#).

Table 2 Comparison of Baseline Features Between Training Set and Verification Set

Variable	Total (n=3202)	Training Set (n=2241)	Validation Set (n=961)	χ^2	P-value
Age (years)				0.9	0.927
<30	254(7.9%)	180(8.0%)	74(7.7%)		
30-39	760(23.7%)	538(24.0%)	222(23.1%)		
40-49	1010(31.5%)	697(31.1%)	313(32.6%)		
50-59	747(23.3%)	526(23.5%)	221(23.0%)		
≥60	431(13.5%)	300(13.4%)	131(13.6%)		
AHPV				7.5	0.113
Negative	440(13.7%)	324(14.5%)	116(12.1%)		
OHR	1865(58.2%)	1274(56.8%)	591(61.5%)		
HPV16	588(18.4%)	427(19.1%)	161(16.8%)		
HPV18/45	275(8.6%)	191(8.5%)	84(8.7%)		
HPV(16+18/45)	34(1.1%)	25(1.1%)	9(0.9%)		
TCT				4.1	0.533
NILM	1012(31.6%)	718(32.0%)	294(30.6%)		
ASC-US	955(29.8%)	658(29.4%)	297(30.9%)		
ASC-H	440(13.7%)	306(13.7%)	134(13.9%)		
LSIL	340(10.6%)	228(10.2%)	112(11.7%)		
AGC	95(3.0%)	70(3.1%)	25(2.6%)		
≥HSIL	360(11.2%)	261(11.6%)	99(10.3%)		
Histopathological Results				1.6	0.205
HSIL-	2354(73.5%)	1633(72.9%)	721(75.0%)		
HSIL+	848(26.5%)	608(27.1%)	240(25.0%)		

Notes: Data are presented as frequencies and percentages [n (%)].

Table 3 Univariate Analysis of Factors Associated with HSIL+ in the Training Set

Variant	HSIL- (n=1633)	HSIL+ (n=608)	OR (95% CI)	P-value
Age (years)				
<30	137(8.4%)	43(7.1%)		
30-39	393(24.1%)	145(23.8%)	1.176(0.794, 1.740)	0.419
40-49	509(31.2%)	188(30.9%)	1.177(0.804, 1.723)	0.403
50-59	404(24.7%)	122(20.1%)	0.962(0.646, 1.4320)	0.849
≥60	190(11.6%)	110(18.1%)	1.845(1.218, 2.794)	0.004
AHPV				
Negative	309(18.9%)	15(2.5%)		
OHR	982(60.1%)	292(48.0%)	6.125(3.590, 10.453)	<0.001
HPV16	187(11.5%)	240(39.5%)	26.439(15.218, 45.932)	<0.001
HPV18/45	140(8.6%)	51(8.4%)	7.504(4.080, 13.802)	<0.001
HPV(16+18/45)	15(0.9%)	10(1.6%)	13.733(5.294, 35.628)	<0.001
TCT				
NILM	648(39.7%)	70(11.5%)		
ASC-US	565(34.6%)	93(15.3%)	1.524(1.095, 2.120)	0.012
ASC-H	186(11.4%)	120(19.7%)	5.972(4.264, 8.364)	<0.001
LSIL	171(10.5%)	57(9.4%)	3.086(2.093, 4.549)	<0.001
AGC	13(0.8%)	248(40.8%)	3.703(2.085, 6.575)	<0.001
≥HSIL	50(3.1%)	20(3.3%)	176.598(95.979, 324.934)	<0.001

Construction and Validation of the Regression Model for Predicting Cervical HSIL+ Lesions

The final multivariate logistic regression model was constructed using the independent risk factors identified above. The regression equation for predicting cervical HSIL+ lesions was: $\text{Logit}(P) = -4.014 + 1.677 \times \text{OHR} + 2.917 \times \text{HPV16} +$

Table 4 Multivariate Logistic Regression for HSIL+ Predictors in the Training Set

