Establishment and analysis of the prediction model for cervical squamous cell carcinoma


Department of Obstetrics and Gynaecology, Xiangyang No.1 People's Hospital, Hubei University of Medicine, Xiangyang, China

Yuhan Zhou and Wufeng Fan contributed equally to this work

Abstract. – OBJECTIVE: This study aimed to construct a prediction model for cervical squamous cell carcinoma and evaluate its accuracy in diagnosing cervical squamous cell carcinoma.

PATIENTS AND METHODS: Fifty patients with initially histopathologically confirmed cervical squamous cell carcinoma and 150 patients with initially histopathologically confirmed cervical intraepithelial neoplasia (CIN) were enrolled. The high-risk human papillomavirus (HR-HPV) infection, human telomerase mRNA component (hTERC) gene and cell-myc (c-myc) gene amplification, and minichromosome maintenance protein 5 (MCM5) protein expression were detected. The indicators related to cervical cancer were screened. The regression model was established to predict cervical squamous cell carcinoma with backward logistic stepwise regression method, and the accuracy of the model was evaluated.

RESULTS: Histograms for HR-HPV infection and viral load, hTERC and c-myc gene amplification, and MCM5 protein expression were constructed. There was a linear relationship between hTERC (X₁), HR-HPV viral load (X₂), MCM5 (X₅) and the regression equation. Also, hTERC (X₁), HR-HPV viral load (X₂) and MCM5 (X₅) were correlated with cervical squamous cell carcinoma. The regression model Logit (p) = -66.283 + 0.042 X₁ + 0.061 X₂ + 0.052 X₅ was established. The model-fitting effect and prediction accuracy were evaluated, HL test p = 1 (p > 0.05). The model fitting effect was good, Cox-Snell R² was 0.643 and Nagelkerke R² was 0.958. The high accuracy of the model was 98.5%.

CONCLUSIONS: The fitting-effect of the regression model established by hTERC gene expression, HR-HPV viral load and MCM5 protein was good. The prediction accuracy of the model for cervical squamous cell carcinoma was high. The combined test of hTERC gene amplification, HR-HPV viral load and MCM5 protein could be used to predict and evaluate cervical squamous cell carcinoma.


Introduction

Cervical cancer is the second most common cancer that occurs in women worldwide. The incidence of cervical cancer also ranks second among malignant tumors in women in developing countries. There were approximately 75,000 newly diagnosed cases of cervical cancer in China each year, and the disease seriously threatened the health of women. Cervical intraepithelial neoplasia (CIN) is a group of precancerous lesions that is closely correlated with cervical cancer and reflects the continuous progress in the development and progression of cervical cancer. According to the International Society of Cancer, the high-risk human papillomavirus (HR-HPV) was a major cause of cervical cancer in 1995. The detection of HR-HPV had been a major item for the screening, diagnosing, and follow-up of cervical disease. However, most of the women with HR-HPV infection do not develop cervical cancer. Researchers have recognized that HR-HPV infection is not the only causative agent of the cells’ malignant transformation. The HR-HPV infection might be the primary causal event in the development of cervical cancer, but there were other pathogenic factors involved. In 2006, Hopman et al found that integration of the HR-HPV oncogene and amplification of the human telomerase mRNA component gene (hTERC) might be a major factor related to the progression from CIN to cervical cancer. Other scholars believed that minichromosome maintenance protein
5 (MCM5) could be used as a biological marker for malignant tumors and atypical hyperplasia. Some studies identified amplification of cell-myc (c-myc) gene expression in malignant tumors. Sagawa et al. hypothesized that c-myc gene amplification might be one of the main indicators for the occurrence and development of cervical cancer. The existing screening tools (including combined screening) could identify cervical lesions, but could not accurately determine and predict the progression and outcome of early lesion, and result in blind treatment and follow-up or even over treatment. The period of progression from CIN to invasive cancer was important for early diagnosis and prevention of development of malignancy. Therefore, clinical prediction methods were urgently needed which could not only detect high-risk patients with low-grade lesions at an early stage, but also more accurately predict progression and prognosis to achieve patients' diversion management and early intervention and treatment. The objective of the study was to screen indicators related to cervical cancer and establish a prediction model by detecting HR-HPV infection and viral load in patients with cervical squamous cell carcinoma and different grades of CIN. Also, hTERC, c-myc gene amplification, and MCM5 protein expression were measured to provide a reference for the evaluation and prognosis of these patients.

