

## RESEARCH ARTICLE

# Development of a Risk Index for Prediction of Abnormal Pap Test Results in Serbia

Dejana Vukovic<sup>1\*</sup>, Ljiljana Antic<sup>2</sup>, Mladenko Vasiljevic<sup>3</sup>, Dragan Antic<sup>2</sup>, Bojana Matejic<sup>1</sup>

### Abstract

**Background:** Serbia is one of the countries with highest incidence and mortality rates for cervical cancer in Central and South Eastern Europe. Introducing a risk index could provide a powerful means for targeting groups at high likelihood of having an abnormal cervical smear and increase efficiency of screening. The aim of the present study was to create and assess validity of a index for prediction of an abnormal Pap test result. **Materials and Methods:** The study population was drawn from patients attending Departments for Women's Health in two primary health care centers in Serbia. Out of 525 respondents 350 were randomly selected and data obtained from them were used as the index creation dataset. Data obtained from the remaining 175 were used as an index validation data set. **Results:** Age at first intercourse under 18, more than 4 sexual partners, history of STD and multiparity were attributed statistical weights 16, 15, 14 and 13, respectively. The distribution of index scores in index-creation data set showed that most respondents had a score 0 (54.9%). In the index-creation dataset mean index score was 10.3 (SD=13.8), and in the validation dataset the mean was 9.1 (SD=13.2). **Conclusions:** The advantage of such scoring system is that it is simple, consisting of only four elements, so it could be applied to identify women with high risk for cervical cancer that would be referred for further examination.

**Keywords:** Cervical cancer - risk - index score - pap test - Serbia

*Asian Pac J Cancer Prev*, 16 (8), 3527-3531

### Introduction

Due to high incidence and mortality rates, cervical cancer represents significant public health problem around the world. Serbia is one of the countries with highest incidence and mortality rates in Central and South Eastern Europe (Vrdoljak, et al., 2011). Despite some efforts and pilot projects, organized screening for cervical cancer failed to cover significant proportion of women (Kesic et al., 2012; Schoueri-Mychasiw and McDonald, 2013). Evidence from research showed that some factors significantly increase risk for cervical cancer, and these findings could help health professionals to focus preventive efforts. Sexual behaviors such as number of lifetime sexual partners, becoming sexually active at early age, history of sexually transmitted diseases (STD) were consistently found to be associated with increased risk for cervical carcinoma (Stone et al., 1995; International collaboration of Epidemiological Studies of Cervical Cancer, 2006, 2007, 2009; Reich, 2005). Multiple sexual partners of woman's partner, along with partner's history of sexually transmitted disease and history of intravenous drug use, were also identified as important risk factors (Maree et al., 2011; Remschmidt

et al., 2013). Along with sexual behavior, some other characteristics proved to be significant for the occurrence of cervical cancer, such as multiparity, smoking and lower educational status (Corral et al., 1996; Sherris et al., 2009; Raychaudhuri and Mandal, 2012). Numerous studies have established positive association between duration of oral contraceptives use and increased risk for cervical cancer (Kjellberg et al., 2000; Castellsague and Munoz, 2003; Smith et al., 2003).

Despite wide range of studies assessing individual risk factors associated with cervical carcinoma, only few studies attempting to calculate composite risk index could be identified. Wilkinson et al. (1994) constructed a scoring system by calculating number of selected risk factors. Introducing such index could provide powerful means for targeting groups at high risk of having abnormal cervical smear and increase efficiency of screening. Patil et al created a risk scoring system for prediction of cervical cancer using only demographic variables of women (Patil et al., 2006). Within the CARE study, Ritter et al developed and validated risky sexual behavior risk index for cervical cancer, based on sexual behavior of the respondents, and using self-reported Pap test results as outcome measure (Reiter et al., 2009).

<sup>1</sup>University of Belgrade, Faculty of Medicine, Institute for Social Medicine, <sup>2</sup>College of Health Studies in Cuprija, <sup>3</sup>University of Belgrade, Faculty of Medicine, University Clinic of Gynaecology and Obstetrics Narodni Front, Serbia \*For correspondence: [dvukovic@med.bg.ac.rs](mailto:dvukovic@med.bg.ac.rs)

Our aim was to create and assess validity of index for prediction of abnormal Pap test using both demographic and sexual behavior characteristics of women.

