

RESEARCH

Open Access



Assessing the risk of high-grade squamous intraepithelial lesions (HSIL+) in women with LSIL biopsies: a machine learning-based study

Dongmei Li^{1†}, Zhichao Wang², Yan Liu³, Meiyuan Zhou⁴, Bo Xia⁴, Lin Zhang¹, Keming Chen^{1*} and Yong Zeng^{1*†}

Abstract

Objective This study aims to analyze factors associated with the missed diagnosis of high-grade squamous intraepithelial lesions (HSIL+) in patients initially diagnosed with low-grade squamous intraepithelial lesions (LSIL) through colposcopic biopsy and to develop a predictive model for assessing the risk of missed HSIL+.

Methods We conducted a retrospective analysis of 505 patients who underwent loop electrical excision procedure (LEEP) following an LSIL diagnosis by colposcopic biopsy. Logistic regression was used to identify demographic and pathological parameters associated with missed diagnoses of HSIL+. Additionally, several machine learning methods were employed to construct and assess the performance of the risk prediction models.

Results The overall rate of missed diagnoses for HSIL+ was 15.2%. Independent risk factors identified were HPV16/18 infection (OR 2.071; 95% CI 1.039–4.127; $p=0.039$), TCT \geq ASC-H (OR 4.147; 95% CI 1.392–12.355; $p=0.011$), TZ3 (OR 1.966; 95% CI 1.003–3.853; $p=0.049$) and Colposcopic impression G2 (OR 3.627; 95% CI 1.350–9.743; $p=0.011$). Among the models tested, the Decision Tree algorithm demonstrated superior performance with an accuracy of 94.7%, sensitivity of 80.0%, specificity of 96.9%, and an area under the curve (AUC) of 0.936 in the validation set.

Conclusion Key independent risk factors for the missed diagnosis of HSIL in patients with LSIL include HPV16/18 infection, TCT \geq ASC-H, TZ3, and colposcopic impression G2. The Decision Tree model offers a cost-effective, reliable, and clinically valuable tool for accurately predicting the risk of missed diagnosis of HSIL+, facilitating early intervention and management.

Keywords LSIL, HSIL, Missed diagnosis, Predictive model, Machine learning, Colposcopic biopsy, Cervical cancer screening, Risk assessment

[†]Dongmei Li and Yong Zeng are Co-first authors.

*Correspondence:

Keming Chen

chenkeming1969@163.com

Yong Zeng

349213101@qq.com

Full list of author information is available at the end of the article



Introduction

Low-grade squamous intraepithelial lesion (LSIL) is a morphological change in the squamous epithelium of the cervix following human papilloma virus (HPV) infection. As a histological manifestation of HPV, LSIL typically has a favorable prognosis: about 60% of these lesions spontaneously regress within 1 year, 30% persist, and only around 10% progress to high-grade squamous intraepithelial lesions (HSIL) within 2 years [1]. Despite this, the management and standardization of LSIL present considerable challenges. For patients diagnosed with LSIL by biopsy, the primary objectives are to prevent missed diagnoses of more severe HSIL+ and to accurately assess the risk of progression to HSIL+, thus avoiding excessive treatment [2, 3].

Research shows that patients with LSIL, identified through abnormal cervical cancer screening and followed by colposcopic biopsy, are at risk of undetected HSIL+ [4]. Without intervention, approximately 5% of CIN2 lesions and 12–33% of CIN3 lesions may progress to invasive cancer over 20–30 years [5]. Identifying risk factors for HSIL+ in the follow-up of patients with LSIL is therefore crucial. In this study, we retrospectively analyzed the clinical data of patients diagnosed with LSIL via colposcopic biopsy and subsequently treated surgically at our hospital. Our goal was to identify potential risk factors for the missed diagnosis of HSIL+ and to develop a robust predictive model. This model aims to estimate the risk of missed HSIL+ diagnoses, thereby enhancing clinical outcomes for these patients.

Materials and methods

We collected clinical data from patients with abnormal cervical cancer screening results who were diagnosed with low-grade squamous intraepithelial lesions (LSIL) via colposcopic biopsy and subsequently treated surgically at a single hospital from January 2017 to December 2022. This study received approval from the Ethics Committee, and informed consent was obtained from all participants and/or their families. Inclusion Criteria: (1) Patients with abnormal cervical cancer screening and colposcopic cervical biopsy; (2) Diagnosis of LSIL confirmed by cervical biopsy; (3) Loop electrical excision procedure (LEEP) performed within 3 months following the colposcopic cervical biopsy; (4) Availability of complete clinicopathological data. Patients were categorized into a training set and a validation set in a 7:3 ratio, based on the sequence of their admission, resulting in 354 cases in the training set and 151 cases in the validation set.

HPV Testing and Genotyping: The Cobas HPV test (Cobas 4800; Roche Molecular Diagnostics) was employed, utilizing a real-time polymerase chain reaction (PCR) system. This test detects 14 high-risk HPV

(HR-HPV) types and provides specific data on HPV16/18 infections. Multiple HR-HPV infections were defined as the presence of two or more HR-HPV types.

Cervical Cytology Diagnosis: Cervical cytology was classified according to the Bethesda system (TBS), which includes the following categories: No abnormal cells (NILM); Atypical squamous cells of undetermined significance (ASCUS); Low-grade squamous intraepithelial lesion (LSIL); Atypical squamous cells—cannot exclude high-grade squamous intraepithelial lesion (ASC-H); High-grade squamous intraepithelial lesion (HSIL); Squamous cell carcinoma (SCC); For clarity, cytological findings are grouped as low-level lesions (ASCUS, LSIL) and high-level lesions (ASC-H, HSIL, SCC).

Colposcopy Indications: (1) Positive for HPV16/18; (2) Cytology indicating atypical squamous cells of undetermined significance (ASC-US) concurrently with high-risk HPV positive; (3) Cytology showing atypical glandular cells (AGC), LSIL, or higher.

Colposcopic Assessment: The initial step in colposcopic evaluation involves identifying the cervical transformation zone. The types of transformation zones are classified as follows: Type I: The transformation zone is entirely external to the cervical canal, allowing complete visualization of both the transformation zone and any lesions under the colposcope. Type II: The transformation zone partially extends inside the cervical canal and partially remains outside. However, all transformation areas can be observed with the aid of auxiliary means during colposcopy. Type III: Only a small portion of the transformation zone is outside the cervical canal, or it is entirely within the canal, making the boundaries of the transformation zone invisible under the colposcope. A biopsy is performed on any suspicious lesions observed during colposcopy. If no suspicious lesions or a Type III transformation zone is present, a random multi-point biopsy combined with endocervical curettage is conducted.

Lesion Extent: The cervix is divided into four quadrants, each representing 25% of the total area. The extent of the lesion is determined based on the number of quadrants affected as indicated in the biopsy's pathological results. A lesion extent of ≥ 3 quadrants suggests involvement across three or more quadrants.

Definition of Missed Diagnosis: Patients diagnosed postoperatively with HSIL or invasive cancer were considered cases of missed diagnosis.