Patients and Methods

Patients

Patients with cervical lesions referred to our hospital from January 2012 to January 2015 were enrolled. The histopathological diagnosis of cervical lesions was made according to previously described standards. There were 357 cases with histopathologically confirmed CIN and cervical cancer through cervical biopsy under a colposcopy. Finally, a total of 200 cases were selected according to the inclusion criteria in this study, including 50 cases with cervical squamous cell carcinoma and 150 cases with CIN (CIN I: 58 cases, CIN II: 50 cases, and CIN III: 42 cases). The age of patients ranged from 21-64 years old, and the average age was 42.7 years. Pathologists examined the exfoliated cervical cells and tissue specimens of all enrolled subjects and measured HR-HPV infection status and viral load, hTERC and c-myc gene amplification, and MCM5 protein expression. This study was approved by the Research Medical Ethics Committee of our Hospital, and all patients signed the informed consent.

Inclusion Criteria

1. All specimens were initial specimens. The specimens of all patients were collected before the initial treatment. Two experienced pathologists confirmed the unified diagnosis of CIN and cervical squamous cell carcinoma.
2. All patients did not receive any physiotherapy or chemotherapy, and surgery of cervix before the above-mentioned examinations. All patients had no history of tumors in other systems, no history of hysterectomy, and no severe diseases in other systems.

Exclusion Criteria

1. The histopathological diagnosis of cervical lesions by biopsy was the other except CIN and cervical squamous cell carcinomas, such as chronic cervicitis, cervical adenocarcinoma and adenosquamous carcinoma.
2. CIN and cervical squamous cell carcinoma had been diagnosed and treated.
3. The patient had a history of tumors in other systems, hysterectomy, or severe diseases in other organ systems.

Testing Methods

HR-HPV test: the hybrid-capture II (HC-II) assay was used to quantitatively detect HR-HPV (including 13 types) in exfoliated cervical cells (liquid-based cytology technology-ThinPrep Cytology Test, TCT). HPV-DNA level was expressed as relative light units/cutoff (RLU/CO). RLU/CO >1 indicated HPV-DNA content >1 pg/ml, and the result was positive. Conversely, it was negative. The viral load was originally recorded as RLU/CO and the discrepancy was larger, so the logarithmic conversion statistical analysis was carried out.

Test of hTERC and c-myc gene amplification: fluorescence in situ hybridization (FISH) was used to measure hTERC and c-myc gene amplification in exfoliated cervical cells (the remaining liquid in the TCT bottle). Preparation of cells on slides, pretreatment, denaturation, and hybridization was carried out according to the instructions of the FISH kit. The hybridization probe was supplied by Beijing Jinpujia Medical Technology (Co., Ltd, Beijing, China). The hTERC locus-specific probe (GLP hTERC) and chromosome 3 centromere probe (CSP 3) were labeled as red
and green, respectively. CSP 3 was used as the control probe. In the same period, 20 cases of cervical cytology results from healthy women were randomly selected in the routine physical examination. At least 100 cells were analyzed per case. More than 2 cells were labeled as red, the cells with ratio of red/green were 3:2, 4:2, 3:3, 4:4 and high hTERC gene copy number were considered abnormal. The threshold value was calculated as follows: threshold value = (the mean of the ratio + 3) × standard deviation × 100%. The patients with threshold value higher than these values were positive for abnormal hTERC gene amplification. Otherwise, they were negative. The threshold was 6.32% based on the aforementioned equation in this study. The c-myc gene amplification was detected and calculated according to the same method; the threshold was 3.18%.