## Materials and Methods

The study population was drawn from the patients attending departments for women's health in two primary health care centers in Serbia. All women who visited department for women health during three month period (1.1.2013-31.3.2013) were selected and their medical records were reviewed for eligibility. Inclusion criteria were: attending primary care gynecologist and having medical records for at least three years, age 18 years and over, not pregnant at the moment of study, not having had hysterectomy, not being diagnosed with invasive cervical cancer. Total number of enrolled patients was 525. All women were informed about the study and signed consent to participate.

*Ethical committee of the Faculty of Medicine, University of Belgrade has approved the study (Decision number 29/1-15).*

Data were obtained using specially designed and pretested questionnaire, as well as patient's medical records. Data regarding socio-demographic characteristics of patients and their behavior was obtained using questionnaire. Data regarding Pap test results were taken from medical records and filled in by the gynecologist. As abnormal Pap test results we considered grade III, IV and V.

As risk factors in present study we analyzed factors which, in previous research, showed significant association with cervical cancer, as well as socio-demographic characteristics. We analyzed: education, marital status (widow, separated, single vs. married), self-assessed socio-economic status (very bad/bad vs. good/very good), low education (elementary or less vs. higher education) multiparity (more than three), ever smoker (smoked at least 100 cigarettes in lifetime), number of lifetime sexual partners (>4 vs.  $\geq 4$ ), sexual partner with history of sexually transmitted disease (STD), personal history of STD, age at first sexual intercourse (<18 vs.  $\geq 18$ ), age at first delivery (<18 vs.  $\geq 18$ ), use of condom (inconsistent vs. consistent), use of oral contraceptives (>5 years), sexual partner with history of injecting drug use.

### Statistical analysis

Out of 525 respondents 350 were randomly selected and data obtained from them were used as index creation dataset. Data obtained from the remaining 175 were used as index validation data set.

### Risk factor index creating

In order to analyze significance of socio-demographic and risk factors, univariate logistic regression was performed with Pap test result as dependent variable. Variables with statistically significant association with abnormal Pap test (grade III, IV and V) were then used to obtain logistic regression model containing eight risk factors as independent, and abnormal Pap test as

dependent variable. The model was constructed to determine association of each risk factor with abnormal Pap test, while controlling for all other factors. Risk factors found to be significantly associated were included in the final model of logistic regression. In order to calculate the risk index based on the final model of logistic regression significant factors in final model were then allotted weights. The weight for factor was calculated based on regression coefficient in the final model. Regression coefficients representing the difference in the log odds of the occurrence of abnormal Pap test for subjects with risk factor vs those without, after controlling for the confounding effects of the covariates in the model.

In order to make risk score more simple to calculate in every day practice, we used linear transformation for weights using the following formula:

$Weight = bx10$ . (Patil et al., 2006). Weights were assigned to risk factors, and calculated for each respondent. For each risk factor respondent was assigned appropriate weight, if she had reported having risk, or 0 if not. Total index score was calculated by adding weighted risk factors for each respondent. Sensitivity and specificity of the index were calculated and Receiver Operating Curve was constructed and used to identify cut-off point.

### Index validation

The validation dataset was used to validate risk factor index. Univariate logistic regression was performed using Pap test as dependent and risk index as independent variable. Hosmer-Lemeshow test was used to assess goodness-of-fit.

## Results

Demographic characteristics of respondents are presented in Table 1. Mean age of respondents was 40.5 (SD 10.8), 30% perceived their socioeconomic status as bad or very bad. Women included in the analyses were most likely to have had secondary level of education (57.1%). One third of respondents were single/divorced/widowed.

