Submitted: 21 September, 2022 Accepted: 08 December, 2022 Published: 15 June, 2024 DOI:10.22514/ejgo.2024.047

ORIGINAL RESEARCH

European Journal of Gynaecological Oncology

CpG methylation of the *PAX1* promoter in exfoliated cervical cells: a potential cervical cancer screening model

Yu Chen[1,†], Yuanyuan Zhang[1,2,†], Xiaochun Chen[1,†], Hong Chu[1], Songqing Guan[1], Qiuni Liang[1], Yuting Xu[1], Xinyue Lv[1], Chunhua Li[1], Yun Li[3], Jin Liu[1], Wenlan Liu[3,*], Zhiying Yu[1,*]

[1]Department of Gynecology, Shenzhen Second People's Hospital/the First Affiliated Hospital of Shenzhen University Health Science Center, 518035 Shenzhen, Guangdong, China
[2]Department of Obstetrics and Gynecology, the Affiliated Hospital of Guizhou Medical University, 550025 Guiyang, Guizhou, China
[3]The Center for Medical Genetics and Molecular Diagnosis, Shenzhen Second People's Hospital/the First Affiliated Hospital of Shenzhen University Health Science Center, 518035 Shenzhen, Guangdong, China

***Correspondence**
2011191208@email.szu.edu.cn (Zhiying Yu);
wlliu@szu.edu.cn (Wenlan Liu)

† These authors contributed equally.

Abstract

Cervical cancer screening is an effective mean of preventing cervical cancer, however, there are limitations of the current screening methods. In this study, an automated high throughput paired box-1 (*PAX1*) methylation detection method was designed by employing matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS), which analyzed 88 *PAX1* methylation sites. In total, 188 subjects were enrolled in the study. The diagnostic efficacy of various *PAX1* methylation sites for detecting cervical intraepithelial neoplasia II or more severe lesions was analyzed and compared with current cervical cancer screening methods. A predictive model was constructed using logistic regression. Five sites with the largest area under receiver operating characteristic curve values were GRCh38/hg38, chr20: X21705687, X21705946, X21706427, X21706285 and X21706637. *PAX1* methylation method had higher diagnostic efficacy than high risk human papillomavirus (hr-HPV) and ThinPrep cytology tests. A predictive model combining hr-HPV and *PAX1* methylation sites was constructed as follows: prediction index = −5.993 + 15.211 × X21705687 + 7.890 × X21706427 + 1.846 × HPV16/18 (1,0) + 1.821 × other hr-HPV (1,0). Area under the curve was 0.868 and sensitivity and specificity were 0.863 and 0.756, respectively. Analysis of *PAX1* promoter methylation alone or in combination with hr-HPV was a promising approach for cervical cancer screening. The study was registered in Chinese Clinical Trial Registry (registration number: ChiCTR2000029231, http://www.chictr.org.cn/index.aspx).

Keywords

Cervical cancer screening; MALDI-TOF MS; Methylation; *PAX1*

1. Introduction

Cervical cancer is the frequently diagnosed gynecological cancer and leading cause of gynecological cancer-related deaths in China and worldwide [1, 2]. It originates in the transformation zone of cervical epithelial tissue and predominantly caused by human papillomavirus (HPV), making it preventable and detectable in pre and early cancer stages [3, 4]. In 2018, World Health Organization announced a global call for action to eliminate cervical cancer. HPV vaccines are effective, however cervical cancer screening is the most effective for preventing cervical cancer. In recent decades, the incidence and mortality of cervical cancer have decreased in affluent countries. In urban areas of China, the threat posed by cervical cancer has declined because of cervical cancer screening [5]. Conversely, never- and under-screenings are common in underprivileged than in privileged populations which has led to the increased incidence and mortality of invasive cervical cancer [6, 7]. Hence, making cervical cancer screening more accessible, reliable, and cost-effective remains a global challenge.

Currently, the common cervical cancer screening methods are liquid-based cytology and HPV DNA testing [8]. However, both methods have limitations. HPV DNA testing is sensitive but lacks specificity as ~80% women are infected with HPV at some point in their lives. Hence, HPV DNA testing requires economic and medical resources which burden the patients with extra hospital visits and psychological distress [9]. Cytological testing has higher specificity than HPV DNA testing, however, it depends on the expertise of pathologists, and accuracy varies from doctor to doctor, making it unsuitable for regions lacking skilled pathologists [10].

PAX1 gene is a tumor suppressor gene located on the short arm of chromosome 20 [11]. A previous study demonstrated that *PAX1* expression inhibited the phosphorylation of multiple kinases including the dual specificity phosphatase 1, 5 and 6, and of epidermal growth factor/mitogen-activated protein kinase signaling [12]. Numerous studies have shown that *PAX1*

This is an open access article under the CC BY 4.0 license (https://creativecommons.org/licenses/by/4.0/).

methylation is related to cervical cancer, however, the clinical significance of *PAX1* methylation detection as cervical cancer screening method remains controversial. *PAX1* methylation testing accuracy fluctuates from 30% to 85% [13]. Literature review revealed that previous studies used methylation-specific polymerase chain reaction (MSP) which restricted detection to few methylation sites. When a methylation site relevant to cervical cancer is tested, the diagnostic value increases, and *vice versa*. The limitations of detection method may contribute to variations in the accuracy.

MALDI-TOF MS can be used to quantify DNA methylation and simultaneously detect multiple methylation sites. Herein, an automated high throughput *PAX1* methylation detection method was designed by using MALDI-TOF MS. The methylation status of 88 sites was analyzed for identifying the most relevant ones and compared the diagnostic efficacy with that of hr-HPV DNA test and ThinPrep cytology test (TCT). Finally, a predictive model was constructed by combining the automatic detection methods for hr-HPV DNA and *PAX1* methylation. The diagnostic efficiency of this automated detection method was substantially better than the traditional cervical cancer screening methods.

2. Materials and methods

2.1 Patient selection

Patients included for this study had abnormal cervical cancer screening results and had undergone cervical biopsy during colposcopy in Department of Gynecology at Shenzhen Second People's Hospital between October 2019 and April 2021. Each patient was asked to fill in a questionnaire for assessing whether they met the inclusion or exclusion criteria and to add other relevant characteristics. The inclusion criteria were as follows: (1) 18–60 years' age; (2) abnormal findings detected for the first time by cervical cancer screening; and (3) patients who were about to undergo cervical biopsy due to abnormal colposcopy findings. Exclusion criteria included: (1) a history of uterine or cervical surgery; (2) patients with other malignant tumors; and (3) women undergoing or within 42 days of pregnancy termination.