Statistical analysis

In this study, data organization and statistical analysis were performed using SPSS (version 24.0), R (version 4.3.1), and Python software. Initially, univariate and multivariate analyses were conducted to identify risk factors

for the missed diagnosis of HSIL+. A nomogram model was subsequently developed based on these identified risk factors, and its validity was tested through a five-fold cross-validation scheme. Several machine learning algorithms were employed to construct the HSIL+ risk prediction models in the training set, including Logistic Regression (LR), Naive Bayes, Support Vector Machine (SVM), K-Nearest Neighbor (KNN), Decision Tree, Random Forest, and Extreme Gradient Boosting (XGBoost). These models were applied to the data from both the training and validation set. A confusion matrix was created for each model to compare the predicted outcomes with the actual data. The performance of each model was evaluated by calculating accuracy, sensitivity, specificity, and other metrics derived from the confusion matrices. Receiver Operating Characteristic (ROC) curves were plotted to assess the predictive value of each model using the Area Under the Curve (AUC). Additionally, the Hosmer–Lemeshow test was utilized to examine the goodness of fit, while calibration curves and decision curve analysis (DCA) were applied to evaluate the calibration and clinical utility of the nomogram model, respectively. Statistical significance was set at $p < 0.05$.

Results

Postoperative histological examination of the LEEP specimens from 505 patients showed that 428 maintained a diagnosis of LSIL, while 77 were upgraded to HSIL+; this included 75 cases of HSIL, one case of adenocarcinoma in situ (AIS), and one case of cervical invasive adenocarcinoma, resulting in a missed diagnosis rate of 15.2%. The average age of the patients was 47 years, ranging from 21 to 78 years. Notably, 207 patients (41%) were aged 50 years or older, and these patients exhibited a higher missed diagnosis rate of HSIL+ at 18.8%. Additionally, 178 patients were postmenopausal, with a missed diagnosis rate of 16.3%. Regarding HPV status, 445 patients (88.1%) tested positive for high-risk HPV (HR-HPV), with 161 (31.9%) having multiple HR-HPV infections and 139 (27.5%) specifically infected with HPV16/18. The missed diagnosis rates for patients with multiple HR-HPV infections and those with HPV16/18 were 20.5% and 25.2%, respectively. Abnormal cytology was observed in 270 patients (53.5%), broken down as follows: 160 with ASCUS, 71 with LSIL, 24 with ASC-H, and 15 with HSIL. The respective missed diagnosis rates were 11.9%, 15.5%, 70.8%, and 60%. Additionally, the missed diagnosis rates among patients with colposcopic impressions of G1 and G2 were 11% and 57.1%, respectively. Across transformation zones 1, 2, and 3 (TZ1, TZ2, and TZ3), the missed diagnosis rates were 8.5%, 16.9%, and 25%, respectively. The patients were divided into a training set and a validation set in a 7:3 ratio based on the time of admission,

comprising 354 and 151 cases, respectively. The analysis revealed no significant difference in baseline characteristics between the two groups ($p > 0.05$), as detailed in Table 1.

Our analysis explored demographic and clinicopathological parameters that could be associated with missed diagnoses of HSIL+. Univariate analysis identified several factors significantly correlated with missed diagnoses. These included HPV16/18 infection ($p < 0.001$), multiple HR-HPV infections ($p = 0.018$), TCT \geq ASC-H ($p < 0.001$), colposcopic impression G2 ($p < 0.001$), transformation zone 3 (TZ3) ($p = 0.004$), and lesion extent ($p = 0.005$). In contrast, factors such as age ($p = 0.143$), menopausal status ($p = 0.68$), immune diseases ($p = 0.591$), smoking ($p = 0.219$), gravidity ($p = 0.573$), parity ($p = 0.78$), and hormone use ($p = 0.342$) did not show significant correlations with missed diagnoses of HSIL+. Details of these findings are summarized in Table 2.

Variables identified as significantly associated with missed diagnoses of HSIL+ in the univariate analysis were further evaluated using a logistic multiple factor regression equation. The logistic regression analysis confirmed several independent risk factors for the missed diagnosis of HSIL+. These included HPV16/18 infection, with an odds ratio (OR) of 2.071 (95% confidence interval [CI] 1.039–4.127; $p = 0.039$), TCT \geq ASC-H (OR 4.147; 95% CI 1.392–12.355; $p = 0.011$), transformation zone 3 (TZ3) (OR 1.966; 95% CI 1.003–3.853; $p = 0.049$), and colposcopic impression G2 (OR 3.627; 95% CI 1.350–9.743; $p = 0.011$). These factors were established as the independent predictors for missed diagnosis of HSIL+. The detailed results are presented in Table 3.

The four independent risk factors identified in the multivariate logistic regression analysis—HPV16/18 infection, TCT \geq ASC-H, TZ3, and colposcopic impression G2—were incorporated as predictors in the construction of the nomogram model. The nomogram assigns corresponding scores to each predictor, which are summed to derive a total score. This total score is then used to estimate the probability of a missed diagnosis of HSIL+ (Fig. 1).

To assess the predictive performance of various models, seven machine learning algorithms were applied to the training set. The results, including prediction accuracy, sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV), were derived from the confusion matrices based on actual and predicted values. Additionally, Receiver Operating Characteristic (ROC) curves were plotted and the area under the curve (AUC) was calculated for these models. The DecisionTree and RandomForest models demonstrated superior performance in the training set, as evidenced by various metrics (Figs. 2 and 3, Table 4).

Table 1 Demographics and clinical characteristics of study population

Parameter	Numbers		Parameter	Numbers	
	Total	Missed diagnosis (%)		Total	Missed diagnosis (%)
Age(years)			Immune disease		
< 50	298	38 (12.8%)	No	437	67 (15.3%)
≥ 50	207	39 (18.8%)	Yes	68	10 (14.7%)
Menopause			Smoking		
No	327	48 (14.7%)	No	464	68 (14.7%)
Yes	178	29 (16.3%)	Yes	41	9 (22%)
HPV statuses			Gravidity		
Negative	48	3 (6.3%)	< 3	334	48 (14.4%)
HR-HPV	445	72 (16.2%)	≥ 3	171	29 (17%)
HPV16/18	139	35 (25.2%)	Parity		
Multiple HR-HPV	161	33 (20.5%)	< 2	314	46 (14.6%)
Colposcopic impression			≥ 2	191	31 (16.2%)
Normal	67	3 (4.5%)	Hormone use		
Low-Grade(G1)	382	42 (11%)	No	458	72 (15.7%)
High-Grade(G2)	56	32 (57.1%)	Yes	47	5 (10.6%)
TZ types			Cytology results		
TZ1	223	19 (8.5%)	NILM	235	21 (8.9%)
TZ2	154	26 (16.9%)	ASC-US	160	19 (11.9%)
TZ3	128	32 (25%)	LSIL	71	11 (15.5%)
Lesion range			ASC-H	24	17(70.8%)
< 3	245	21 (8.6%)	HSIL	15	9 (60%)
≥ 3	260	56 (21.5%)			

The same seven machine learning models were then employed on an validation set to further evaluate their predictive accuracy, sensitivity, specificity, PPV, and NPV. ROC curves were again drawn using the validation set data, and the AUC was calculated. The DecisionTree model consistently showed the best predictive performance in the validation set, making it the optimal prediction model for this study (Table 5, Fig. 4).

Thus, based on comprehensive evaluations using both training and validation set, the DecisionTree emerges as the most effective model for predicting missed diagnoses of HSIL+ in this study.

The nomogram model’s goodness of fit and calibration were evaluated using the Hosmer–Lemeshow goodness of fit test and calibration curves. In the training set, the Hosmer–Lemeshow test yielded a χ^2 value of 0.102 with a *p*-value of 1, and in the validation set, a χ^2 of 4.09 with a *p*-value of 0.905; both *p*-values being greater than 0.05 indicate a good fit of the model. The calibration curves for the training set demonstrated high consistency between the model’s predicted risks of missed diagnosis of HSIL+ and the actual observed risks, confirming the model’s accuracy. The calibration curve for the validation set also showed good consistency, underscoring the model’s robustness across different datasets (Fig. 5A, B).