Distribution of risk factors is presented in Table 2. Most prevalent risk factors were inconsistent condom use

**Table 1. Socio-Demographic Characteristics of the Respondents**

| Characteristic                      | Index Creation Dataset | Validation Dataset |
|-------------------------------------|------------------------|--------------------|
| Age (Mean, Sd)                      | 40.0 (10.8)            | 41.6 (10.8)        |
| Self-Assessed Socio-Economic Status |                        |                    |
| Very Good                           | 53 (15.3%)             | 26 (14.9%)         |
| Good                                | 202 (58.4%)            | 106 (60.9%)        |
| Bad                                 | 77 (22.3%)             | 33 (19.0%)         |
| Very Bad                            | 14 (4.0%)              | 9 (5.2%)           |
| Education                           |                        |                    |
| Elementary                          | 41 (11.7%)             | 17 (9.7%)          |
| High School                         | 203 (58.0%)            | 97 (55.4%)         |
| College/University                  | 106 (30.3%)            | 61 (34.9%)         |
| Marital Status                      |                        |                    |
| Single/Divorced/Widowed             | 114 (32.6%)            | 53 (30.3%)         |
| Married/Common Law                  | 236 (67.4%)            | 122 (69.7%)        |

(82.0%), smoking (53.0%) and having had first intercourse under the age of 18 years (22.5%).

Univariate logistic regression models were created for all socio-demographic and risk factors as independent variables, and abnormal Pap test as dependent variable. All factors that were significantly associated with abnormal Pap test were included in multiple logistic regression model, as independent variables and abnormal Pap test as dependent variable. Multiple logistic regression model showed that multiparity, more than 4 lifetime sexual partners, personal history of STD and age at first intercourse were significantly associated with abnormal Pap test, even when controlling for other factors (Table 3). Final model was created using 4 before mentioned risk factors as independent variables, in order to obtain b coefficients. Final model was used to calculate weights for risk factors. Weights were derived from b coefficients (Table 3). Age at first intercourse under 18, more than

4 sexual partners, history of STD and multiparity were attributed statistical weights 16, 15, 14 and 13, respectively. The distribution of index scores in index-creation data set showed that most of respondents had score 0 (54.9%).

*In index-creation dataset mean index score was 10.3 (SD=13.8), and in validation dataset mean was 9.1 (SD=13.2)*

The ROC curve was constructed from index-creation dataset, and area under the curve was 0.808. Sensitivity and specificity for selected cutoff points are presented in Table 4. Maximum Cohen's Kappa score was at the value of index of  $\geq 28.5$ . This was validated in index-validation dataset, where overall predictive accuracy of the risk scoring system calculated with Wilcoxon statistics was 0.926 (CI 0.823-0.957).

The Hosmer-Lemeshow goodness-of-fit test didn't indicate a lack of fit in index-validation dataset (n=175, p=0.130), successfully validating scoring index created.

**Table 2. Frequency of Risky Sexual Behavior among Respondents in Index-Creation and Validation Dataset**

| Risk behaviour                      | Index creation dataset | Validation dataset |
|-------------------------------------|------------------------|--------------------|
| Multiparity                         | 44 (12.6%)             | 15 (8.7%)          |
| Ever smoker                         | 190 (54.3%)            | 88 (50.3%)         |
| Number of sexual partners >4        | 68 (19.4%)             | 30 (17.1%)         |
| Sexual partner with STD             | 11 (3.1%)              | 7 (4.0%)           |
| Sexual partner IV drug user         | 3 (0.86%)              | 0                  |
| History of STD                      | 42 (12.0)              | 23 (13.1)          |
| Age at first intercourse<18         | 89 (25.4%)             | 39 (22.3%)         |
| Age at first delivery<18            | 29 (8.3%)              | 18 (10.3%)         |
| Inconsistent condom use             | 284 (81.14%)           | 147 (84.0%)        |
| Use of oral contraceptives >5 years | 10 (2.86%)             | 1 (0.57%)          |

**Table 3. Multiple Logistic Regression Coefficients and Weights for the Risky Sexual Behavior in the Index Creation Dataset (n=350)**

| Risk behavior                            | b      | sig   | weight |
|--|--------|-------|--------|
| Marital status (not in the relationship) | 0.253  | 0.511 |        |
| Multiparity                              | 1.413  | 0.002 |        |
| Ever smoker                              | 0.043  | 0.91  |        |
| Number of sexual partners >4             | 1.445  | 0     |        |
| Sexual partner with STD                  | 0.439  | 0.609 |        |
| History of STD                           | 1.263  | 0.009 |        |
| Age at first intercourse<18              | 1.675  | 0     |        |
| Age at first delivery<18                 | -0.279 | 0.757 |        |
| Final model                              |        |       |        |
| Multiparity                              | 1.297  | 0.003 | 13     |
| Number of sexual partners >4             | 1.522  | 0     | 15     |
| History of STD                           | 1.398  | 0.001 | 14     |
| Age at first intercourse<18              | 1.637  | 0     | 16     |