Ultimately, 188 participants were enrolled in the study. The study conformed to The Code of Ethics of the World Medical Association (Declaration of Helsinki), published by British Medical Journal (1964).

Cervical biopsy is the gold standard for pathological diagnosis. Samples included 108 inflamed cervical epithelia (non-malignant), 78 cervical intraepithelial neoplasms (CINs), and two cervical cancers. The 78 CINs comprised of 29 CIN I (less than 1/3 abnormal cells), 21 CIN II (1/3 to 2/3 abnormal cells), and 28 CIN III (2/3 to full cervical surface layer thickness) lesions.

2.2 Sample collection and DNA extraction

Exfoliated cervical epithelial cells were obtained by following the colposcopic examination prior to cervical biopsy. After exposing the cervix, blood and vaginal discharge were gently wiped off. Cytology swabs were inserted into the cervical squamocolumnar junction and rotated clockwise for five to ten times. The cells were immediately stored in cell preservation solution (Cat# 51306; Qiagen, Hilden, Germany) at 4 °C until DNA extraction.

2.3 PAX1 methylation detection

DNA from cervical epithelial cells was extracted using QIAmp Tissue Kit (Qiagen). Cytosines were converted to uracils using EZ DNA Methylation-Gold Kit (Zymo Research, Irvine, CA, USA). Methylated cytosines remained unmodified during the process. According to the University of California Santa Cruz genome browser database (http://genome.ucsc.edu/cgi-bin/hgGateway), the methylation sites in *PAX1* promoter contained 141 5-cytosine-phosphate-guanine-3 (CpG) islands (chr20: 21705562–21707051) (Fig. 1A). Methylation specific primers were designed using EpiDesigner (https://epidesigner.com/start3.html, Agena Bioscience, San Diego, CA, USA). Five primer pairs were used to cover 88 CpG islands: (1) left: ATTTTTAAAGTTAGAGGAGGAAGGG, right: CCCAAACCCAAAATAAACTTCATC, CpG coverage: 18; (2) left: ATGAAGTTTATTTTGGGTTTGGG, right: CCAAACTTACTCAACTACCCCCTA, CpG coverage: 34; (3) left: TTTAGTTTTGGGTTTTGGAGAAGTT, right: TTCCAAAAATAACCTATAAATCCCC, CpG coverage: 10; (4) left: GAGTTGTTTGGGGATTTATAGGTTA, right: CAAAAATACCAAAATCTCCTTACTT, CpG coverage: 30; and (5) left: ATTTTGGTATTTTTGTTTGGGAGAT, right: CTAACCAATATAAAACCCTCCCCTA, CpG coverage: 38 (Fig. 1B). Exfoliated cervical cell DNA was extracted by enzymatic lysis, transformed with bisulfite, and amplified using PCR. Base-specific cleavage was performed using homogeneous Mass Cleave assay. The 88 sites were cut into 52 pieces. The base fragments were analyzed using MassARRAY time-of-flight MS (Agena Bioscience).

2.4 Liquid-based cytology

Exfoliated cervical epithelial cells were obtained as described in Section 2.2. Samples for cytology were prepared according to the standard ThinPrep (Hologic, Marlborough, MA, USA) protocol. The samples were evaluated by two experienced pathologists. The results were classified as negative intraepithelial lesion or malignancy, low-grade squamous intraepithelial lesion, or high-grade squamous intraepithelial lesion, when both pathologists agreed with the cell morphology. Atypical squamous cells of undetermined significance and atypical squamous cells categorized as high-grade, that could not be excluded based on suspected underlying lesion (low-grade squamous intraepithelial lesion *vs*. high-grade squamous intraepithelial lesion, respectively), were used when pathologists were uncertain or in disagreement with the results.

2.5 HPV DNA testing

HPV DNA was amplified using Hema 9600 gene amplification instrument (Hema Medical Instrument Co. Ltd., Zhuhai, China). YN-H18 automatic nucleic acid molecular hybridization instrument (Yaneng Biosciences, Guangdong, China) was employed for detecting 18 hr-HPV genotypes,

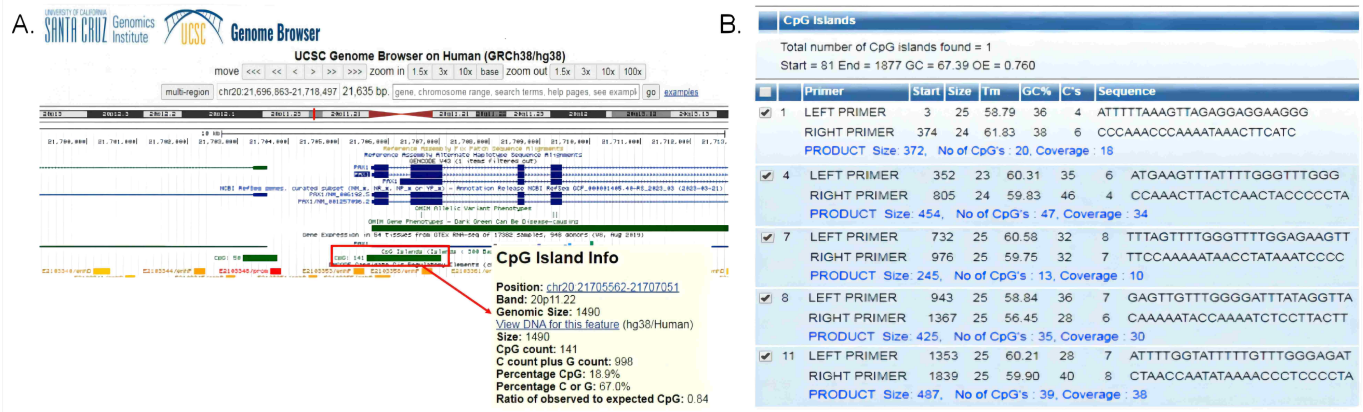

FIGURE 1. ***PAX1*** **promoter analysis.** (A) Methylation of CpG islands in the University of California Santa Cruz genome browser database (http://genome.ucsc.edu/cgi-bin/hgGateway). (B) Five pairs of methylation specific primers were designed using EpiDesigner software (Agena Bioscience, San Diego, CA, USA). Abbreviations: CpG: 5-cytosine-phosphate-guanine-3; *PAX1*: paired box-1.

including HPV16, 18, 31, 33, 35, 39, 45, 51, 52, 53, 56, 58, 59, 66, 68, 73, 82 and 83.