The clinical utility of the prediction model was further assessed by drawing clinical decision curves. The results depicted the dotted line in the upper right quadrant relative to the ‘All’ and ‘None’ lines, indicating that the HSIL+ nomogram prediction model possesses substantial clinical practical value in predicting the risk of missed diagnosis (Fig. 6A, B).

Discussion

Low-grade squamous intraepithelial lesions (LSIL) arise as histological changes secondary to HPV infection, with approximately 80% of LSIL cases attributed to high-risk HPV (HR-HPV) infections [6]. This is supported by our study, which reported an HR-HPV infection rate of 88.1%. The complexity and transient nature of HPV infections make the management of LSIL challenging, often leading to missed diagnoses and overtreatment. Overtreatment can result in complications such as bleeding and stenosis, which may impede subsequent follow-ups [7]. Conversely, missed diagnoses increase the risk of progressing to cervical cancer [4, 5]. Consequently, clinical follow-up observation is the recommended management approach for LSIL, provided more severe lesions are excluded. In clinical practice, accurately identifying patients at risk of missed HSIL+ remains problematic. This challenge

Table 2 Univariate analyses for prediction of missed diagnosis HSIL+ in patients with LSIL diagnosed by colposcopy biopsy in the training set

Parameter	Numbers		OR (95%CI)	p-values
	Total	Missed diagnosis (%)		
Age (years)			0.654 (0.370–1.156)	0.143
< 50	211	29 (13.7%)		
≥ 50	143	28 (19.6%)		
Menopause			0.883 (0.490–1.592)	0.68
No	232	36 (15.5%)		
Yes	122	21 (17.2%)		
HPV-16/18 infection			2.813 (1.569–5.042)	0
No	255	30 (11.8%)		
Yes	99	27 (27.3%)		
Multiple HR-HPV infection			1.992 (1.118–3.549)	0.018
No	240	31 (12.9%)		
Yes	114	26 (22.8%)		
TZ types			2.368 (1.305–4.296)	0.004
TZ1/2	265	34 (13.3%)		
TZ3	89	23 (26.1%)		
Lesion range			2.369 (1.285–4.366)	0.005
< 3	166	17 (10.2%)		
≥ 3	188	40 (21.3%)		
TCT ≥ ASC-H			11.200 (4.762–26.340)	< 0.001
NO	328	41 (12.5%)		
Yes	26	16 (61.5%)		
Colposcopic impression			9.743 (4.765–19.921)	< 0.001
Normal/G1	314	35 (11.2%)		
G2	40	22 (55%)		
Immune disease			0.806 (0.367–1.772)	0.591
No	306	48 (15.7%)		
Yes	48	9 (18.8%)		
Smoking			0.571 (0.232–1.409)	0.219
No	325	50 (15.4%)		
Yes	29	7 (24.1%)		
Gravidity			1.185 (0.657–2.137)	0.573
< 3	119	21 (17.6%)		
≥ 3	235	36 (15.3%)		
Parity			0.921 (0.516–1.642)	0.78
< 2	217	34 (15.7%)		
≥ 2	137	23 (16.8%)		
Hormone use			0.556(0.163–1.897)	0.342
No	324	54 (16.7)		
Yes	30	10 (10%)		

has led to the adoption of “risk stratification” as a strategy to standardize the management of LSIL [8]. Literature suggests that the rate of missed HSIL+ diagnoses in patients with LSIL diagnosed by colposcopic biopsy ranges from 10 to 55% [4, 9]; our study aligns with these

findings, presenting a missed diagnosis rate of 15.2%. For patients whose pathology was upgraded to HSIL after LEEP, we referred to the “Chinese Expert Consensus on the Management of High-grade Cervical intraepithelial Lesions” for management [2]. Factors associated with

Table 3 Multivariate analyses for prediction of missed diagnosis HSIL+ in patients with LSIL diagnosed by colposcopy biopsy in the training set

Parameter	B	S.E	Wald	P	OR	95% C.I
HPV16/18	0.728	0.352	4.277	0.039	2.071	1.039 4.127
TCT ≥ ASC-H	1.422	0.557	6.520	0.011	4.147	1.392 12.355
TZ3	0.676	0.343	3.875	0.049	1.966	1.003 3.853
Colposcopic impression G2	1.289	0.504	6.533	0.011	3.627	1.350 9.743

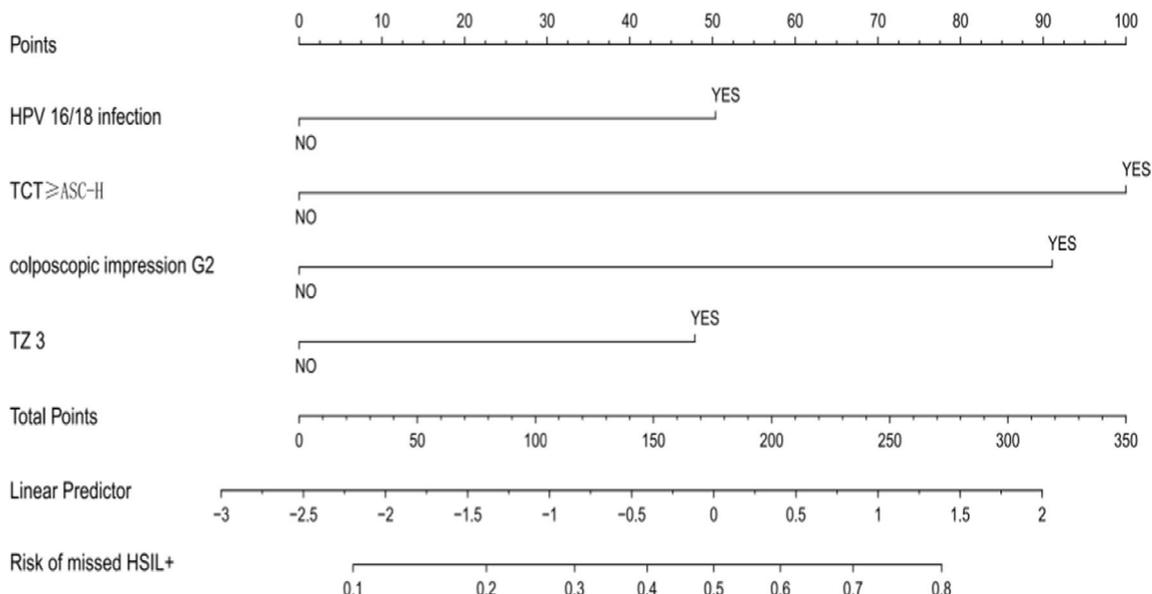


Fig. 1 Nomogram prediction model of missed diagnosis HSIL+ in patients with LSIL diagnosed by colposcopic biopsy. The four independent risk factors identified in the multivariate logistic regression analysis—HPV16/18 infection, TCT ≥ ASC-H, TZ3, and colposcopic impression G2—were included as final predictors in the model. Then R software was used to construct a nomogram prediction model for the risk of missed diagnosis of HSIL+. Result interpretation: each factor took a vertical line, corresponding to the top “Points” score, and then added the four factor scores to get the Total Points; Then the total score was taken as a vertical line, and the point corresponding to Risk was the risk of missed diagnosis of HSIL+

missed diagnoses in the literature include severe cytological abnormalities, extensive lesion range, transformation zone, immunosuppression, and gland involvement [10, 11]. Our findings confirm that HPV16/18 infection, TCT ≥ ASC-H, TZ3, and colposcopic impression G2 are independent risk factors for the missed diagnosis of HSIL+. The nomogram model developed based on these factors demonstrated high predictive value.