Variant	B	SE	Wald χ^2	P	Exp (B)	95% CI
Age (years)						
<30	-	-	-	-	1.000	-
30-39	-0.058	0.236	0.059	0.807	0.944	0.594-1.500
40-49	0.104	0.232	0.203	0.653	1.110	0.705-1.749
50-59	-0.287	0.249	1.324	0.250	0.751	0.461-1.223
≥ 60	0.364	0.261	1.954	0.162	1.440	0.864-2.400
AHPV						
Negative	-	-	-	-	1.000	-
OHR	1.701	0.315	29.243	<0.001	5.480	2.958-10.152
HPV16	2.964	0.328	81.813	<0.001	19.383	10.196-36.846
HPV18/45	1.973	0.361	29.856	<0.001	7.194	3.545-14.600
HPV(16+18/45)	2.296	0.609	14.228	<0.001	9.937	3.013-32.767
TCT						
NILM	-	-	-	-	1.000	-
ASC-US	0.308	0.175	3.092	0.079	1.361	0.965-1.918
ASC-H	1.625	0.181	80.294	<0.001	5.080	3.560-7.248
LSIL	1.138	0.209	29.605	<0.001	3.119	2.071-4.699
AGC	1.655	0.324	26.087	<0.001	5.235	2.774-9.881
\geq HSIL	4.913	0.319	236.562	<0.001	136.019	72.730-254.381
Constant	-4.032	0.383	110.909	<0.001	0.018	

Table 5 Multivariate Logistic Regression Analysis of Independent Risk Factors for Predicting Cervical HSIL+ Lesions in the Training Set

Variant	B	SE	Wald χ^2	P	Exp (B)	95% CI
AHPV						
Negative	-	-	-	-	1.000	-
OHR	1.677	0.312	28.860	<0.001	5.350	2.902-9.865
HPV16	2.917	0.325	80.684	<0.001	18.491	9.784-34.947
HPV18/45	1.938	0.358	29.255	<0.001	6.948	3.442-14.025
HPV(16+18/45)	2.343	0.601	15.197	<0.001	10.412	3.206-33.818
TCT						
NILM	-	-	-	-	1.000	-
ASC-US	0.326	0.174	3.508	0.061	1.386	0.985-1.950
ASC-H	1.676	0.180	86.747	<0.001	5.345	3.756-7.606
LSIL	1.161	0.207	31.354	<0.001	3.193	2.127-4.794
AGC	1.593	0.322	24.498	<0.001	4.919	2.617-9.243
\geq HSIL	4.939	0.319	240.153	<0.001	139.623	74.761-260.759
Constant	-4.014	0.319	158.126	<0.001	0.018	

$1.938 \times \text{HPV18/45} + 2.343 \times \text{HPV(16+18/45)} + 0.326 \times \text{ASC-US} + 1.676 \times \text{ASC-H} + 1.161 \times \text{LSIL} + 1.593 \times \text{AGC} + 4.939 \times \geq \text{HSIL}$ (Table 5). The R 4.4.3 rms package (version 6.8.1) was used to create a nomogram for predicting cervical HSIL+ (Figure 1).

The model demonstrated excellent performance in internal validation. In the training set, the AUC was 0.843 (95% CI: 0.824-0.863) with good calibration (Hosmer-Lemeshow test $\chi^2 = 6.389$, $P = 0.604$). In the validation cohort ($n = 961$), the AUC was 0.833 (95% CI: 0.813-0.873) with similarly good calibration (Hosmer-Lemeshow test $\chi^2 = 6.201$, $P = 0.516$) (Figure 2). Bootstrap validation showed minimal absolute errors between predicted and actual values (0.008 and 0.005 for training and validation sets, respectively) (Figure 3). DCA demonstrated that the prediction model provided positive net