2.6 Tissue collection and pathological diagnosis

All abnormal lesions were evaluated according to the American Society for Colposcopy and Cervical Pathology and the Chinese Society for Colposcopy and Cervical Pathology guidelines. The lesions were biopsied by cervical disease specialists. Tissues were fixed in 4% formaldehyde. Paraffin embedded tissues were cut into 5-μm sections for hematoxylin and eosin staining.

Pathological slides were analyzed by the two experienced pathologists. If two pathologists disagreed, a third pathologist examined the slides and performed immunohistochemical staining as supporting method, if necessary. The tissues samples in this study were fairly typical and above scenario was not created.

2.7 Statistical analysis

The data were analyzed using SPSS Statistics 27.0.1 (SPSS Inc., Chicago, IL, USA). Figs were plotted using GraphPad Prism 9.4.0 (GraphPad Software, San Diego, CA, USA). Quantitative variables were expressed as mean ± standard deviation (SD) or median (interquartile range). Between-group differences were assessed using *t*-tests or one-way analysis of variance for normally distributed variables and non-parametric tests for non-normally distributed variables. The Chi-square test or Fisher's exact test was used to analyze the qualitative data.

Receiver operating characteristic (ROC) curves were plotted, and cut-off values were determined using maximum Youden index. The sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) were calculated based on optimal cut-off values. The predictive model was constructed using multivariate logistic regression and backward elimination. The model was internally evaluated using ROC curves.

3. Results

3.1 Participant characteristics

A total of 188 patients were included in this study and placed in five groups based on their pathological diagnoses: 108 cases of inflamed cervical epithelia (non-malignant), 29 of CIN I, 21 of CIN II, 28 of CIN III, and 2 of cervical cancer (Fig. 2). The baseline characteristics of patients are listed in Table 1. The median (interquartile range) age of enrolled patients was 35 (29–42) years. Differences in parity, hr-HPV counts, and TCT results were observed among the patients ($p = 0.027$, $p < 0.01$, and $p < 0.01$, respectively). No significant differences were found in age, weight, height, the onset of menstruation, or gravity. HPV52 (19.7%) was the most prevalent HPV genotype, followed by HPV16 (17.0%) and HPV58 (13.3%). This was consistent with the overall distribution of HPV genotypes in China but inconsistent with the most prevalent genotypes worldwide (HPV16 and HPV18) [14, 15]. The detailed distribution of HPV genotypes in relation to pathological diagnoses is described in **Supplementary Table 1**.

3.2 Diagnostic value of hr-HPV and TCT

According to cervical cancer screening guidelines, HPV16/18 testing alone or combined with other genotypes can be used as an independent screening method. TCT can be employed as an independent screening method or as a secondary screening method for hr-HPV-positive patients [16]. The diagnostic efficacy of these methods was evaluated. At least one positive test for either HPV16 or HPV18 was considered positive for HPV16/18 test. Patients tested positive for at least one hr-HPV genotype were considered hr-HPV-positive. A negative intraepithelial lesion or malignancy was defined as TCT-negative. Patients with equal to or more severe than atypical squamous cells of undetermined significance were considered TCT-positive.

The participants were divided into control and patient groups as per their pathological diagnosis. CIN II is the commonly accepted threshold for treatment and its detection is the primary

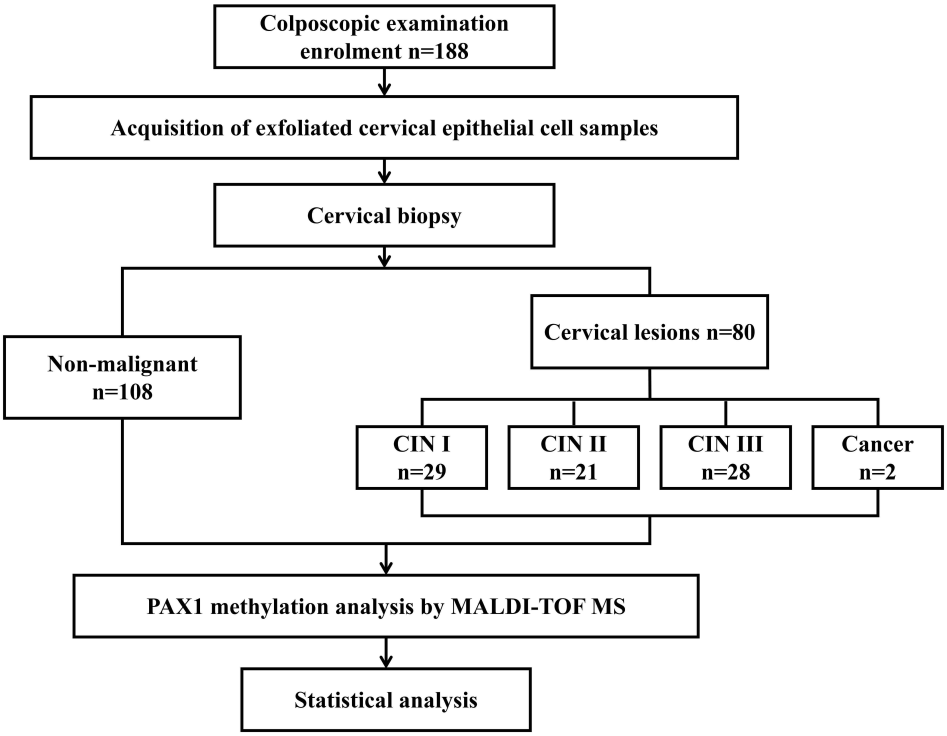

FIGURE 2. Study flowchart. Women requiring cervical biopsy were evaluated to determine whether they met the criteria of this study. Before biopsy, exfoliated cervical epithelial cells were obtained. Patients were diagnosed as having inflamed cervical epithelia (non-malignant), cervical intraepithelial neoplasia CIN I, CIN II, CIN III, or cervical cancer based on pathological analysis. Abbreviations: MALDI-TOF MS: matrix-assisted laser desorption/ionization time-of-flight mass spectrometry; *PAX1*: paired box-1.

goal of cervical cancer screening [16, 17]. Participants diagnosed with inflamed cervical epithelia (non-malignant) and CIN I were categorized as the control group. Participants diagnosed with CIN II or above were placed in the patient group. The sensitivity, specificity, PPV, NPV and accuracy of hr-HPV DNA testing are shown in Table 2.