MCM5 protein expression test: immunohistochemical streptavidin-peroxidase (SP) method was used. Sections were deparaffinized and incubated at room temperature for 10 min. Ethylene Diamine Tetraacetic Acid (EDTA) antigen retrieval buffer was added and sections were treated under high temperature and pressure. Sections were blocked and treated with primary and second antibodies. 3,3’-Diaminobenzidine was used for color development, and hematoxylin was used for counter-staining. Stained sections were evaluated microscopically based on the proportion of positive cells in the sections and the staining depth. If the results were inconsistent, they were re-evaluated. PBS was used as the negative control instead of primary antibody. A nucleus that stained brown was considered a positive result. Cells were then grouped according to staining intensity and the number of positive cells as follows: 0 point for staining similar to the background; 1 point for light staining and those slightly higher than the background; 2 points for moderate staining and those significantly higher than the background; 3 points for strong staining and those in dark brown. 0 point for the number of positive cells < 10%; 1 point for 10-50%; 2 points for 51-75%; and 3 points for > 75%.

Statistical Analysis
All the data analyses were performed using SPSS version 20 (SPSS Inc., Chicago, IL, USA). Backward logistic stepwise regression analysis was adopted, the test standards were α = 0.05. The predicting indicators were screened and the regression model was established. The Hosmer-Lemeshow (HL) test was used to evaluate the fitting-effect of the model. The generalized determination coefficient (R²) and the prediction probability sheet were used to evaluate the prediction accuracy of the model.

Results
HR-HPV infection status and viral load, hTERC gene amplification, c-myc gene amplification, and MCM5 protein expression tended to increase with the severity of cervical lesions (Figure 1). The backward logistic regression analysis was carried out with the grade of cervical lesions as the dependent variable and hTERC, HR-HPV load, HR-HPV infection status, c-myc, and MCM5 as the independent variables: X1, X2, X3, X4 and X5 respectively. The results showed that the regression coefficients including hTERC (X1), HR-HPV load (X2), and MCM5 (X5) were 0.042, 0.061 and 0.052; the p-values were 0.024, 0.005, and 0.005, respectively (p < 0.05). There was a linear relationship between hTERC (X1), HR-HPV viral load (X2), MCM5 (X5) and the regression equation. It indicated that hTERC (X1), HR-HPV viral load (X2), and MCM5 (X5) were associated with cervical squamous cell carcinoma. For HR-HPV infection status (X3) and c-myc (X4), the p-values were 0.856 and 0.682 (p > 0.05), and they were ruled out by the regression equation. It indicated that there was no significant correlation between HR-HPV infection status or c-myc and the regression equation. Then, hTERC (X1), HR-HPV load (X2), and MCM5 (X5) were used to establish the regression model Logit (p) = -66.283 + 0.042 X1 + 0.061 X2 + 0.052 X5 (Tables I, II).

The fitting-effect and prediction accuracy of the model were evaluated. HL test was carried out, and the result showed that p = 1 (p > 0.05). The fitting-effect of the model was good, the determination coefficients of the model Cox-Snell R² was 0.643 and Nagelkerke R² was 0.958. The surprisingly high prediction accuracy was 98.5% (Tables III-V).

Discussion
At present, HPV, TCT and other diagnostic methods combined with biomarkers were applied in the early diagnosis of CIN and cervical cancer. Domestic and international studies also focused on expanding the scope of screening, improving
Establishment and analysis of the prediction model for cervical squamous cell carcinoma

screening quality, sensitivity and specificity, simplifying the screening process and scheme, and improving the diagnostic rate of cervical lesions. Epidemiological studies had confirmed that HPV was the direct cause of cervical cancer. Most scholars\(^9,10\) believed that HPV load was positively correlated with the incidence of CIN and invasive cervical cancer. According to Woodman et al\(^11\), the virus in the free state was not sufficient to cause cancer, and the integration of HR-HPV DNA in host cells was the leading cause of cervical cancer. HR-HPV might integrate into the host cell DNA, which could result in activating oncogenes, inactivating tumor suppressor genes\(^12,13\) inducting telomerase expression, and causing abnormal cell proliferation and apoptosis, leading

![Figure 2. High-risk type HPV infection, hTERC/C - MYC amplification and MCM 5 expression results.](image)

<table>
<thead>
<tr>
<th>Score</th>
<th>df</th>
<th>(p)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(X_1)</td>
<td>0.033</td>
<td>1</td>
</tr>
<tr>
<td>(X_2)</td>
<td>0.167</td>
<td>1</td>
</tr>
<tr>
<td>The total statistical magnitude</td>
<td>0.326</td>
<td>2</td>
</tr>
</tbody>
</table>

Table I. Logistic regression analysis of hTERC, HR-HPV, c-myc, and MCM 5.