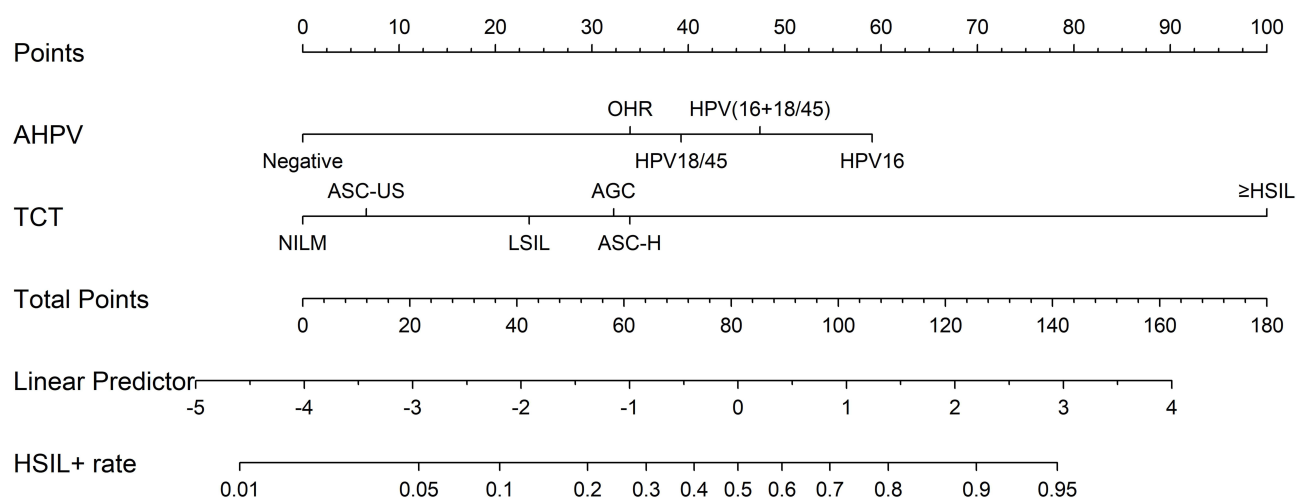


Figure 1 Nomogram for predicting high-grade squamous intraepithelial lesion or worse (HSIL+) risk. To use the nomogram: (1) Locate patient values on each variable axis, (2) Draw a line upward to the “Points” axis to determine points for each variable, (3) Sum the points from all variables, (4) Locate the total on the “Total Points” axis, (5) Draw a line downward to the “HSIL+ rate” axis to determine the predicted probability.

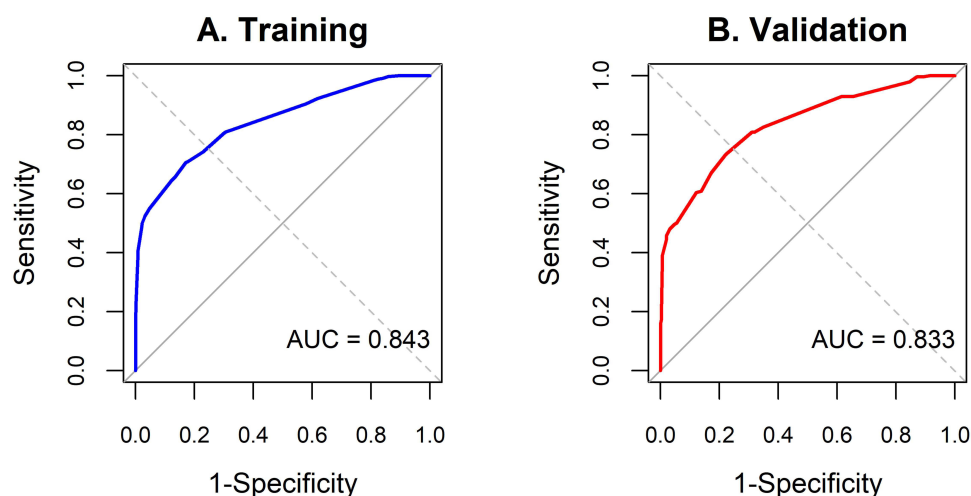


Figure 2 Receiver operating characteristic (ROC) curves of the prediction model. (A) ROC curve for the training cohort (AUC = 0.843). (B) ROC curve for the validation cohort (AUC = 0.833).

benefit across a clinically relevant risk threshold range of approximately 2–50% in both training and validation sets (Figure 4). The model consistently outperformed both the “treat all” strategy (representing immediate colposcopy for all patients) and the “treat none” strategy (representing no intervention for any patient), indicating its potential to optimize clinical decision-making by reducing unnecessary procedures while appropriately identifying high-risk individuals.

Discussion

This study aimed to construct and validate a predictive model for cervical HSIL+ by integrating age, AHPV status, and TCT results. The findings highlight the significant role of HPV infection and abnormal cytology in the development of cervical HSIL+ lesions, providing a robust tool for risk stratification and clinical decision-making.