3.3 Diagnostic value of PAX1 promoter methylation

Eighty-eight *PAX1* methylation sites were evaluated for each sample and their methylation percentages were recorded (Fig. 3). After classifying participants to control and patient groups, ROC curves were plotted for each methylation site (Fig. 4 and **Supplementary Fig. 1**). The area under ROC curve (AUC) and Spearman correlation between *PAX1* methylation and pathological diagnosis were analyzed. GRCh38/hg38 chr20: X21705687 (X21705687) had the largest AUC (0.792). The optimal cut-off values for five sites with the largest AUCs and their corresponding sensitivity, specificity, PPV, NPV, and accuracy are shown in Table 3. *PAX1* methylation status of these sites had higher diagnostic efficacy levels than conventional screening methods. The hr-HPV genotypes had higher sensitivity, however their specificity of 0.299 was considered unsatisfactory as a false-positive result may consume unnecessary resources to perform further colposcopy and biopsy. In contrast, HPV16/18 had relatively high specificity but its low sensitivity of 0.333 was ineffective for cervical cancer screening and prevention.

3.4 Predictive model construction and evaluation

Based on the high sensitivity of hr-HPV, high specificity of HPV16/18, and advantages of automated detection process in clinical practice, a predictive model was constructed using multivariate logistic regression with backward elimination that combined HPV testing and *PAX1* methylation. Using Chi-square test, HPV16/18 ($p = 0.003$) and other hr-HPV genotypes ($p = 0.034$) were confirmed to be associated with pathological diagnosis. *PAX1* methylation sites associated with pathological diagnosis, as well as HPV16/18 and other hr-HPV genotypes, were initially included in the regression model. By gradually eliminating the variables having little effect on regression model until all variables having significant effect on regression coefficient, it was determined that HPV16/18, other hr-HPV genotypes, X21705687, and X21706427 had significant impact on regression model ($p < 0.05$) which allowed the prediction index (PI) to reach its maximum predictive value. The logistic regression equation was as follows:

$$PI = lnP/(1–P) = -5.933 + 15.211 \times X21705687 + 7.890 \times X21706427 + 1.846 \times HPV16/18(1,0) + 1.821 \times other\ hr - HPV\ (1,0)\,. \quad (1)$$

TABLE 1. Participant characteristics.

Pathology	Non-malignant (n = 108)	CIN I (n = 29)	CIN II (n = 21)	CIN III (n = 28)	Cervical cancer (n = 2)	Total (n = 188)	*p*-value
Age (yr)	36.0 (30.0–43.0)	30.0 (26.0–35.5)	32.0 (27.0–39.5)	35.5 (29.3–48.5)	44.5 (39.0–50.0)	35.0 (29.0–42.0)	0.118
Height (cm)	160.0 (156.0–162.0)	163.0 (158.5–165.0)	160.0 (155.5–65.5)	160.0 (154.3–165.0)	162.0 (160.0–164.0)	160.0 (156.0–163.0)	0.104
Weight (kg)	53.0 (49.0–57.0)	54.0 (47.0–60.8)	57.8 (48.5–63.5)	56.5 (63.0–60.0)	65.5 (53.0–60.0)	53.0 (49.0–59.0)	0.604
Age at menarche (yr)	13.0 (12.0–14.0)	13.0 (12.0–14.0)	13.0 (13.0–14.0)	13.0 (12.0–14.0)	13.0 (13.0–13.0)	13.0 (12.0–14.0)	0.696
Gravidity							
0	22 (20.4%)	10 (34.5%)	10 (47.6%)	7 (25.0%)	0 (0.0%)	49 (26.1%)	
1	19 (17.6%)	5 (17.2%)	0 (0.0%)	4 (14.3%)	0 (0.0%)	28 (14.9%)	0.132
>1	67 (62.1%)	14 (48.2%)	11 (52.5%)	17 (60.7%)	2 (100.0%)	111 (59.0%)	
Parity							
0	31 (28.7%)	16 (55.2%)	11 (52.4%)	8 (28.6%)	0 (0.0%)	66 (35.1%)	
1	42 (38.9%)	8 (27.6%)	4 (19.0%)	7 (25.0%)	0 (0.0%)	61 (32.4%)	0.027*
>1	35 (32.4%)	5 (17.2%)	6 (28.6%)	13 (46.5%)	2 (100.0%)	61 (32.4%)	
Contraception							
None	26 (24.1%)	7 (24.1%)	7 (33.3%)	7 (25.0%)	0 (0.0%)	47 (25.0%)	
Condoms	45 (41.7%)	14 (48.3%)	11 (52.4%)	11 (39.3%)	1 (50.0%)	82 (43.6%)	0.378
Other	32 (29.6%)	7 (24.1%)	2 (9.6%)	6 (21.4%)	0 (0.0%)	47 (25.0%)	
Unknown	5 (4.6%)	1 (3.4%)	1 (4.8%)	4 (14.3%)	1 (50.0%)	12 (6.4%)	
hr-HPV count							
0	39 (36.1%)	2 (6.9%)	0 (0.0%)	1 (3.6%)	0 (0.0%)	42 22.3%	
1	62 (57.4%)	18 (62.1%)	16 (76.2%)	22 (78.6%)	1 (50.0%)	119 63.3%	<0.010**
>1	7 (5.7%)	9 (31.0%)	5 (23.9%)	5 (17.9%)	1 (50.0%)	27 14.4%	
TCT							
NILM	76 (70.4%)	17 (58.6%)	10 (47.6%)	8 (28.6%)	1 (50.0%)	112 (59.6%)	
ASC-US	18 (16.7%)	7 (24.1%)	6 (28.6%)	7 (25.0%)	0 (0.0%)	38 (20.2%)	
L-SIL	12 (11.1%)	3 (10.3%)	3 (14.3%)	3 (10.7%)	1 (50.0%)	22 (11.7%)	<0.010**
ASC-H	1 (0.9%)	1 (3.4%)	1 (4.8%)	4 (14.3%)	0 (0.0%)	7 (3.7%)	
H-SIL	1 (0.9%)	1 (3.4%)	1 (4.8%)	6 (21.4%)	0 (0.0%)	9 (4.8%)	

**p < 0.05, **p < 0.01 (Kruskal-Wallis or Fisher's exact test). Abbreviations: ASC-H: atypical squamous cells of high-grade; ASC-US: atypical squamous cells of undetermined significance; CIN: cervical intraepithelial neoplasm; hr-HPV: high-risk human papillomavirus; L/H-SIL: low/high-grade squamous intraepithelial lesion; NILM: negative intraepithelial lesion or malignancy; TCT: ThinPrep cytology test.*

TABLE 2. Diagnostic efficacy of hr-HPV and TCT.