<table>
<thead>
<tr>
<th>Step 4</th>
<th>B</th>
<th>S.E.</th>
<th>Wals</th>
<th>df</th>
<th>(p)</th>
<th>(Exp(B))</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Upper limit</td>
</tr>
<tr>
<td>X_1</td>
<td>0.042</td>
<td>0.019</td>
<td>5.073</td>
<td>1</td>
<td>0.024</td>
<td>1.043</td>
</tr>
<tr>
<td>X_2</td>
<td>0.061</td>
<td>0.021</td>
<td>8.066</td>
<td>1</td>
<td>0.005</td>
<td>1.063</td>
</tr>
<tr>
<td>X_3</td>
<td>0.052</td>
<td>0.019</td>
<td>7.740</td>
<td>1</td>
<td>0.005</td>
<td>1.053</td>
</tr>
<tr>
<td>Constant</td>
<td>-66.283</td>
<td>19.990</td>
<td>10.995</td>
<td>1</td>
<td>0.001</td>
<td>0.000</td>
</tr>
</tbody>
</table>

Table II. Variables not in the equation.
Y.-H. Zhou, W.-F. Fan, J. Deng, H.-L. Xi

The incidence of cervical cancer finally. Most HPV tests were performed by measuring the sum of free and integrated HPV. Although HPV test was an important diagnostic method, it was characterized by low specificity. The hTERC gene was an important gene involved in cervical tumorigenesis. Some studies had shown that the occurrence of cervical cancer was closely related to the gain of copy number on the long arm of human chromosome 3. The chromosome change could prevent apoptosis and lead to cell immortalization, and thus cause the occurrence of tumors. Heselmeyer et al. found that the sensitivity and specificity of hTERC gene amplification for diagnosing CIN were over 90%. Therefore, they believed that hTERC gene amplification could be used as an independent indicator to predict CIN. Abnormal expression of the c-myc gene closely related to the occurrence and development of a variety of malignant tumors, which could be caused by gene amplification, chromosomal rearrangement, and overexpression. The result was that cells proliferated indefinitely and become immortalized, thereby leading to tumorigenesis. Ding et al. found that the c-myc gene did not express in normal cervical epithelium, but higher expression of the c-myc gene began to appear in severe hyperplastic epithelium, and the expression of c-myc further increased in cervical cancer. He et al. showed that c-myc gene expression level significantly correlated with HPV16/18 expression, suggesting that the carcinogenic effect of HR-HPV infection might be achieved through activation of the c-myc gene. The c-myc gene might be the boot gene of HPV16/18 infection. MCM5 protein expression could accurately reflect cell proliferation. It did not express in quiescent, senescent or differentiated cells, suggesting that MCM5 protein could be used as a marker of cell proliferation. Freeman et al. found that the expression of MCM5 was limited to the base portion and corresponding sites of proliferation in normal cervical tissue, while MCM5 expressed in the entire epithelial layer and inversely proportioned to the degree of differentiation in atypical hyperplasia and tumor tissues. MCM5 protein could be used as a biological indicator for screening cervical lesions. In this study, HR-HPV infection, viral load, hTERC, c-myc gene amplification, and MCM5 protein expression were detected in different grades of CIN and cervical squamous cell carcinoma tissues. The result was consistent with other reports. All indicators showed increased amplification and expression with the severity of cervical lesions, suggesting that HR-HPV, the hTERC gene, the c-myc gene, and MCM5 played an important adjunct diagnostic role in screening cervical lesions.