Cervical intraepithelial neoplasia (CIN) progresses from abnormal hyperplasia to cervical cancer, with LSIL and HSIL carrying a 10-year risk of progression to invasive cancer of 15% and 30–45%, respectively.^{4,18,19} Therefore, early detection and intervention are crucial for halting disease progression and reducing cervical cancer risk. The results of this study show that the highest proportions of histologically confirmed LSIL and HSIL cases were in the 30–39 age group

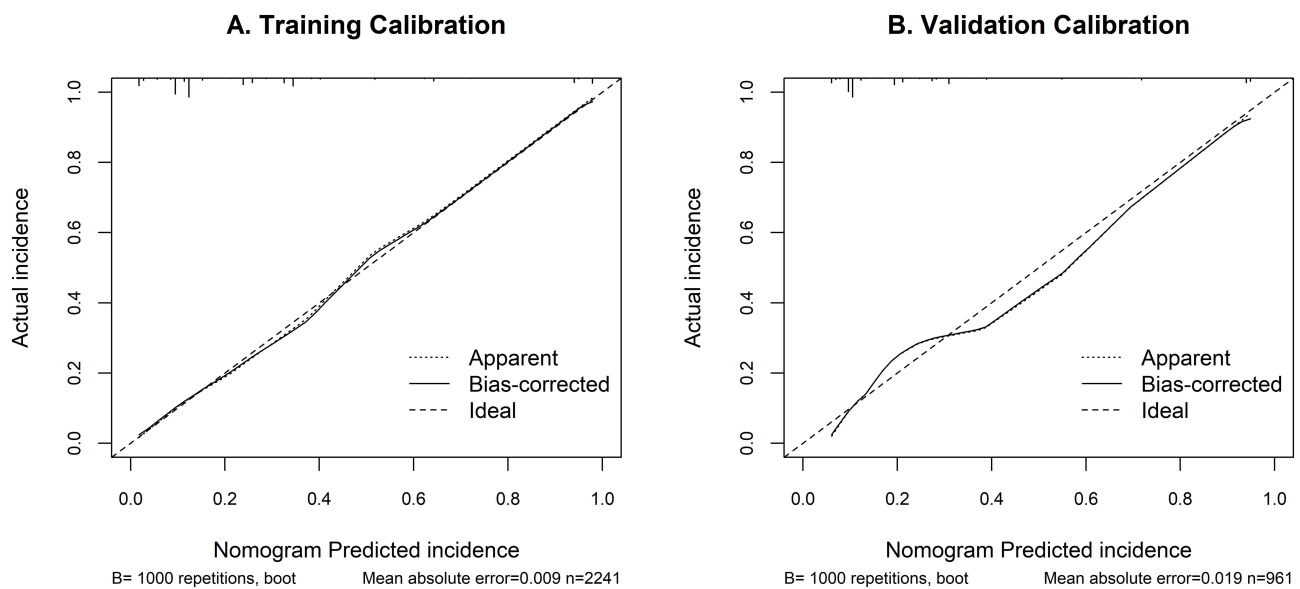


Figure 3 Calibration curves of the nomogram. **(A)** Training set calibration. **(B)** Validation set calibration. The dashed line represents the perfect reference, the solid line represents the apparent performance of the nomogram, and the dotted line indicates the bias-corrected performance after bootstrapping (1000 repetitions). Both curves demonstrate good agreement between nomogram-predicted probability and observed outcome.

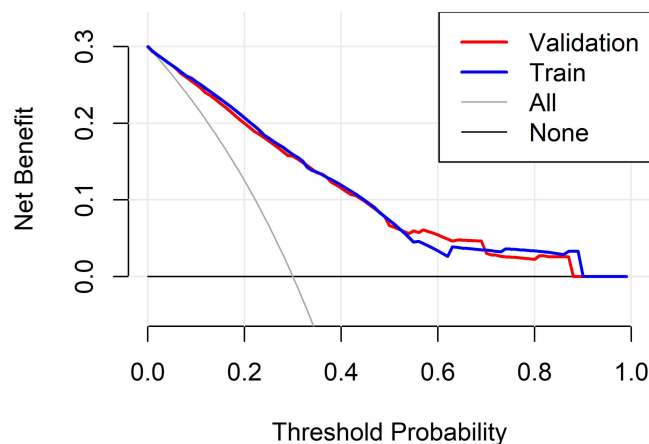


Figure 4 Decision curve analysis for the prediction model. The figure shows the net benefit of the model across different threshold probabilities in the training and validation cohorts. The “Train” curve represents the performance in the training cohort, while the “Validation” curve represents the performance in the validation cohort. The “All” and “None” lines indicate the net benefit of strategies to treat all patients or no patients, respectively. The model demonstrates clinical utility when its curve lies above both the “All” and “None” reference lines across a range of threshold probabilities.