**Table 4. Sensitivity and Specificity of Index in Index-Creation Sample**

| Value      | Sensitivity | Specificity | Cohen's Kappa (SE) |
|------------|-------------|-------------|--------------------|
| $\geq 6.5$ | 0.794       | 0.624       | 0.2627 (0.0436)    |
| 13.5       | 0.73        | 0.697       | 0.2979 (0.0471)    |
| 14.5       | 0.714       | 0.739       | 0.3348 (0.0488)    |
| 15.5       | 0.683       | 0.812       | 0.4084 (0.0514)    |
| 22         | 0.603       | 0.927       | 0.5435 (0.0534)    |
| 28.5       | 0.603       | 0.934       | 0.6687 (0.0754)    |
| 29.5       | 0.54        | 0.965       | 0.5722 (0.0523)    |
| 30.5       | 0.508       | 0.976       | 0.568 (0.051)      |

## Discussion

The overall objective of this study was to create and validate an index to identify women with high risk of cervical cancer based on known risk factors for cervical cancer. As the study was cross sectional, we used abnormal Pap test as proxy for risk of cervical cancer. We included in the analysis known socio-demographic and behavioral factors (Raychaudhuri and Mandal, 2012; Antic et al., 2014). Regarding demographic characteristics of the respondents being single (never married, divorced or widowed) as well as multiparity were significantly associated with abnormal Pap test in univariate analysis. However, after controlling for other behavioral factors marital status was no longer significantly associated with abnormal Pap test. The explanation is probably that marital status has high correlation with some behavioral factors. Multiparity remained significant and was included in the final model. Significance of multiparity was recognized by other researchers as well (Patil et al, 2006; Uysal and Birsal, 2009; Paçarada et al, 2010). Looking at behavioral risk factors smoking (Sherris et al, 2009; International collaboration of Epidemiological Studies of Cervical Cancer, 2006; Collins et al, 2010), life-time number of sexual partners (Maree et al, 2011; Reiter et al, 2009; Petry et al, 2013), personal history of STD (International collaboration of Epidemiological Studies of Cervical Cancer, 2009; Al-Daraji and Smith, 2009), as well as sexual partner with history of STD (Maree et al, 2011; Reiter et al, 2009), age at first intercourse (Stone et al, 1995; International collaboration of Epidemiological Studies of Cervical Cancer, 2009) and age at first delivery under 18 (International collaboration of Epidemiological Studies of Cervical Cancer, 2006) were significantly associated with abnormal Pap test, which was consistent with findings in previous studies. However, having sexual partner with history of intravenous drug use was indicated in previous studies as significantly associated with cervical cancer (Reiter et al, 2009; Cole et al, 2007; Conti et al, 1993), in our study only three respondents indicated having had sexual partner with history of drug use. In

the final model four factors were extracted representing the main contributors associated with the abnormal Pap test: multiparity, more than 4 lifetime sexual partners, personal history of STD and having had first sexual intercourse before age of 18. Taking into account these four factors it was possible to predict abnormal Pap test in the validation dataset with the overall accuracy of 0.926. The value of index of 28.5 or above was the point from which there was significantly higher risk of abnormal Pap test. Combination of having had first intercourse before age of 18 with any other risk factor, or having more than four lifetime sexual partners combined with a personal history of STD renders a woman at higher risk of having abnormal Pap test.

Authors stress the importance of sexual partner's behavior as risk factor for cervical cancer in women (18). We did not include them in the index as those factors were not significantly associated with abnormal Pap test. The explanation is probably related to very small number of respondents who indicated that they have had sexual partner with history of STD or sexual partner with history of intravenous drug use. Such a small percentage is probably underestimated due to either reluctance of respondents to reveal it to investigators or the fact that they were not aware of their partner's risk behavior.

In our study we created risk index using combined demographic and sexual behavior factors. Even though some studies used self-reported Pap test results (Reiter et al., 2009), we decided to use Pap test results from the respondents' medical records. It has been demonstrated that self-report is much less valid than medical records (Schroder et al., 2003; Nelson et al., 2009; Lofters et al., 2013).