	Sensitivity	Specificity	PPV	NPV	Accuracy
hr-HPV	0.980	0.299	0.342	0.976	0.484
HPV16/18	0.333	0.861	0.472	0.776	0.718
TCT	0.627	0.679	0.421	0.830	0.665

Abbreviations: hr-HPV: high-risk human papillomavirus; NPV: negative predictive value; PPV: positive predictive value; TCT: ThinPrep cytology test.

TABLE 3. Diagnostic value of *PAX1* promoter methylation.

	AUC	95% CI	Optimal cut-off	Sensitivity	Specificity	PPV	NPV	Accuracy	*p*-value
X21705687	0.792	0.719–0.866	0.125	0.843	0.622	0.457	0.913	0.683	<0.0001
X21705946	0.792	0.718–0.865	0.145	0.725	0.730	0.500	0.877	0.729	<0.0001
X21706427	0.790	0.721–0.859	0.175	0.608	0.825	0.564	0.850	0.766	<0.0001
X21706285	0.783	0.705–0.861	0.155	0.647	0.781	0.524	0.856	0.745	<0.0001
X21706637	0.777	0.698–0.855	0.165	0.745	0.696	0.481	0.879	0.710	<0.0001

Abbreviations: AUC: area under the curve; CI: confidence interval; NPV: negative predictive value; PPV: positive predictive value.

TABLE 4. Diagnostic efficacy of predictive model.

	AUC	95% CI	Optimal cut-off	Sensitivity	Specificity	PPV	NPV	Accuracy	*p*-value
Model	0.868	0.817–0.920	0.300	0.843	0.756	0.566	0.927	0.780	<0.0001

Abbreviations: AUC: area under the curve; CI: confidence interval; NPV: negative predictive value; PPV: positive predictive value.

For HPV16/18 and other hr-HPV genotypes, 1 = positive for one or more HPV genotypes; and 0 = negative for one or more HPV genotypes. AUC for this predictive model was 0.868 (95% confidence interval: 0.817–0.920) (Fig. 5). The cut-off, sensitivity, specificity, PPV, NPV, and accuracy were 0.300, 0.843, 0.756, 0.566, 0.927 and 0.780, respectively (Table 4). The variables adopted by this model could be inserted automatically; they were quantifiable, stable, and independent of the pathologist's interpretation. Internal validation of predictive model showed higher accuracy than conventional screening methods which proved its clinical value.

4. Discussion

DNA methylation is a major form of epigenetic regulation and regarded as a symbolic event in carcinogenesis. It occurs when methyl group is added to cytosine by DNA methyltransferase to form CpG dinucleotide [13]. DNA methylation plays role in the development of cervical cancer, for example, the methylation of *PAX1*, sex-determining region Y-box 1, hsa-miR-124, telomerase reverse transcriptase, LIM homeobox transcription factor 1A [18], septin 9 [19], Wilms' tumor 1, NK6 transcription factor related locus 1, deleted in bladder cancer 1 [20], junctional adhesion molecule 3 [21], erythrocyte membrane protein band 4.1 like 3, cell adhesion molecule 1, and maturation associated protein [22]. Among these, *PAX1* methylation is the most correlated with CIN progression and cervical carcinogenesis [11]. *PAX* gene family interacts with homeobox gene family and is critical for embryonic tissue development and cellular differentiation [23, 24]. The mechanism of *PAX1* methylation in cancer development has yet not been elucidated. Compelling evidence has suggested that *PAX1* methylation could have diagnostic potential in cancer screening, especially for cervical cancer. Previous studies have examined the methylation status of *PAX1* in cervical samples, however the diagnostic value of this data was found inconsistent as the sensitivity ranged from 0.640–0.941 and the specificity from 0.600–0.910, for diagnosing CIN II and higher or CIN III and higher-grade lesions [11, 13, 25–32].

The approaches used to detect DNA methylation can be categorized based on their pretreatment methods like restriction enzyme digestion, bisulfite conversion, and affinity enrichment. Mainstream DNA methylation detection methods are based on bisulfite conversion including MSP, digital MSP, nucleic acid MS, methylation chips, and first-, second-, or third-generation sequencing [33]. MALDI-TOF MS is an analytical technique in which a laser irradiates the crystalline film formed between sample and matrix to ionize biomolecules in the sample. The charged biomolecules move under an electric field, and their mass-to-charge (m/z) ratios are calculated by measuring their movement speed [34]. Compared with PCR and methylation chips, MALDI-TOF MS has accurate absolute quantification, higher throughput, and higher detection stability. Previous studies have largely used MSP as the detection method. Most studies did not clearly state the exact methylation sites. Only a scant number of methylation sites were analyzed because of the limitations of approach. Time-of-flight MS was used to detect *PAX1* methylation in 122 cervical samples, in which 20 CpG islands were analyzed [35]. The lack of absolute quantification and coverage of methylation sites may result in large

