

NEW APPROACH TO DETECTION OF ABNORMAL CERVICAL CELLS

*Branislava Jeftić**, Lidija Matija, Djuro Koruga

University of Belgrade, Faculty of Mechanical Engineering,
Kraljice Marije 16, 11000 Belgrade, Serbia

Abstract: Optomagnetic Imaging Spectroscopy demonstrated high percentages of accuracy in biological sample classification, namely cervical, oral and colon samples. It enables detection of abnormal tissue and cells, and thus can be used as a diagnostic tool in screening programs. Papanicolaou smears and liquid based cytology samples were analysed in previous studies on cervical cancer detection by Optomagnetic Imaging Spectroscopy and it was shown that this method can differentiate normal healthy tissue from the cancer tissue. So far, only binary classification of the cervical samples was performed based on optomagnetic spectra of the samples. In this paper, classification of the Papanicolaou smears into four groups (II, III, IV and V Papanicolaou groups) was tested with the Random Forest classification model that demonstrated interclass sensitivity of 49.25%, 58.97%, 50%, 44.44% for II, III, IV and V Papanicolaou group respectively, and specificity of 65.26%, 54.76%, 98.70% and 98.69% for II, III, IV and V Papanicolaou group respectively.

Keywords: Optomagnetic Imaging Spectroscopy, cervical cancer, screening test.

1. INTRODUCTION

Cervical cancer still remains the fourth most common cancer in women for both incidence and mortality regardless of all efforts made through the last decades to improve screening methods, prevention, diagnosis and treatment. According to WHO (World Health Organization) estimated number of new cases in 2018 was 570,000 [1]. Effective screening programs are implemented in developed countries, where early detection of precancerous lesions enables on-time treatment and leads to decrease of the death rate of cervical cancer. Vaccines against HPV infections have been introduced in developed countries since 2006. They serve as primary prevention of cervical cancer and are aimed to successfully decrease the number of HPV infections i.e. the cervical cancer incidence. Recent studies show high effectiveness of the HPV vaccine and lower risk of CIN2+ in vaccinated women compared to unvaccinated women [2–4]. Still, the cost of the vaccine programmes is high and the implementation in low income countries is restricted. In less developed countries patients usually don't have access even to the Papanicolaou screening test, making the disease undetectable until it is in a progressed stage. This is why 90% of all cervical cancer deaths occurs in less developed countries [1].

Optomagnetic imaging spectroscopy has demonstrated high accuracy in detecting abnormal cells in cervical smears, in colon tissues and oral cavity [5–7]. In previous studies Optomagnetic Imaging spectroscopy was used to screen cervical smears and to obtain optomagnetic spectra of the samples. Based on the optomagnetic fingerprint of the samples, binary classification was done: first class consisted of negative cases (Papanicolaou group II), while the second class consisted of abnormal samples (III, IV and V Papanicolaou group individually or three groups combined). In this study, classification of the cervical samples into four classes is tested, where four classes correspond to four Papanicolaou groups: Papanicolaou group II (samples that have no cell abnormalities, reported as negative for intraepithelial lesion or malignancy), Papanicolaou group III (samples containing abnormal cells), Papanicolaou group IV (carcinoma *in situ*) and Papanicolaou group V (cancer).

2. MATERIALS AND METHODS

Cervical samples were collected in Gynaecology-Obstetric clinic "Narodni Front", Belgrade, Serbia from 2011 to 2016. Two cervical smears were made for each patient, one from the cells collected from the cervical canal and the other

* Corresponding author: jeftic.branislava@gmail.com

from the exocervix. Samples were fixated and stained according to the standard Papanicolaou procedure. Total number of 1632 samples was collected: 672 samples reported as Papanicolaou group II, 784 reported as Papanicolaou group III, 82 samples reported as Papanicolaou group IV and 94 samples reported as Papanicolaou group V. All considered cervical samples were screened by Optomagnetic imaging spectroscopy and based on the sample spectra classification of the samples was conducted.

A set of unstained samples was also tested with a total number of 336 cervical samples: 254 samples reported as II Papanicolaou group, 50 samples reported as III Papanicolaou group, 16 samples reported as IV Papanicolaou group and 16 samples reported as V Papanicolaou group. Unstained samples were collected the same way as conventional stained samples, smeared on the microscopic slide, but they were not stained nor fixated.

Optomagnetic Imaging Spectroscopy (OMIS) is a nanophysical method for the material characterization. It is based on light-matter interaction, i.e. interaction between white diffuse light and the valence electrons within the material. According to the principles of Optomagnetic Imaging spectroscopy method, when the sample is illuminated with the white diffuse light perpendicular to the sample and then with the white diffuse light under the Brewster's angle, unique spectral fingerprint of the sample will be captured in the optomagnetic sample spectrum based on spectral convolution of the RGB (Red, Green and Blue) color channels in the digital images of the sample [8]. White diffuse light is emitted by the diodes integrated into the OMIS device, while the polarisation of the white diffuse light is achieved through the reflection of the white diffuse light emitted under the Brewster's angle. Digital images of the sample under the white diffuse light and polarised white light are taken and by combining the information captured in the digital images, i.e. RGB color channels in the white (W) and polarised (P) mode, the (W-P)(R-B) spectral fingerprint is obtained. OMIS device for cervical sample scanning is shown in the Figure 1.