The advantage of such scoring system is that it is simple, consisting of only four elements, so could be applied to identify women with high risk for cervical cancer that would be referred for further examination. It could be especially important in developing countries where significant proportion of women is not covered with screening for cervical cancer, particularly in rural areas (Smith et al., 2003; Reiter et al., 2009). Using the index to assess potential risk for cervical cancer could serve to identify individuals as well as population groups with increased risk, and focus educational efforts in order to stimulate those with increased risk to modify behavior and undertake Pap test regularly (Reiter et al., 2009).

There were few limitations of this study. The study population was limited to women who were visiting gynecologist. Data regarding sexual behavior were gathered using questionnaire, so the validity of answers could not be confirmed. Some respondents could have underreported risk sexual behavior, however the index was successfully validated in validation dataset indicating that index could be of use in health care practice. In order to validate the index for predicting abnormal Pap test, additional longitudinal studies are preferable.

In summary, our study has identified four risk factors independently associated with risk for cervical cancer (early sexual activity, lifetime number of sexual partners, history of STD and multiparity) which were used to create risk index. Practical application of the index would be

simple. It would be serve primary health care workers to identify women with increased risk for cervical cancer, which could be significant for health education, screening recommendation and future research.

## References

- Al-Daraji WI, Smith JH (2009). Infection and cervical neoplasia: facts and fiction. *Int J Clin Exp Pathol*, **2**, 48-64.
- Antic LG, Vukovic DS, Vasiljevic MD, Antic DZ, Aleksopoulos HG (2014). Differences in risk factors for cervical dysplasia with the applied diagnostic method in Serbia. *Asian Pac J Cancer Prev*, **15**, 6697-701.
- Castellsague X, Muñoz N (2003). Chapter 3: Cofactors in human papillomavirus carcinogenesis--role of parity, oral contraceptives, and tobacco smoking. *J Natl Cancer Inst Monogr*, **31**, 20-8.
- Cole J, Logan TK, Shannon L (2007). Risky sexual behavior among women with protective orders against violent male partners. *AIDS Behav*, **11**, 103-12.
- Collins S, Rollason TP, Young LS, Woodman CB (2010). Cigarette smoking is an independent risk factor for cervical intraepithelial neoplasia in young women: a longitudinal study. *Eur J Cancer*, **46**, 405-11.
- Conti M, Agarossi A, Parazzini F, Muggiasca ML, Boschini A, Negri E, et al (1993). HPV, HIV infection, and risk of cervical intraepithelial neoplasia in former intravenous drug abusers. *Gynecol Oncol*, **49**, 344-8.
- Corral F, Cueva P, Yepez J, Montes E (1996). Limited education as a risk factor in cervical cancer. *Bull Pan Am Health Organ*, **30**, 322-329.
- International Collaboration of Epidemiological Studies of Cervical Cancer (2006). Carcinoma of the cervix and tobacco smoking: collaborative reanalysis of individual data on 13,541 women with carcinoma of the cervix and 23,017 women without carcinoma of the cervix from 23 epidemiological studies. *Int J Cancer*, **118**, 1481-95.
- International Collaboration of Epidemiological Studies of Cervical Cancer (2006). Cervical carcinoma and reproductive factors: collaborative reanalysis of individual data on 16,563 women with cervical carcinoma and 33,542 women without cervical carcinoma from 25 epidemiological studies. *Int J Cancer*, **119**, 1108-24.
- International Collaboration of Epidemiological Studies of Cervical Cancer (2007). Comparison of risk factors for invasive squamous cell carcinoma and adenocarcinoma of the cervix: collaborative reanalysis of individual data on 8,097 women with squamous cell carcinoma and 1,374 women with adenocarcinoma from 12 epidemiological studies. *Int J Cancer*, **120**, 885-91.
- International Collaboration of Epidemiological Studies of Cervical Cancer (2009). Cervical carcinoma and sexual behavior: collaborative reanalysis of individual data on 15,461 women with cervical carcinoma and 29,164 women without cervical carcinoma from 21 epidemiological studies. *Cancer Epidemiol Biomarkers Prev*, **18**, 1060-9.
- Kesic V, Poljak M, Rogovskaya S (2012). Cervical cancer burden and prevention activities in Europe. *Cancer Epidemiol Biomarkers Prev*, **21**, 1423-33.
- Kjellberg L, Hallmans G, Ahren AM, et al (2000). Smoking, diet, pregnancy and oral contraceptive use as risk factors for cervical intra-epithelial neoplasia in relation to human papillomavirus infection. *Br J Cancer*, **82**, 1332-8.
- Lofters AK, Moineddin R, Hwang SW, Glazier RH (2013). Does social disadvantage affect the validity of self-report for cervical cancer screening? *Int J Womens Health*, **5**, 29-33.
- Maree JE, Wright SC, Makua TP (2011). Men's lack of