(27.3% and 27.5%, respectively). This aligns with previous epidemiological reports indicating a trend of increasing HSIL incidence among younger women and a first peak in cervical cancer incidence occurring in the 30–39 age group.^{4,18–20} Furthermore, this study found that the 30–49 age group accounted for the majority of combined LSIL and HSIL cases, and cervical cancer cases began to appear after age 30, highlighting the necessity of enhanced screening for sexually active women.

Notably, this study found that the highest proportion of cervical cancer cases was in the ≥ 60 age group (38.5%), followed by the 50–59 age group (26.9%). This finding is consistent with existing evidence that cervical cancer incidence remains high among postmenopausal women. Currently, debate persists regarding the appropriate age to discontinue cervical cancer screening, even for women with a regular screening history.²¹ Our results underscore the importance of fully considering age-specific factors when formulating screening strategies and indicate that older women still face a substantial risk of cervical cancer, warranting continued attention.

The HPV E6/E7 oncogenes drive viral replication and tumor progression by inhibiting apoptosis and disrupting DNA repair mechanisms. HPV E6/E7 mRNA levels directly reflect the active expression of these viral oncogenes and are closely associated with the severity of cervical lesions and disease progression.^{22–25} Studies have shown that HPV E6/E7 mRNA levels increase with the progression of cervical lesions, and this assay demonstrates high sensitivity and specificity, particularly in older women.^{26,27} Our findings provide strong support for this: in histologically confirmed HSIL cases, positivity for OHR types was most prevalent (55.1%), followed by HPV16 positivity (35.1%); whereas in CC cases, HPV16 positivity was highest (53.2%), highlighting its pivotal role in carcinogenesis. This study further confirms that a positive AHPV test significantly increases the risk of HSIL and above, with OHR, HPV16, HPV18/45, and HPV(16+18/45) co-infections increasing the risk by 5-fold, 19-fold, 7-fold, and 10-fold, respectively. This indicates that HPV E6/E7 mRNA 16/18/45 genotyping can effectively predict HSIL and cervical cancer, supporting its use in screening high-risk individuals.²⁸

TCT is widely used in cervical cancer screening due to its simplicity and relatively high detection rate. Our study observed a clear correlation between cytologic and histologic findings: the majority of inflammatory cases (47.6%) showed NILM results; cytologic ASC-US was most prevalent in histologic LSIL cases (39.5%); whereas histologic HSIL and CC cases predominantly exhibited \geq HSIL cytology results (36.3% and 58.3%, respectively). However, the accuracy of TCT is limited by subjective interpretation and sample quality, suggesting that it often needs to be combined with other methods, such as HPV testing, to improve the detection efficiency of lesions.^{12,29–31}

Although several prediction models for cervical cancer exist, few incorporate both TCT and HPV E6/E7 mRNA testing, and many are limited by lack of validation or small sample sizes.^{32–36} The key innovation of our study lies in developing a quantitative risk-stratification tool that integrates these established biomarkers into a single, intuitive nomogram. This directly addresses a critical translational gap in current practice, particularly for managing the substantial “intermediate-risk” population (eg, AHPV-positive with normal cytology), where clinical decisions are often highly subjective. While WHO recommendations emphasize primary HPV screening strategies,⁶ there remains a pressing need for tools that can refine risk assessment in screen-positive women. Only two previous studies included HPV genotyping, achieving modest accuracy,^{32,37} and Song et al³⁸ reported that combining HPV E6/E7 mRNA and TCT improved diagnostic efficiency but did not develop a predictive model.

This study successfully constructed and validated a logistic regression model based on TCT findings and AHPV with 16/18/45 genotyping, using histopathology as the gold standard. Multivariate analysis identified both HPV status and TCT findings as independent predictors of HSIL+. The final model demonstrated excellent performance in both training and validation sets, with AUC values of 0.843 and 0.833, respectively, indicating robust discriminative ability. The model also showed good calibration and provided positive net benefit across a clinically relevant risk threshold range of approximately 2%-50%, consistently outperforming both the “treat all” and “treat none” strategies. Compared to traditional linear regression, logistic regression better handles categorical and nonlinear data. While machine learning models (eg, random forests) offer advantages in handling complex data, they often lack interpretability and require larger datasets. The logistic regression model developed in this study effectively balances predictive performance with interpretability and clinical applicability.