Persistent HR-HPV infection, particularly with HPV16 and HPV18, is a well-recognized precursor to precancerous lesions and cervical cancer. In our cohort, 88.1% were infected with HR-HPV, 27.5% had HPV16/18 infections, and 31.9% had multiple HR-HPV infections, highlighting the significant role of HR-HPV in the etiology of cervical precancerous lesions. Studies have repeatedly shown that HPV16/18 infections significantly elevate the risk of developing HSIL+, with HPV16/18 positive women, even those with negative cytology, more likely to develop

CIN2+ compared to those infected with other HR-HPV types [12]. Our analysis found that the rate of missed HSIL+ diagnoses was notably higher in women positive for HPV16/18 (25.2%) compared to those who were not (11.5%). Moreover, HPV16, known for its higher integration frequency with human genes, poses the greatest risk among HR-HPVs, making it the most prevalent type in HSIL+ cases [13]. A meta-analysis found that HPV16 accounts for 34–52% of all high-grade cervical lesions [14, 15], and a prospective cohort study indicated that women with HPV16 are significantly more likely to develop CIN2+ compared to those without HPV16 [16]. Given these data, HPV16/18 infections, particularly HPV16, emerge as crucial predictive markers for missed HSIL+ diagnoses in patients with LSIL, underscoring the importance of targeted HPV screening in this group.

Studies such as the Kaiser Permanente Northern California Medical Care Plan (KPNC) have highlighted the

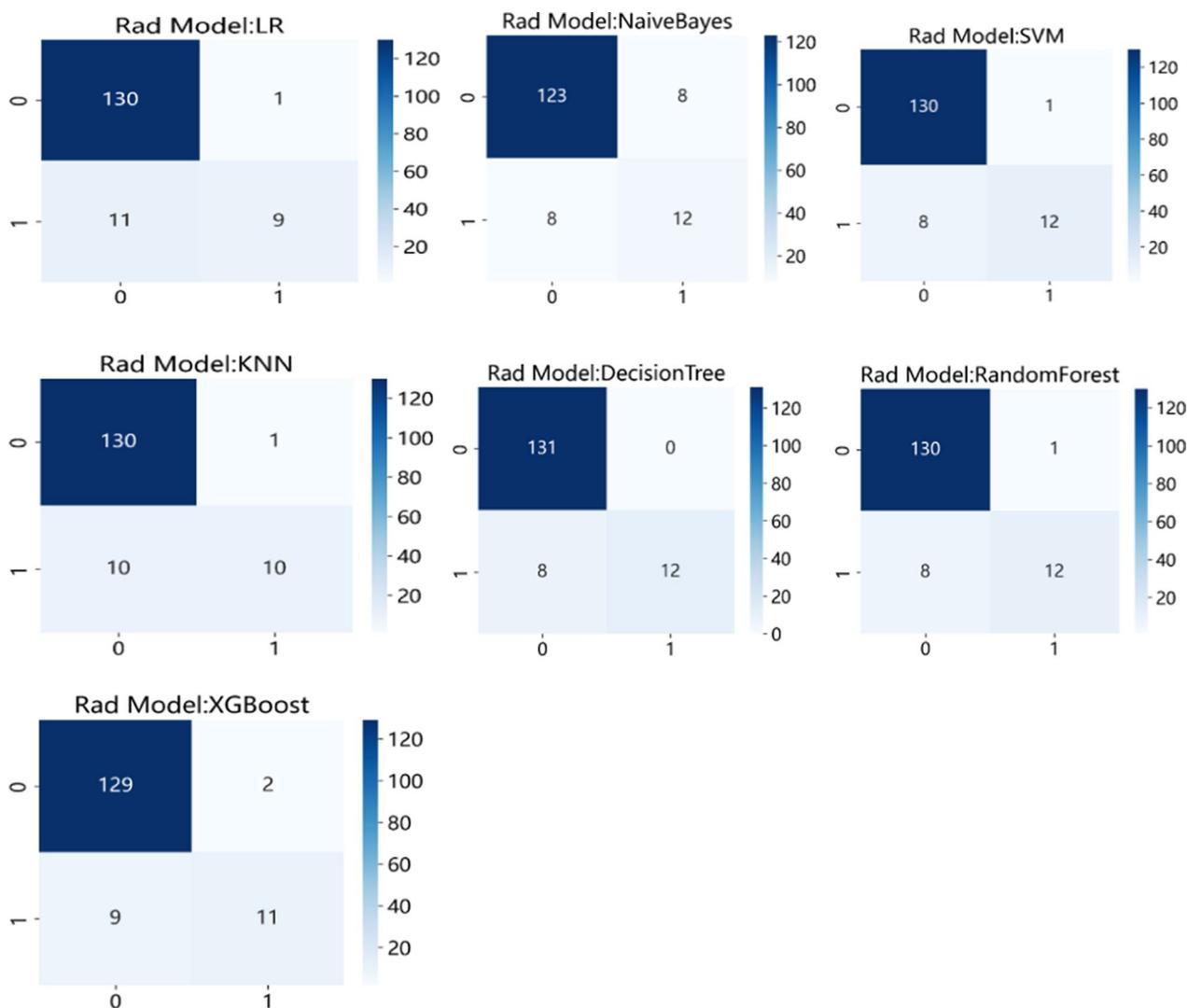


Fig. 2 Confusion matrix of the 7 models in training set. The confusion matrix of seven kinds of machine learning models were constructed according to the real values of the samples in the training set and the predicted values of the model. A line label of "0" meant that the true result of the sample was no missed diagnosis, and a line label of "1" meant that the true result of the sample was missed diagnosis. The column label "0" meant that the model predicted no missed diagnosis, and the row label "1" meant that the model predicted missed diagnosis. The upper left corner showed the number of cases no missed diagnosis by both the model predicted value and the sample true value, and the right corner showed the number of cases with missed diagnosis by both the model predicted value and the sample true value. The darker the color of the cell, the more times of correct classification

importance of prior cytology screening results in stratifying the risk of CIN3+ in women diagnosed with normal or \leq CIN1 lesions by colposcopic biopsy [17]. For instance, women diagnosed with CIN1 and concurrent cytology of LSIL or HPV-positive ASCUS exhibited a 5-year risk of developing CIN3+ of 3.8%, whereas this risk escalated to 15% for those with CIN1 and cytology of HSIL [17]. These findings emphasize the significant role of cytology screening results in the stratified management of LSIL, an approach also endorsed by the Chinese expert consensus [1]. In our study, the missed diagnosis

rates of HSIL+ corresponding to cytology of ASCUS, LSIL, ASC-H, and HSIL were 11.9%, 15.5%, 70.8%, and 60%, respectively. Notably, as cytological abnormalities increased, so did the rates of missed diagnoses. For patients with cytological findings of HSIL and ASC-H on a background of histologic LSIL, the 1-year risks of progressing to CIN3+ were 3.9% and 1.4%, respectively [18]. Furthermore, KPNC's findings suggest that women with cytologic ASC-H are at a risk level closer to HSIL than to LSIL [17], indicating that management protocols for cytology ASC-H should align more closely with