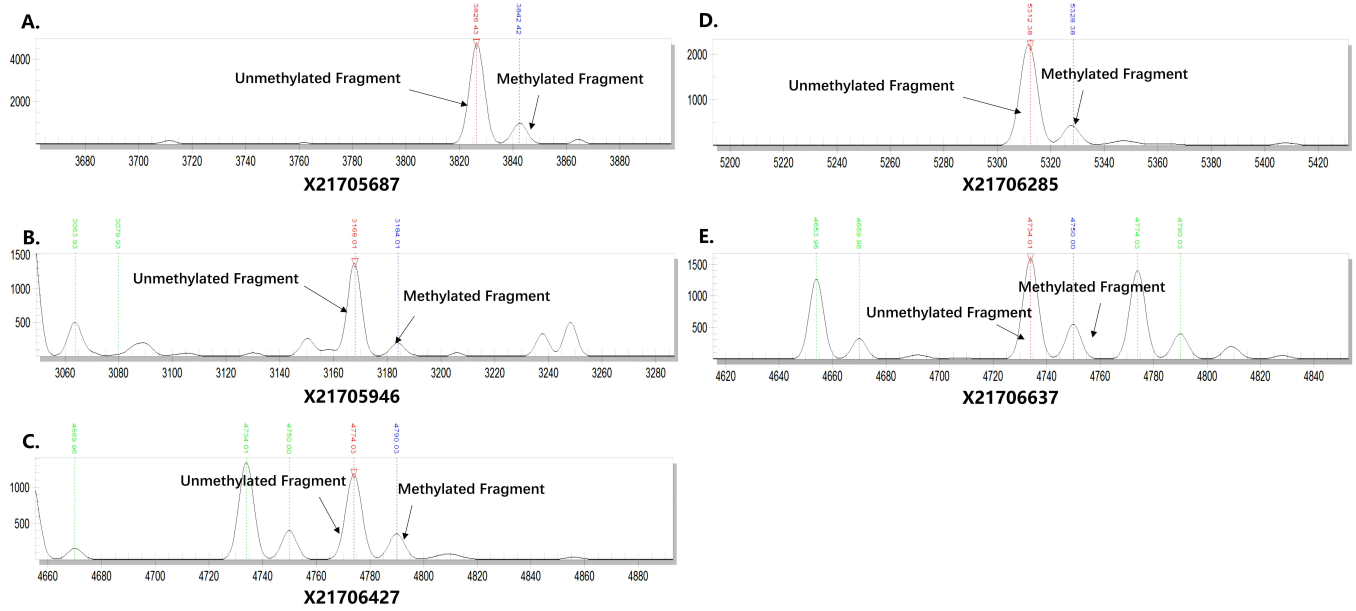

FIGURE 3. Representative mass spectra for *PAX1* methylation sites. Mass spectra of (A) X21705687, (B) X21705946, (C) X21706427, (D) X21706285, and (E) X21706637. The abscissa is mass-to-charge ratio and ordinate is the ion abundance. Abbreviation: *PAX1*: paired box-1.

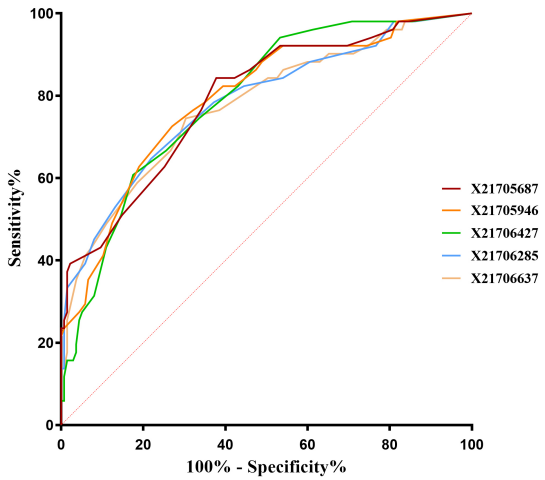

FIGURE 4. ROC curves of five methylation sites with the highest AUCs. Abbreviations: AUC: area under the curve; ROC: receiver operating characteristic.

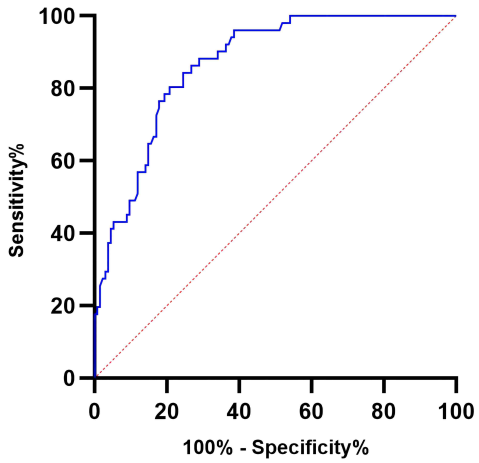

FIGURE 5. ROC curve of the predictive model. Abbreviation: ROC: receiver operating characteristic.

differences of sensitivity and specificity in previous studies. In addition, owing to the lack of comprehensive coverage of *PAX1* methylation sites and accurate methylation estimation, these results have less referential value for guiding the design of primers used to detect *PAX1* promoter methylation with MSP or MALDI-TOF MS.

There were 141 methylation sites determined in *PAX1* promoter. A total of 88 methylation sites were detected using this method as some sites were lost because of inconsistent coverage of CpG islands by each primer pair. The diagnostic efficacy of each site was evaluated as a cervical cancer screening method and determined that X21705687 had the highest AUC of 0.792. To the best of our knowledge, this is the first time that AUC values for specific *PAX1* methylation sites have been reported. Although the sensitivity of *PAX1* was marginally lower than that of hr-HPV, the false-positive rate of hr-HPV was as high as 0.701. When used as an independent screening method, hr-HPV can lead to unnecessary colposcopies and cervical biopsies which increases the pain and psychological burden on patients and wastes medical resources. Although HPV16/18 testing has high specificity, HPV16 and HPV18 are not the only prevalent genotypes among Chinese women which result in low sensitivity [14]. Consequently, if only HPV16/18 testing is used, many patients will not be diagnosed and may develop serious lesions. Thus, our results do not support independent HPV testing as the sole means of cervical cancer screening. In contrast to HPV and TCT testing, *PAX1* methylation has high and stable sensitivity and specificity which makes it an ideal cervical cancer screening marker.