- knowledge adds to the cervical cancer burden in South Africa. *Eur J Cancer Care (Engl)*, **20**, 662-8.
- Nelson W, Moser RP, Gaffey A, Waldron W (2009). Adherence to cervical cancer screening guidelines for U.S. women aged 25-64: data from the 2005 Health Information National Trends Survey (HINTS). *J Womens Health*, **18**, 1759-68.
- Pacarada M, Lulaj S, Kongjeli G, Kongjeli N, Qavdarbasha H (2010). Factors associated with pathologic colposcopic and cytologic changes in 500 clinically asymptomatic women. *Int J Gynaecol Obstet*, **108**, 7-11.
- Patil V, Wahab SN, Zodpey S, Vasudeo ND (2006). Development and validation of risk scoring system for prediction of cancer cervix. *Indian J Public Health*, **1**, 38-43.
- Petry KU, Luyten A, Justus A, et al (2013). Prevalence of high-risk HPV types and associated genital diseases in women born in 1988/89 or 1983/84--results of WOLVES, a population-based epidemiological study in Wolfsburg, Germany, **13**, 135.
- Raychaudhuri S, Mandal S (2012). Socio-demographic and behavioural risk factors for cervical cancer and knowledge, attitude and practice in rural and urban areas of North Bengal, India. *Asian Pac J Cancer Prev*, **13**, 1093-6.
- Reich O (2005). Is early first intercourse a risk factor for cervical cancer? *Gynakol Geburtshilfliche Rundsch*, **45**, 251-6.
- Reiter P, Katz M, Ferketich A, et al (2009). Measuring cervical cancer risk: development and validation of the CARE risky sexual behavior index. *Cancer Causes Control*, **20**, 1865-71.
- Renschmidt C, Kaufmann AM, Hagemann I, et al (2013). Risk factors for cervical human papillomavirus infection and high-grade intraepithelial lesion in women aged 20 to 31 years in Germany. *Int J Gynecol Cancer*, **23**, 519-26.
- Schoueri-Mychasiw N, McDonald PW (2013). Factors associated with underscreening for cervical cancer among women in Canada. *Asian Pac J Cancer Prev*, **14**, 6445-50.
- Schroder KE, Carey MP, Venable PA (2003). Methodological challenges in research on sexual risk behavior: II. Accuracy of self-reports. *Ann Behav Med*, **26**, 104-23.
- Sherris J, Wittet S, Kleine A, et al (2009). Evidence-based, alternative cervical cancer screening approaches in low-resource settings. *Int Perspect Sex Reprod Health*, **35**, 147-54.
- Smith JS, Green J, Berrington de Gonzalez A, et al (2003). Cervical cancer and use of hormonal contraceptives: a systematic review. *Lancet*, **5**, 1159-67.
- Stone KM, Zaidi A, Rosero-Bixby L, et al (1995). Sexual behavior, sexually transmitted diseases, and risk of cervical cancer. *Epidemiol*, **6**, 409-14.
- Uysal A, Birsal A (2009). Knowledge about cervical cancer risk factors and pap testing behavior among Turkish women. *Asian Pac J Cancer Prev*, **10**, 345-50.
- Vrdoljak E, Wojtkiewicz MZ, Bodoky G, et al (2011). Cancer epidemiology in Central and South Eastern European Countries. *Croat Med J*, **52**, 478-87.
- Wilkinson CE, Peters TJ, Stott NC, Harvey IM (1994). Prospective evaluation of a risk scoring system for cervical neoplasia in primary care. *Br J Gen Pract*, **44**, 341-4.