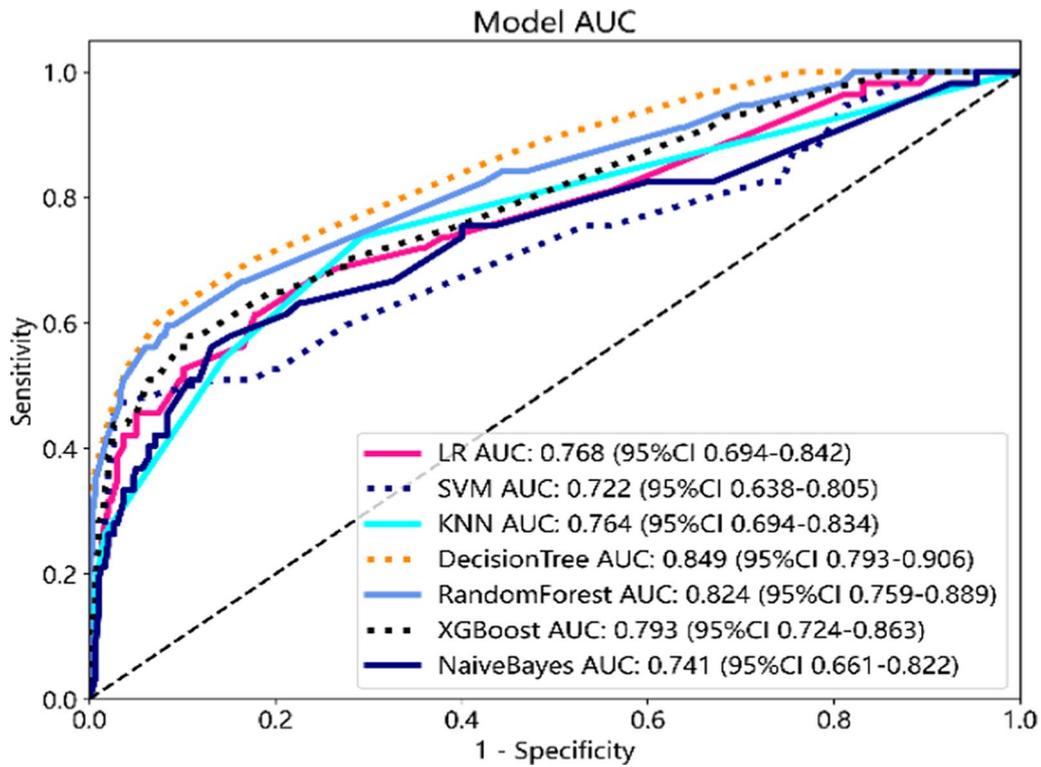


Fig. 3 AUC curves of different machine learning prediction models in training set. The ROC curves of different models were drawn using the training set. The area under the ROC curve (AUC) was calculated to evaluate the differentiation of different models in the training set. The lines in different colors represented different models, and the AUC values and 95% confidence intervals for each model were listed at the bottom right. In the figure, the area under the curve corresponding to the yellow line was the largest, and the AUC value was 0.849 (95% CI 0.793–0.906), indicating that the Decision Tree model had the best differentiation in the training set

Table 4 Predictive performance of different machine learning models in training set

Model	ACC (%)	AUC	Sensitivity	Specificity	PPV	NPV
Logistic regression	79.1	0.768	0.561	0.835	0.395	0.908
Naïve bayes	88.2	0.741	0.509	0.882	0.453	0.903
SVM	89.0	0.722	0.456	0.973	0.765	0.903
KNN	80.5	0.764	0.544	0.855	0.419	0.907
Decision tree	87.6	0.849	0.596	0.929	0.618	0.923
Random forest	86.4	0.824	0.579	0.919	0.579	0.919
XGBoost	83.9	0.793	0.561	0.892	0.500	0.914

Table 5 Predictive performance of different machine learning models in validation set

Model	ACC (%)	AUC	Sensitivity	Specificity	PPV	NPV
Logistic regression	72.1	0.843	0.600	0.969	0.750	0.941
Naïve bayes	79.5	0.814	0.700	0.809	0.359	0.946
SVM	93.4	0.801	0.600	0.985	0.857	0.942
KNN	92.7	0.861	0.500	0.992	0.909	0.929
Decision tree	94.7	0.936	0.800	0.969	0.800	0.969
Random forest	90.7	0.940	0.800	0.924	0.615	0.968
XGBoost	88.1	0.893	0.800	0.893	0.533	0.967

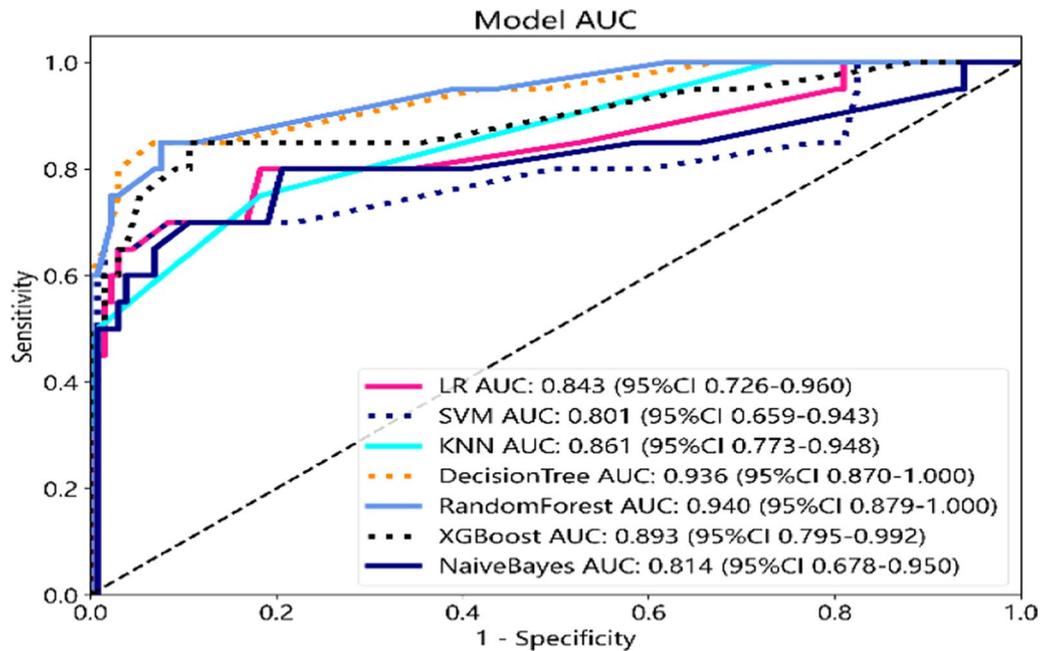


Fig. 4 AUC curves of different machine learning prediction models in validation set. The ROC curves of different models were drawn using the validation set. The area under the ROC curve (AUC) was calculated to evaluate the differentiation of different models in the validation set. The lines in different colors represented different models, and the AUC values and 95% confidence intervals for each model were listed at the bottom right. In the figure, the area under the curve corresponding to the yellow dotted line and the blue solid line was the largest, and the AUC value was 0.936(95% CI 0.870–1.000) and 0.940(95% CI 0.879–1.000), respectively, indicating that the Decision Tree and Random Forest models had the best differentiation in the validation set