The nomogram visualization allows clinicians and patients to easily assess individual risk. More importantly, its primary clinical utility is to serve as a triage tool for women who have already tested positive on primary AHPV screening and/or have abnormal TCT results. Within this large referral population, the model helps identify those at the highest risk who are most in need of colposcopy and biopsy, thereby optimizing the use of limited healthcare resources by preventing unnecessary procedures in low-risk individuals. This facilitates personalized management and aligns with the WHO’s call for more precise screening strategies.

Strengths of this study include the large sample size, rigorous statistical methods, and internal validation of the predictive model. The use of histopathological diagnosis as the gold standard further strengthens the reliability of the findings. Limitations include the single-center design, which may limit generalizability, and the lack of external validation. While the requirement for both AHPV testing and TCT could pose challenges in some resource-constrained settings adopting primary HPV screening, our model is precisely intended to optimize subsequent management in screen-positive women, thus potentially enhancing resource efficiency rather than mandating universal co-

testing. Future studies should validate the model in diverse populations and explore the integration of additional biomarkers to further improve predictive accuracy.

Conclusion

The constructed logistic regression model, incorporating AHPV status and TCT results, effectively predicts cervical HSIL+ lesions. This model provides a valuable tool for risk stratification, enabling targeted screening and early intervention to reduce the incidence and mortality of cervical cancer. Future research should focus on external validation and the integration of novel biomarkers to enhance the model's clinical utility.

Data Sharing Statement

The datasets generated and/or analyzed during this study are available from the corresponding author (Yingyu Mao; myy905@qq.com) upon reasonable request.

Ethics Approval and Informed Consent

This study was conducted in accordance with the tenets of the Declaration of Helsinki. Informed consent was obtained from all individual participants enrolled in the study. This study was reviewed and approved by the Ethics Committee of Mindong Hospital, Ningde City, affiliated to Fujian Medical University (No: K2025032403).

Acknowledgments

Lili Zheng and Yingyu Mao contributed equally to this work as the co-correspondents' authors.

Funding

This study received no funding.

Disclosure

The authors report no conflicts of interest in this work.

References

1. Sung H, Ferlay J, Siegel RL, 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
2. Arbyn M, Weiderpass E, Bruni L, et al. Estimates of incidence and mortality of cervical cancer in 2018: a worldwide analysis. *Lancet Glob Health.* 2020;8(2):e191–e203. doi:10.1016/S2214-109X(19)30482-6
3. Bruni L, Albero G, Serrano B, et al. ICO/IARC information centre on HPV and Cancer (HPV information centre). human papillomavirus and related diseases in the world. *Summary Report.* 2023. Available from: <https://hpcvcentre.net/statistics/reports/XWX.pdf>. Accessed February 23, 2025.
4. Chen W, Zheng R, Baade PD, et al. Cancer statistics in China, 2015. *CA Cancer J Clin.* 2016;66(2):115–132. doi:10.3322/caac.21338
5. World Health Organization. *WHO Guideline for Screening and Treatment of Cervical Pre-Cancer Lesions for Cervical Cancer Prevention.* 2nd ed. Geneva: World Health Organization;2021.
6. World Health Organization. *WHO Guideline for Screening and Treatment of Cervical Pre-Cancer Lesions for Cervical Cancer Prevention: Use of mRNA Tests for Human Papillomavirus (HPV).* 2nd ed. Geneva: World Health Organization;2021.
7. Burd EM. Human papillomavirus laboratory testing: the changing paradigm. *Clin Microbiol Rev.* 2016;29(2):291–319. doi:10.1128/CMR.00013-15
8. Oliveira A, Verdasca N, Pista Â. Use of the NucliSENS EasyQ HPV assay in the management of cervical intraepithelial neoplasia. *J Med Virol.* 2013;85(7):1235–1241. doi:10.1002/jmv.23590
9. Benevolo M, Vocaturo A, Caraceni D, et al. Sensitivity, specificity, and clinical value of human papillomavirus (HPV) E6/E7 mRNA assay as a triage test for cervical cytology and HPV DNA test. *J Clin Microbiol.* 2011;49(7):2643–2650. doi:10.1128/JCM.02570-10
10. Ge Y, Christensen P, Luna E, Armylagos D, Schwartz MR, Mody DR. Performance of Aptima and Cobas HPV testing platforms in detecting high-grade cervical dysplasia and cancer. *Cancer Cytopathol.* 2017;125(8):652–657. doi:10.1002/ncy.21875
11. Zhang J, Zhao Y, Dai Y, et al. Effectiveness of high-risk human papillomavirus testing for cervical cancer screening in china: a multicenter, open-label, randomized clinical trial. *JAMA Oncol.* 2021;7(2):263–270. doi:10.1001/jamaoncol.2020.6575
12. Liu Y, Zhang L, Zhao G, Che L, Zhang H, Fang J. The clinical research of Thinprep Cytology Test (TCT) combined with HPV-DNA detection in screening cervical cancer. *Cell Mol Biol (Noisy-le-Grand).* 2017;63(2):92–95. doi:10.14715/cmb/2017.63.2.14
13. Iftner T, Neis KJ, Castanon A, et al. Longitudinal clinical performance of the RNA-Based Aptima Human Papillomavirus (AHPV) assay in comparison to the DNA-based hybrid capture 2 HPV test in two consecutive screening rounds with a 6-year interval in Germany. *J Clin Microbiol.* 2019;57(1):e01177–18. doi:10.1128/JCM.01177-18