Despite the low accuracy of HPV testing as independent screening method, we have noted the high sensitivity of hr-HPV and high specificity of HPV16/18 testing. This information has been used to construct a predictive model for improving the diagnostic efficacy of cervical cancer screening. Both tests are combined using logistic regression to build a predictive model. This model improves the specificity and accuracy when compared with *PAX1* methylation alone. All variables included in the model can be automatically detected. This provides more reliable method than TCT or other similar cervical cytology tests. Our method is independent of human factors and is thus suitable for regions lacking skilled pathologists. Another model with marginally higher diagnostic efficacy was also constructed as follows:

$$PI = -5.911 + 13.172 \times X21705687 + 3.498 \times X21705946 + 6.371 \times X21706427 + 1.834 \times HPV16/18(1,0) + 1.791 \times other\ hr - HPV\ (1,0) \quad (2)$$

AUC of this model was 0.869, the optimal cut-off value was 0.273, and sensitivity, specificity, PPV, NPV, and accuracy were 0.863, 0.756, 0.571, 0.936 and 0.785, respectively. Neither X21705946 (p = 0.433) nor X21706427 (p = 0.145) significantly contributed to this model, and thus was not adopted.

5. Conclusions

The diagnostic efficacy of *PAX1* methylation was higher than those of HPV DNA testing and TCTs. X21705687 had the largest AUC value compared to other methylation sites. The predictive model combining HPV testing and *PAX1* methylation improved the accuracy of cervical cancer predictions. Although further evaluations will be required before this or similar models can be practically applied in cervical cancer screening. HPV52 was the most prevalent genotype among the patients included in this study. This is consistent with the distribution of HPV genotypes among Han Chinese women (the major nationality in China), whereas it is in contrast to global data that has identified HPV16 and HPV18 as the most prevalent genotypes. Therefore, we cannot rule out the possibility of varying results in different populations, which is a major limitation of this study [14, 15]. Nonetheless, the presented approach offers a promising alternative to traditional cervical cancer screening methods.

ABBREVIATIONS

AUC, area under the curve; CIN, cervical intraepithelial neoplasm; CpG, 5-cytosine-phosphate-guanine-3; HPV, human papillomavirus; hr-HPV, high-risk HPV; MALDI-TOF MS, matrix-assisted laser desorption/ionization time-of-flight MS; MS, mass spectrometry; MSP, methylation-specific PCR; NPV, negative predictive value; *PAX1*, paired box-1; PI, prediction index; PPV, positive predictive value; ROC, receiver operating characteristic; TCT, ThinPrep cytology test.

AVAILABILITY OF DATA AND MATERIALS

The datasets used and/or analyzed during the current study are available from our data administrator Yu Chen (chenyu20171@email.szu.edu.cn) upon reasonable request.

AUTHOR CONTRIBUTIONS

ZYY and WLL—designed and supervised the investigation. YC and YYZ—recruited participants and analyzed the data. YC and XCC—wrote the manuscript. YTX, CHL, XYL and YL—performed *PAX1* methylation detection. HC, SQG, QNL and JL—collected samples. All authors contributed to the editorial changes in the manuscript. All authors read and approved the final manuscript.

ETHICS APPROVAL AND CONSENT TO PARTICIPATE

The study was approved by the Ethics Committee of the Shenzhen Second People's Hospital (approval number: KS20190521002-FS2019052903) and registered in the Chinese Clinical Trial Registry (registration number: ChiCTR2000029231).

ACKNOWLEDGMENT

We would like to thank all the participants that were enrolled in this study.

FUNDING

This research was funded by Shenzhen Science and Technology Program (JCYJ20180228163529609); Sanming Project of Medicine in Shenzhen (SZSM201812041); Clinical Research Funding from Shenzhen Second People's Hospital/the First Affiliated Hospital of Shenzhen University Health Science Center (20193357011).

CONFLICT OF INTEREST

The authors declare no conflict of interest.

SUPPLEMENTARY MATERIAL

Supplementary material associated with this article can be found, in the online version, at https://oss.ejgo.net/files/article/1801487262440931328/attachment/Supplementary%20material.docx.

REFERENCES

[1] Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, *et al*. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA: A Cancer Journal for Clinicians. 2021; 71: 209–249.

[2] Arbyn M, Weiderpass E, Bruni L, de Sanjosé S, Saraiya M, Ferlay J, *et al*. Estimates of incidence and mortality of cervical cancer in 2018: a worldwide analysis. The Lancet Global Health. 2020; 8: e191–e203.

[3] Burd EM. Human papillomavirus and cervical cancer. Clinical Microbiology Reviews. 2003; 16: 1–17.

[4] Reich O, Regauer S, McCluggage WG, Bergeron C, Redman C. Defining the cervical transformation zone and squamocolumnar junction. International Journal of Gynecological Pathology. 2017; 36: 517–522.

[5] Saslow D, Solomon D, Lawson HW, Killackey M, Kulasingam SL, Cain J, *et al*. American Cancer Society, American Society for Colposcopy and Cervical Pathology, and American Society for Clinical Pathology screening guidelines for the prevention and early detection of cervical cancer. CA: A Cancer Journal for Clinicians. 2012; 62: 147–172.

[6] Jolidon V, De Prez V, Willems B, Bracke P, Cullati S, Burton-Jeangros C. Never and under cervical cancer screening in Switzerland and Belgium: trends and inequalities. BMC Public Health. 2020; 20: 1517.

[7] Landy R, Sasieni PD, Mathews C, Wiggins CL, Robertson M, McDonald YJ, *et al*. Impact of screening on cervical cancer incidence: a population-based case—control study in the United States. International Journal of Cancer. 2020; 147: 887–896.

[8] Practice Bulletin No. 168: cervical cancer screening and prevention. Obstetrics and Gynecology. 2016; 128: e111–e130.

[9] Chesson HW, Dunne EF, Hariri S, Markowitz LE. The estimated lifetime probability of acquiring human papillomavirus in the United States. Sexually Transmitted Diseases. 2014; 41: 660–664.

[10] Nanda K, McCrory DC, Myers ER, Bastian LA, Hasselblad V, Hickey JD, *et al*. Accuracy of the papanicolaou test in screening for and follow-up of cervical cytologic abnormalities. Annals of Internal Medicine. 2000; 132: 810.

[11] Fang C, Wang SY, Liou YL, Chen MH, Ouyang W, Duan KM. The promising role of *PAX1* (aliases: HUP48, OFC2) gene methylation in cancer screening. Molecular Genetics & Genomic Medicine. 2019; 7: e506.

[12] Su PH, Lai HC, Huang RL, Chen LY, Wang YC, Wu TI, *et al*. Paired box-1 (*PAX1*) activates multiple phosphatases and inhibits kinase cascades in cervical cancer. Scientific Reports. 2019; 9: 9195.