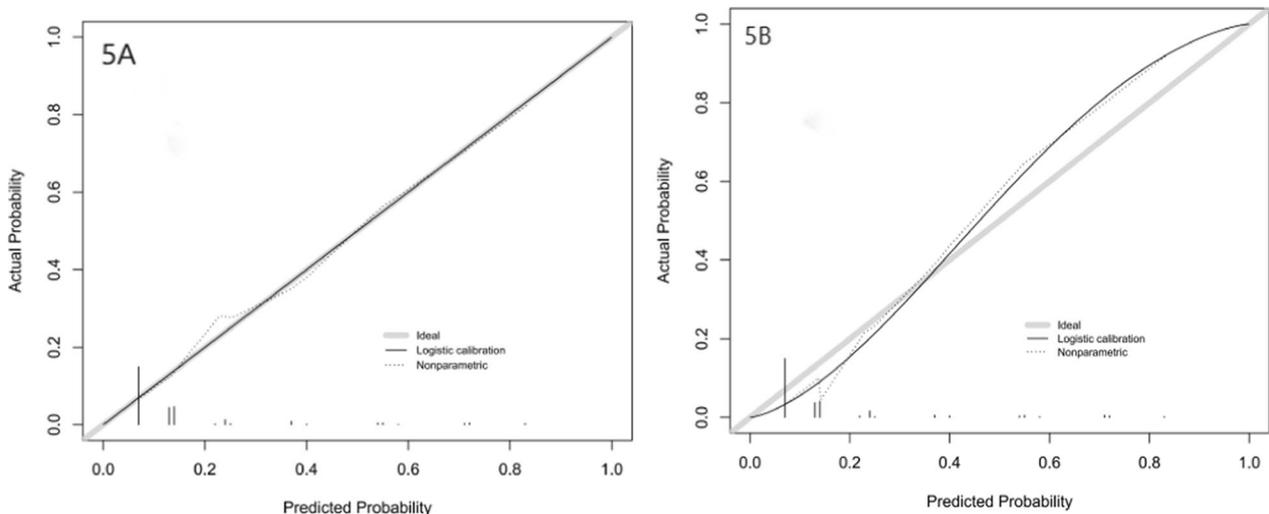


Fig. 5 Evaluation and validation of nomogram calibration curves (A training set, B validation set). Calibration curve was used to evaluate the calibration degree of the nomogram model. **A:** Calibration curve of training set; **B:** Calibration curve of the validation set. The missed diagnosis HSIL+ risk predicted by the model was plotted as the horizontal axis, and the actual missed diagnosis HSIL+ risk was plotted as the vertical axis. The prediction performance in the ideal state was represented by the gray diagonal line, which indicated that the prediction results were completely consistent with the actual situation. The dotted line showed the actual predictive performance of the model, while the black line showed the calibrated predictive performance of the model. The closer the model's prediction curve (black and dotted lines) was to the diagonal, the model was the more accurate

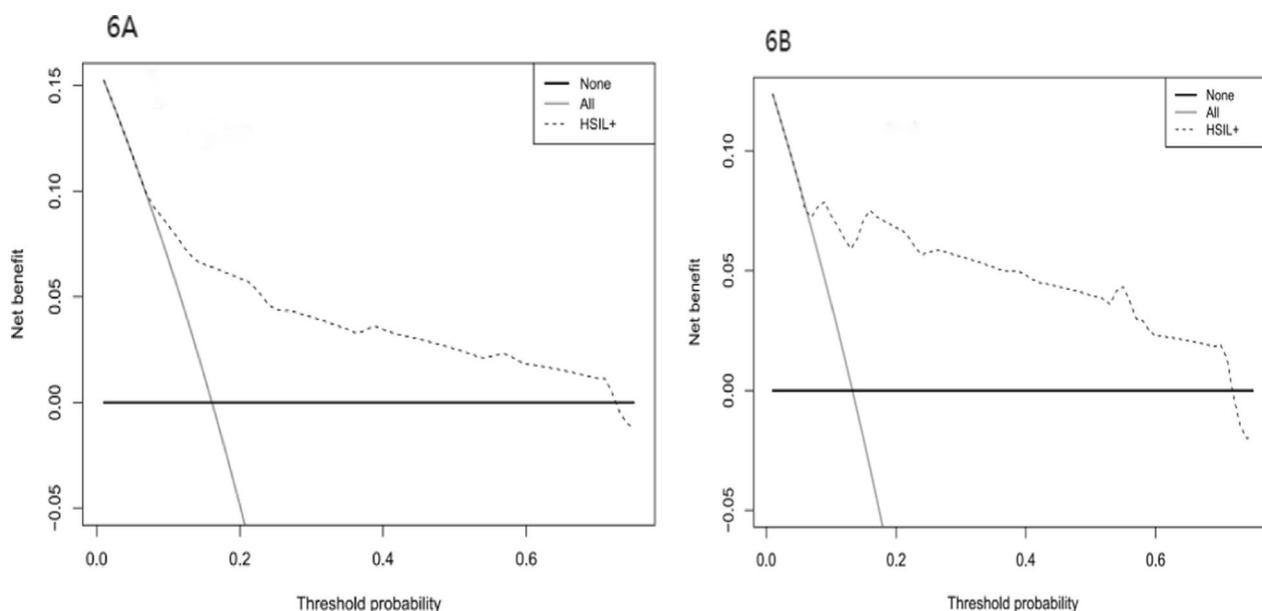


Fig. 6 Evaluation and validation of clinical decision curve of nomogram (**A** training set, **B** validation set). The clinical decision curve was constructed to evaluate the clinical practicability of the prediction model. **A** Clinical decision curve of training set; **B** Clinical decision curve of the validation set. In the DCA curve, the horizontal coordinate represented the threshold probability and the vertical coordinate represented the net benefit rate. The None line represented that none of the samples were patients with missed HSIL+ and that the clinical net benefit rate was 0 without treatment. The ALL line represented the net benefit rate when all samples were missed patients. The dotted line represented the net benefit rate of patients in this column chart model, and the dotted in the upper right, indicating that the model had good clinical utility

those for cytology HSIL. In our cohort, the rate of missed diagnoses of HSIL+ was significantly higher in patients with TCT \geq high-level lesion (66.7% vs. 10.9%). Moreover, patients with TCT \geq ASC-H were found to have a 4.147-fold increased risk of a missed diagnosis of HSIL+. These insights underscore the necessity to enhance the stratified management of previous cytological results in patients with LSIL to mitigate the potential risk of occult HSIL+. Therefore, a more aggressive management approach is warranted for patients with cytologic findings of ASC-H/HSIL.

Transformation zones (TZs) are critical anatomical sites where precancerous lesions and invasive carcinomas typically develop. However, TZ visibility can vary, particularly in postmenopausal women, where age-related changes often retract the TZ into the cervical canal and, combined with epithelial atrophy, can complicate biopsy collection, thereby increasing the risk of missed diagnosis of cervical intraepithelial neoplasia (CIN). Studies have reported varying accuracy rates for detecting CIN2+ across different TZ types: 92.2% for TZ1, 90.5% for TZ2, and 76.5% for TZ3 [19]. In cases with low-grade colposcopic impressions and TZ3, missed diagnosis rates for CIN2 and CIN3 are notably higher (52.6% and 31.6%, respectively) compared to TZ1/2 (27.5% and 18.8%) [20]. In our study, the rates of missed HSIL+ diagnoses were 8.5% for TZ1, 16.9% for TZ2, and 25% for TZ3, indicating

a significant increase in missed diagnoses in patients with TZ3. Caution is therefore necessary when interpreting nonrepresentative biopsies from women with TZ3, and employing endocervical curettage (ECC) may enhance HSIL+ detection in these cases [21].