14. Pan D, Zhang CQ, Liang QL, Hong XC. An efficient method that combines the ThinPrep cytologic test with E6/E7 mRNA testing for cervical cancer screening. *Cancer Manag Res.* 2019;11:4773–4780. doi:10.2147/CMAR.S197749
15. Kaufman HW, Alagia DP, Chen Z, Onisko A, Austin RM. Contributions of liquid-based (Papanicolaou) cytology and human papillomavirus testing in cotesting for detection of cervical cancer and precancer in the United States. *Am J Clin Pathol.* 2020;154(4):510–516. doi:10.1093/ajcp/aqaa074
16. Moscicki AB, Ma Y, Wibbelsman C, et al. Rate of and risks for regression of cervical intraepithelial neoplasia 2 in adolescents and young women. *Obstet Gynecol.* 2010;116(6):1373–1380. doi:10.1097/AOG.0b013e3181fe777f
17. Loopik DL, Bentley HA, Eijgenraam MN, IntHout J, Bekkers RLM, Bentley JR. The natural history of cervical intraepithelial neoplasia grades 1, 2, and 3: a systematic review and meta-analysis. *J Low Genit Tract Dis.* 2021;25(3):221–231. doi:10.1097/LGT.0000000000000604
18. Lawson HW. Practice bulletin No. 157: cervical cancer screening and prevention. *Obstet Gynecol.* 2016;127(1):e1–e20. doi:10.1097/AOG.0000000000001263
19. Schiffman M, Castle PE, Jeronimo J, Rodriguez AC, Wacholder S. Human papillomavirus and cervical cancer. *The Lancet.* 2007;370(9590):890–907. doi:10.1016/S0140-6736(07)61416-0
20. Arbyn M, Ronco G, Anttila A, et al. Evidence regarding human papillomavirus testing in secondary prevention of cervical cancer. *Vaccine.* 2012;30(5):F88–F99. doi:10.1016/j.vaccine.2012.06.095
21. Castañón A, Landy R, Cuzick J, Sasieni P. Cervical screening at age 50–64 years and the risk of cervical cancer at age 65 years and older: population-based case control study. *PLoS Med.* 2014;11(1):e1001585. doi:10.1371/journal.pmed.1001585
22. Zur Hausen H. Papillomaviruses and cancer: from basic studies to clinical application. *Nat Rev Cancer.* 2002;2(5):342–350. doi:10.1038/nrc798
23. Moody CA, Laimins LA. Human papillomavirus oncoproteins: pathways to transformation. *Nat Rev Cancer.* 2010;10(8):550–560. doi:10.1038/nrc2886
24. Almeida AM, Queiroz JA, Sousa F, Sousa Â. Cervical cancer and HPV infection: ongoing therapeutic research to counteract the action of E6 and E7 oncoproteins. *Drug Discov Today.* 2019;24(10):2044–2057. doi:10.1016/j.drudis.2019.07.011
25. Cuschieri K, Wentzensen N. Human papillomavirus mRNA and p16 detection as biomarkers for the improved diagnosis of cervical neoplasia. *Cancer Epidemiol Biomarkers Prev.* 2008;17(10):2536–2545. doi:10.1158/1055-9965.EPI-08-0306
26. Cuschieri K, Ronco G, Lorincz A, et al. Eurogin roadmap 2017: triage strategies for the management of HPV-positive women in cervical screening programs. *Int J Cancer.* 2018;143(4):735–745. doi:10.1002/ijc.31261
27. Pruski D, Millert-Kalinska S, Lewek A, Kedzia W. Sensitivity and specificity of HR HPV E6/E7 mRNA test in detecting cervical squamous intraepithelial lesion and cervical cancer. *Ginekol Pol.* 2019;90(2):66–71. doi:10.5603/GP.2019.0011
28. Arbyn M, Snijders PJ, Meijer CJ, et al. Which high-risk HPV assays fulfil criteria for use in primary cervical cancer screening? *Clin Microbiol Infect.* 2015;21(9):817–826. doi:10.1016/j.cmi.2015.04.015