[13] Nikolaidis C, Nena E, Panagopoulou M, Balgkouranidou I, Karaglani M, Chatzaki E, *et al*. *PAX1* methylation as an auxiliary biomarker for cervical cancer screening: a meta-analysis. Cancer Epidemiology. 2015; 39: 682–686.

[14] Zhang J, Cheng K, Wang Z. Prevalence and distribution of human papillomavirus genotypes in cervical intraepithelial neoplasia in China: a meta-analysis. Archives of Gynecology and Obstetrics. 2020; 302: 1329–1337.

[15] Guo C, Du H, Belinson JL, Wang C, Huang X, Qu X, *et al*. The prevalence and distribution of human papillomavirus among 10,867 Chinese Han women. Infectious Agents and Cancer. 2021; 16: 21.

[16] Fontham ETH, Wolf AMD, Church TR, Etzioni R, Flowers CR, Herzig A, *et al*. Cervical cancer screening for individuals at average risk: 2020 guideline update from the American Cancer Society. CA: A Cancer Journal for Clinicians. 2020; 70: 321–346.

[17] Wright TC, Massad LS, Dunton CJ, Spitzer M, Wilkinson EJ, Solomon D. 2006 consensus guidelines for the management of women with abnormal cervical screening tests. Journal of Lower Genital Tract Disease. 2007; 11: 201–222.

[18] Rogeri CD, Silveira HCS, Causin RL, Villa LL, Stein MD, de Carvalho AC, *et al*. Methylation of the hsa-miR-124, SOX1, TERT, and LMX1a genes as biomarkers for precursor lesions in cervical cancer. Gynecologic Oncology. 2018; 150: 545–551.

[19] Jiao X, Zhang S, Jiao J, Zhang T, Qu W, Muloye GM, *et al*. Promoter methylation of SEPT9 as a potential biomarker for early detection of cervical cancer and its overexpression predicts radioresistance. Clinical Epigenetics. 2019; 11: 120.

[20] Wu D, Zhang J, Fan P, Li H, Li D, Pan H, *et al*. Methylation in the promoter regions of WT1, NKX6-1 and DBC1 genes in cervical cancer tissues of Uygur women in Xinjiang. Genetics and Molecular Biology. 2018; 41: 9–17.

[21] Yin A, Zhang Q, Kong X, Jia L, Yang Z, Meng L, *et al*. JAM3 methylation status as a biomarker for diagnosis of preneoplastic and neoplastic lesions of the cervix. Oncotarget. 2015; 6: 44373–44387.

[22] Holubekova V, Mersakova S, Grendar M, Snahnicanova Z, Kudela E, Kalman M, *et al*. The role of CADM1 and MAL promoter methylation in inflammation and cervical intraepithelial neoplasia. Genetic Testing and Molecular Biomarkers. 2020; 24: 256–263.

[23] Cillo C, Cantile M, Faiella A, Boncinelli E. Homeobox genes in normal and malignant cells. Journal of Cellular Physiology. 2001; 188: 161–169.

[24] Robson EJD, He S, Eccles MR. A PANorama of *PAX* genes in cancer and development. Nature Reviews Cancer. 2006; 6: 52–62.

[25] Huang J, Liou YL, Kang YN, Tan ZR, Peng MJ, Zhou HH. Real-time colorimetric detection of DNA methylation of the *PAX1* gene in cervical scrapings for cervical cancer screening with thiol-labeled PCR primers and gold nanoparticles. International Journal of Nanomedicine. 2016; 11: 5335–5347.

[26] Kan YY, Liou YL, Wang HJ, Chen CY, Sung LC, Chang CF, *et al*. *PAX1* methylation as a potential biomarker for cervical cancer screening. International Journal of Gynecological Cancer. 2014; 24: 928–934.

[27] Lai H, Lin Y, Huang R, Chung M, Wang H, Liao Y, *et al*. Quantitative DNA methylation analysis detects cervical intraepithelial neoplasms type 3 and worse. Cancer. 2010; 116: 4266–4274.

[28] Liou Y, Zhang T, Yan T, Yeh C, Kang Y, Cao L, *et al*. Combined clinical and genetic testing algorithm for cervical cancer diagnosis. Clinical Epigenetics. 2016; 8: 66.

[29] Tian X, Chen D, Zhang R, Zhou J, Peng X, Yang X, *et al*. Quantitative survey of multiple CpGs from 5 genes identifies CpG methylation panel discriminating between high- and low-grade cervical intraepithelial neoplasia. Clinical Epigenetics. 2015; 7: 4.

[30] Tian Y, Yuan Wu N, Liou Y, Yeh C, Cao L, Kang Y, *et al*. Utility of gene methylation analysis, cytological examination, and HPV-16/18 genotyping in triage of high-risk human papilloma virus-positive women. Oncotarget. 2017; 8: 62274–62285.

[31] Zhang L, Yu J, Huang W, Zhang H, Xu J, Cai H. A sensitive and simplified classifier of cervical lesions based on a methylation-specific PCR assay: a chinese cohort study. Cancer Management and Research. 2020; 12: 2567–2576.

[32] Lai HC, Ou YC, Chen TC, Huang HJ, Cheng YM, Chen CH, *et al*. *PAX1*/SOX1 DNA methylation and cervical neoplasia detection: a Taiwanese gynecologic oncology group (TGOG) study. Cancer Medicine. 2014; 3: 1062–1074.

[33] Forero DA, Patrinos GP, Science D. Genome plasticity in health and disease. 1st edn. Academic Press: London. 2020.

[34] Singhal N, Kumar M, Kanaujia PK, Virdi JS. MALDI-TOF mass spectrometry: an emerging technology for microbial identification and diagnosis. Frontiers in Microbiology. 2015; 6: 791.
[35] Zhang L, Hu D, Wang S, Zhang Y, Pang L, Tao L, *et al*. Association between dense *PAX1* promoter methylation and HPV16 infection in cervical squamous epithelial neoplasms of Xin Jiang Uyghur and Han women. Gene. 2020; 723: 144142.

How to cite this article: Yu Chen, Yuanyuan Zhang, Xiaochun Chen, Hong Chu, Songqing Guan, Qiuni Liang, *et al*. CpG methylation of the *PAX1* promoter in exfoliated cervical cells: a potential cervical cancer screening model. European Journal of Gynaecological Oncology. 2024; 45(3): 48-57. doi: 10.22514/ejgo.2024.047.