Colposcopy plays a vital role in cervical cancer screening by localizing lesions in the lower genital tract, guiding biopsies, and informing management strategies. It also assists in the follow-up after treatment. Historical data indicate that the agreement between histopathological diagnosis and colposcopic diagnosis ranges from 52 to 99% [22]. In this study, among 382 patients with a low-grade colposcopic impression, the concordance between cervical biopsy diagnosis and colposcopy was 75.6%, while the agreement between postoperative diagnosis and colposcopy diagnosis was 89%, aligning with prior findings. Given that colposcopy's effectiveness can be influenced by the examiner's experience and is somewhat subjective, approximately 10% of HSIL cases may still be overlooked, even when biopsy pathology under colposcopy indicates LSIL [9]. For patients referred for colposcopy with LSIL, 9.9% were found to have CIN2+ under a low-grade colposcopic impression [23]. A meta-analysis revealed that the overall risk of CIN2+ for women with a low-grade colposcopic impression ranges widely from 11 to 69%, depending on the screening context [24]. Furthermore, 30% of women with grade 2 colposcopy

findings have HSIL/CIN3, irrespective of screening test results [20]. In our study, 11% of women with a low-grade colposcopic impression and 57.1% with a high-grade impression were diagnosed with HSIL+. These findings underscore the importance of standardized, safe, and accurate colposcopy, with quality control being essential to maximize colposcopy's value in preventing and treating cervical cancer.

Machine learning is an important branch of artificial intelligence. In medicine, machine learning technology is changing the way we diagnose, treat, and manage disease. Machine learning can help us extract valuable information from massive medical data, and build predictive models based on the processing data to improve the diagnosis accuracy of diseases, predict the development trend of diseases and optimize treatment plans. Machine learning has brought tremendous insight and predictive power to the medical industry and is increasingly being used in the medical field. For example, Our center has done some research on machine learning and HSIL. Zeng constructed a logistic regression model based on risk factors to predict the risk coefficient of residual lesions after cervical conization in patients with HSIL, The AUC calculated by logistic regression model was 0.78 [25]. This model is being further validated and is expected to be used in clinical practice. Zhang conducted a retrospective analysis of 3343 patients who underwent CKC for HSIL, and then applied seven machine learning methods to construct a positive margin risk prediction model, logistic regression model has the best predictive performance, with an accuracy of 74.7%, sensitivity of 76.7%, specificity of 74.4%, and AUC of 0.826 [26]. Machine learning can often build some better prediction models with improved accuracy and sensitivity due to its diversity and uniqueness of methods and relatively advanced algorithms.

Through univariate and multivariate analyses of collected data, we identified independent risk factors for the missed diagnosis of HSIL+. Utilizing these factors, we constructed predictive models for HSIL+ using machine learning methods. Machine learning, with its diverse methodologies and advanced algorithms, often surpasses traditional logistic regression models in accuracy and sensitivity. Our comparative analysis across the training and validation set revealed that the Decision Tree model demonstrated the most effective predictive performance.

The predictive models we developed offer two significant advantages: simplicity of operation and cost-effectiveness. The model can be programmed and stored on a computer. Clinicians simply input the patient's data for specific indicators, and the program automatically calculates the risk of missed diagnosis of HSIL+. This aids in the early identification of patients at high risk

of missed diagnosis of HSIL+. Additionally, the indicators required by our model are straightforward, derived from routine medical inquiries and examinations, eliminating the need for costly tests, invasive procedures, or any invasion of patient privacy. For patients identified as high-risk through preoperative examinations, clinicians should diligently assess the risk of HSIL+ and consider selective diagnostic coning if necessary. Moreover, such patients should be closely monitored post-surgery to minimize the chances of residual disease and recurrence.

However, our study is not without limitations. Firstly, the sample size is relatively small, characterizing this as a small-sample study. Secondly, instead of using a prospective cohort design and following up with patients predicted to be at high risk, we conducted a retrospective study, which may introduce selection bias. Additionally, the training and validation set data were sourced from the same hospital, which could limit the applicability of our model in other settings with different surgical techniques and protocols.

In conclusion, our study identifies several independent risk factors for the missed diagnosis of HSIL+ in patients with LSIL diagnosed by colposcopic biopsy, including HPV16/18 infection, TCT \geq ASC-H, TZ3, and colposcopic impression G2. The clinical prediction model developed in this study demonstrates robust consistency and practical value, offering significant guidance for clinicians in reducing the risks of missed diagnosis of HSIL+ and the overtreatment of LSIL.

Abbreviations

HSIL	High-grade squamous intraepithelial lesion
LSIL	Low-grade squamous intraepithelial lesion
LEEP	Loopelectrical excision procedure
CAP	American Society for Pathology
ASCCP	American Society for Colposcopy and Cervical Pathology
CIN	Cervical intraepithelial neoplasia
HR-HPV	High-risk human papillomavirus
TCT	Thin cytologic test
AIS	Adenocarcinoma in situ
NILM	No abnormal cells
ASC-US	Atypical squamous cells of unknown significance
ASC-H	Atypical squamous cells of cannot exclude high-grade squamous intraepithelial lesion
ROC	Receiver operator characteristic
AUC	Area under receiver operator characteristic curve
DCA	Decision curve analysis

Acknowledgements

We thank all the patients who participated in this study, their families, and the investigators.

Author contributions

DmL, KmC and YZ had full access to all of the data in the study and take responsibility for the integrity of the data and the accuracy of data analysis. Conception and design: DmL, KmC and YZ. Acquisition, analysis, or interpretation of data: DmL, ZcW, YL, MyZ, B X, YZ. Drafting of the manuscript: DmL, YZ. Critical revision of the manuscript for important intellectual content: DmL, KmC and YZ. Statistical analysis: ZcW, YL. Data Collection: DmL, MyZ, BX, LZ. Supervision: KmC, YZ. All authors read and approved the final manuscript. xz.

Funding

This work was supported by "Jingzhou city science and technology guidance project(2023HC51)" and "Jingzhou City Joint Research Fund Project (2024LHY29)".

Availability of data and materials

No datasets were generated or analysed during the current study.

Declarations

Ethics approval and consent to participate

This study was approved by the Ethics Committee of the First Affiliated Hospital of Yangtze University. Informed consent was obtained for all patients or family members. In addition to this, all methods in the study were performed in accordance with the relevant guidelines and regulations.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Author details

¹Department of Obstetrics and Gynecology, The First Affiliated Hospital of Yangtze University, 8 Hangkong Road, Shashi District, Jingzhou, Hubei, China. ²Oncology Department, The First Affiliated Hospital of Yangtze University, 8 Hangkong Road, Shashi District, Jingzhou, Hubei, China. ³Neurology Intensive Care Unit, The First Affiliated Hospital of Yangtze University, 8 Hangkong Road, Shashi District, Jingzhou, Hubei, China. ⁴Pathology Department, The First Affiliated Hospital of Yangtze University, 8 Hangkong Road, Shashi District, Jingzhou, Hubei, China.