In recent years, some scholars had proposed to increase the early diagnosis of cervical cancer through combined testing methods (combination of TCT test with HPV, hTERC, HER-2, and c-myc gene). Zhao et al. proposed that the normal amplification of the hTERC and c-myc genes correlated with the severity of cervical lesions, which could more accurately reflect the grade of cervical lesions. According to Sun et al., the sensitivity of combination of TCT test with

<table>
<thead>
<tr>
<th>Step</th>
<th>$\chi^2$</th>
<th>df</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.417</td>
<td>8</td>
<td>1.000</td>
</tr>
<tr>
<td>4</td>
<td>0.417</td>
<td>8</td>
<td>1.000</td>
</tr>
</tbody>
</table>

Table IV. Generalized correlation factors of the model.

<table>
<thead>
<tr>
<th>Step</th>
<th>Cox-Snell $R^2$</th>
<th>Nagelkerke $R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.644</td>
<td>0.959</td>
</tr>
<tr>
<td>4</td>
<td>0.643</td>
<td>0.958</td>
</tr>
</tbody>
</table>

Table V. The model prediction accuracy.

<table>
<thead>
<tr>
<th>Observation</th>
<th>Degree of cervical lesions</th>
<th>Correction percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Step 4</td>
<td>CIN</td>
<td>149</td>
</tr>
<tr>
<td></td>
<td>Cancer</td>
<td>2</td>
</tr>
<tr>
<td>Total percentage</td>
<td></td>
<td>149</td>
</tr>
</tbody>
</table>
Establishment and analysis of the prediction model for cervical squamous cell carcinoma

HPV, hTERC, HER-2, and c-myc gene test for screening high-grade cervical lesions was higher, and the misdiagnostic rate was lower compared with TCT screening alone. Especially TCT and hTERC combined testing scheme was the best, and it could be used as an adjunct method for early screening cervical cancer. However, there were few studies about methods of scientific prediction for cervical lesions at present. This article focused on studying the indicators for screening cervical cancer, establishing a model to predict cervical squamous cell carcinoma, and evaluating the effectiveness and accuracy of the model. We found that hTERC, HR-HPV load, and MCM5 protein expression correlated with progression of CIN and the occurrence of cervical squamous cell carcinoma, and they could be used to establish a model to predict the occurrence of CIN and cervical squamous cell carcinoma. HR-HPV infection status and c-myc gene amplification did not significantly correlate with cervical cancer. This might be due to most HPV infections were transient, and c-myc might be a target gene of HPV. Therefore, neither of the two could be used as an indicator to predict cervical squamous cell carcinoma. In this study, we also found that the fitting-effect of the model was good and the prediction accuracy was high. We proposed to test patients with CIN at different stages through combining hTERC gene amplification, HR-HPV load and MCM5 protein expression. We also suggested predicting the occurrence and outcome of cervical squamous cell carcinoma through establishing a model. According to the results of prediction, the management of high-risk patients should be strengthened, the follow-up should be strictly carried out, and the intervention and treatment actively should be adopted. It will further improve the early diagnosis of cervical squamous cell carcinoma, reduce blind treatment and improve prognosis, to screen and early detect cervical lesions combined cytology with histopathology.

Conclusions

With the understanding of the etiology and pathogenesis of cervical squamous cell carcinoma gradually deepening, it could provide a guidance for the early diagnosis of CIN and cervical squamous cell carcinoma, observation of the disease outcome and correct prediction of the prognosis, to detect the important marker in the development from normal cervical to CIN, then to cervical cancer. However, because the occurrence and development of cervical cancer were not the result of a single factor, each biomarker might have limitations. Therefore, it could provide suggestion and theoretical foundation for improving the diagnosis rate and predicting the occurrence and development of cervical squamous cell carcinoma, to combine test and establish an evaluation model for prediction. It has high clinical and social value to strengthen diversion management of high-risk patients, properly apply intervention and treatment and correctly evaluate prognosis in clinical practice.

Conflict of Interest
The Authors declare that they have no conflict of interests.

References