29. Nayar R, Wilbur DC. *The Bethesda System for Reporting Cervical Cytology: Definitions, Criteria, and Explanatory Notes.* 3rd ed. Cham: Springer International Publishing; 2015.
30. Trzeszcz M, Mazurec M, Jach R, et al. Liquid-based screening tests results: HPV, liquid-based cytology, and P16/Ki67 dual-staining in private-based opportunistic cervical cancer screening. *Diagnostics.* 2021;11(8):1420. doi:10.3390/diagnostics11081420
31. Hu H, Zhao J, Yu W, et al. Human papillomavirus DNA, HPV L1 capsid protein and p16INK4a protein as markers to predict cervical lesion progression. *Arch Gynecol Obstet.* 2019;299(1):141–149. doi:10.1007/s00404-018-4931-1
32. Kahng J, Kim EH, Kim HG, Lee W. Development of a cervical cancer progress prediction tool for human papillomavirus-positive Koreans: a support vector machine-based approach. *J Int Med Res.* 2015;43(4):518–525. doi:10.1177/0300060515577846
33. Sun L, Yang L, Liu X, et al. Optimization of cervical cancer screening: a stacking-integrated machine learning algorithm based on demographic, behavioral, and clinical factors. *Front Oncol.* 2022;12:821453. doi:10.3389/fonc.2022.821453
34. Koeneman MM, van Lint FHM, van Kuijk SMJ, et al. A prediction model for spontaneous regression of cervical intraepithelial neoplasia grade 2, based on simple clinical parameters. *Hum Pathol.* 2017;59:62–69. doi:10.1016/j.humpath.2016.09.012
35. Charlton BM, Carwile JL, Michels KB, Feldman S. A cervical abnormality risk prediction model: can we use clinical information to predict which patients with ASCUS/LSIL Pap tests will develop CIN 2/3 or AIS? *J Low Genit Tract Dis.* 2013;17(3):242–247. doi:10.1097/LGT.0b013e3182730fec
36. Vukovic D, Antic L, Vasiljevic M, Antic D, Matejic B. Development of a risk index for prediction of abnormal pap test results in Serbia. *Asian Pac J Cancer Prev.* 2015;16(8):3527–3531. doi:10.7314/apjcp.2015.16.8.3527
37. van der Waal D, Bekkers RLM, Dick S, et al. Risk prediction of cervical abnormalities: the value of sociodemographic and lifestyle factors in addition to HPV status. *Prev Med.* 2020;130:105927. doi:10.1016/j.yjmed.2019.105927
38. Song Y, Zhang M, Zhang C, Du S, Zhai F. HPV E6/E7 mRNA combined with thin-prep cytology test for the diagnosis of residual/recurrence after loop electrosurgical excision procedure in patients with cervical intraepithelial neoplasia. *Diagn Microbiol Infect Dis.* 2024;108(1):116119. doi:10.1016/j.diagmicrobio.2023.116119

International Journal of Women's Health

Publish your work in this journal

The International Journal of Women's Health is an international, peer-reviewed open-access journal publishing original research, reports, editorials, reviews and commentaries on all aspects of women's healthcare including gynecology, obstetrics, and breast cancer. The manuscript management system is completely online and includes a very quick and fair peer-review system, which is all easy to use. Visit <http://www.dovepress.com/testimonials.php> to read real quotes from published authors.

Submit your manuscript here: <https://www.dovepress.com/international-journal-of-womens-health-journal>

Dovepress
Taylor & Francis Group