Received: 21 August 2024 Accepted: 28 November 2024

Published online: 05 December 2024

References

- Wei L, Shen D, Zhao F, et al. Abnormal cervical cancer screening in China and management related expert consensus. *Chin J Obstet Gynecol Clin*. 2017;19(3):286–8.
- Zhao C, Bi H, Zhao Y, et al. Chinese expert consensus on the management of cervical high-grade intraepithelial lesions. *Chin J Clin Obstet Gynecol*. 2022;23(2):220–4. <https://doi.org/10.13390/j.issn.1672-1861.2022.02.038>.
- Bi H, Li M, Zhao C, et al. Cervix low grade squamous intraepithelial lesions in the management of China expert consensus 2022;23(4):443–445
- Duesing N, Schwarz J, Choschzick M, et al. Assessment of cervical intraepithelial neoplasia (CIN) with colposcopic biopsy and efficacy of loop electro-surgical excision procedure (LEEP). *Arch Gynecol Obstet*. 2012;286(6):1549–54. <https://doi.org/10.1007/s00404-012-2493-1>.
- McCredie MR, Sharples KJ, Paul C, et al. Natural history of cervical neoplasia and risk of invasive cancer in women with cervical intraepithelial neoplasia 3: a retrospective cohort study. *Lancet Oncol*. 2008;9(5):425–34. [https://doi.org/10.1016/S1470-2045\(08\)70103-7](https://doi.org/10.1016/S1470-2045(08)70103-7).
- World Health Organization. Female genital tumours. IARC. WHO classification of tumours. 5th Edition [EB/OL]. <https://tumourclassification.iarc.who.int/9789283245049>.
- Tanaka Y, Ueda Y, Kakuda M, et al. Predictors for recurrent/persistent high-grade intraepithelial lesions and cervical stenosis after therapeutic conization: a retrospective analysis of 522 cases. *Int J Clin Oncol*. 2017;22(5):921–6. <https://doi.org/10.1007/s10147-017-1124-z>.
- Demarco M, Lorey TS, Fetterman B, et al. Risks of CIN 2+, CIN3+, and cancer by cytology and human papillomavirus status: the foundation of risk-based cervical screening guidelines. *J Low Genit Tract Dis*. 2017;21(4):261–7. <https://doi.org/10.1097/LGT.0000000000000343>.
- Kurman RJ, Carcangiu ML, Herrington CS, et al. WHO classification of tumours of female reproductive organs. 4th eds. Ly• on, IARC Press; 2014. pp. 8–253.
- Perkins RB, Guido RS, Castle PE, et al. 2019 ASCCP risk-based management consensus guidelines for abnormal cervical cancer screening tests and cancer precursors. *J Low Genit Tract Dis*. 2020;24(2):102–31. <https://doi.org/10.1097/LGT.0000000000000525>.
- Alonso I, Torné A, Puig-Tintoré LM, et al. High-risk cervical epithelial neoplasia grade 1 treated by loop electrosurgical excision: follow-up and value of HPV testing. *Am J Obstet Gynecol*. 2007;197(4):359.e1–6. <https://doi.org/10.1016/j.ajog.2007.01.023>.
- Bonde J, Sandri M-T, Gary DS, Andrews JC. Clinical utility of human papillomavirus genotyping in cervical cancer screening: a systematic review. *J Low Genit Tract Dis*. 2020;24:1–13. <https://doi.org/10.1097/LGT.0000000000000494>.
- Schiffman M, Castle PE, Jeronimo J, et al. Human papillomavirus and cervical cancer. *Lancet*. 2007;370(9590):890–907. [https://doi.org/10.1016/S0140-6736\(07\)61416-0](https://doi.org/10.1016/S0140-6736(07)61416-0).
- Clifford G, Franceschi S, Diaz M, et al. Chapter 3: PV type-distribution in women with and without cervical neoplastic diseases. *Vaccine*. 2006;24(Suppl 3):S26–34. <https://doi.org/10.1016/j.vaccine.2006.05.026>.
- Wentzensen N, Walker J, Smith K, et al. A prospective study of risk-based colposcopy demonstrates improved detection of cervical precancers. *Am J Obstet Gynecol*. 2018;218(6):604.e1–604.e8. <https://doi.org/10.1016/j.ajog.2018.02.009>.
- Tidy JA, Lyon R, Ellis K, et al. The impact of age and high-risk human papillomavirus (hrHPV) status on the prevalence of high-grade cervical intraepithelial neoplasia (CIN2+) in women with persistent hrHPV-positive, cytology-negative screening samples: a prospective cohort study. *BJOG*. 2020;127(10):1260–7. <https://doi.org/10.1111/1471-0528.16250>.
- Katki HA, Schiffman M, Castle PE, et al. Benchmarking CIN 3+ risk as the basis for incorporating HPV and Pap contesting into cervical screening and management guidelines. *J Low Genit Tract Dis*. 2013;17(5 Suppl 1):S28–35. <https://doi.org/10.1097/LGT.0b013e318285423c>.
- Egemen D, Cheung LC, Chen X, et al. Risk estimates supporting the 2019 ASCCP risk-based management consensus guidelines. *J Low Genit Tract Dis*. 2020;24(2):132–43. <https://doi.org/10.1097/LGT.0000000000000529>.
- Stuebs FA, Schulmeyer CE, Mehlhorn G, et al. Accuracy of colposcopy-directed biopsy in detecting early cervical neoplasia: a retrospective study. *Arch Gynecol Obstet*. 2019;299(2):525–32. <https://doi.org/10.1007/s00404-018-4953-8>.
- Del PM, Angeles MA, Marti C, et al. Colposcopic impression has a key role in the estimation of the risk of HSIL/CIN3. *Cancers (Basel)*. 2021;13(6):1224. <https://doi.org/10.3390/cancers13061224>.
- Wei B, Li Q, Seery S, et al. Endocervical curettage for diagnosing high-grade squamous intraepithelial lesions or worse in women with type 3 transformation zone lesions: a retrospective, observational study. *BMC Womens Health*. 2023;23(1):245. <https://doi.org/10.1186/s12905-023-02297-0>.
- Li Y, Zhang H, Zheng R, et al. Agreement between colposcopic diagnosis with 2011 international terminology of colposcopy and cervical pathology in cervical lesions. *Chin J Obstet Gynecol*. 2015;5:361–6. <https://doi.org/10.3760/cma.j.issn.0529-567x.2015.05.009>.
- Phianpiset R, Ruengkhachorn I, Jareemit N, et al. ASCCP risk based colposcopy recommendations applied in Thai women with atypical squamous cells of undetermined significance or low-grade squamous intraepithelial lesion cytology. *Obstet Gynecol*. 2020;136(3):510–7. <https://doi.org/10.1097/AOG.0000000000003982>.
- Silver MI, Andrews J, Cooper CK, et al. Risk of cervical intraepithelial neoplasia 2 or worse by cytology, human papillomavirus 16/18, and colposcopy impression: a systematic review and meta-analysis. *Obstet Gynecol*. 2018;132(3):725–35. <https://doi.org/10.1097/AOG.0000000000002812>.
- Zeng Y, Jiang T, Zheng Y, Yang J, Wei H, Yi C, Liu Y, Chen K. Risk factors predicting residual lesion in subsequent hysterectomy following cold knife conization (CKC) for high-grade squamous intraepithelial lesion (HSIL). *BMC Womens Health*. 2022;22(1):358. <https://doi.org/10.1186/s12905-022-01939-z>.
- Zhang L, Zheng Y, Lei L, Zhang X, Yang J, Zeng Y, Chen K. Development of a machine learning-based model for predicting positive margins in high-grade squamous intraepithelial lesion (HSIL) treatment by cold knife conisation (CKC): a single-center retrospective study. *BMC Womens Health*. 2024;24(1):332. <https://doi.org/10.1186/s12905-024-03180-2>.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.