The sample is placed on the plate, illuminated with white diffuse light perpendicular to the sample, and a digital image of the sample is captured. Then, the sample is illuminated with white difusse light

under the Brewster's angle and the second image of the sample is taken. Two images are processed with the convolution algorithm in MATLAB® 2013a (Math-Works, USA) and the optomagnetic spectra of the sample is obtained.



Figure 1. OMIS device

Characteristic parameters in the optomagnetic spectra of the samples are used to determine whether the sample contains normal or abnormal cervical cells. Based on the parameters classification of the samples is done. Intensities of the maximum and minimum peaks and areas under the positive and negative peaks were used as features in a Random Forest classification model developed in R, a free programming language and environment for statistical computing and graphics.

3. RESULTS AND DISCUSSION

Cervical smears were collected and screened by Optomagnetic imaging spectroscopy. Digital images of the samples were then processed with convolution algorithm in order to produce optomagnetic spectra of the samples. Typical optomagnetic spectra for the unstained and stained cervical samples reported as II, III, IV and V Papanicolaou (PA) group are shown on the Figure 2 and Figure 3.

In the case of unstained cervical samples, optomagnetic spectra obtained for II Papanicolaou group sample coincides with the III Papanicolaou group sample (Figure 2), with small differences present in the maximal and minimal peak intensities, while spectra obtained for the IV and V Papanicolaou group samples notably differ compared to the II and III Papanicolaou group sample spectra.

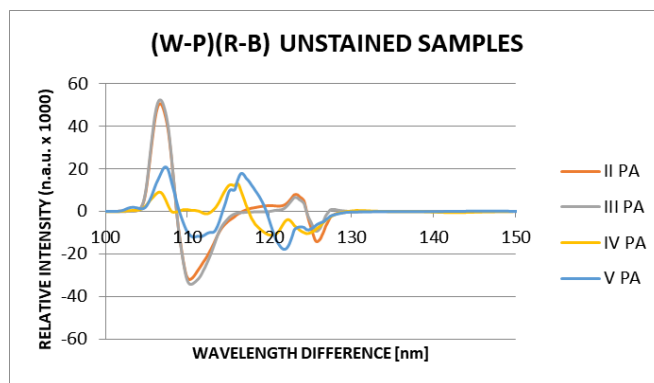


Figure 2. Optomagnetic spectra obtained for the unstained cervical samples reported as II, III, IV and V Papanicolaou (PA) group based on convolution algorithm that uses R and B colour channels in the digital images of the samples and combines them (W-P)(R-B) in order to get a unique spectral fingerprint of the sample

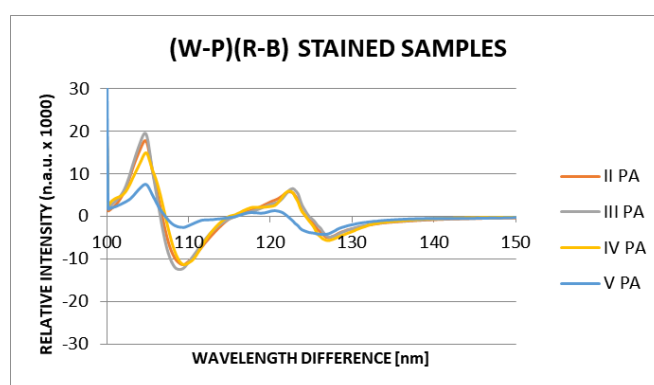


Figure 3. Optomagnetic spectra obtained for the stained cervical samples reported as II, III, IV and V Papanicolaou (PA) group based on convolution algorithm that uses R and B colour channels in the digital images of the samples and combines them (W-P)(R-B) in order to get a unique spectral fingerprint of the sample

Stained samples spectra for cancer samples show lower peak intensities compared to the spectra obtained for the II and III Papanicolaou group samples, same as the spectra obtained for the unstained cancer samples.

Following values from the spectra were used as feature sets in four class classification problems based on a Random Forest model: maximum peak

intensity values, minimum peak intensity values, wavelengths where maximum and minimum peaks occur, area under the positive peaks and area under the negative peaks. Classification algorithm was developed in R software package. For the model performance evaluation, interclass sensitivity, specificity, and accuracy were used (Table 1 and Table 2).

Table 1. Performance of the Random Forest classification model in the four class classification problem solved for the unstained cervical samples. Four classes are designated as II, III, IV and V Papanicolaou group.

Random Forest Model	Classes			
	II PA	III PA	IV PA	V PA
Sensitivity	100%	20%	33.33%	33.33%
Specificity	25%	100%	100%	100%
Accuracy	81.82%			
95% CI	(0.7039, 0.9024)			

Table 2. Performance of the Random Forest classification model in the four class classification problem solved for the stained cervical samples. Four classes are designated as II, III, IV and V Papanicolaou group.

Random Forest Model	Classes			
	II PA	III PA	IV PA	V PA
Sensitivity	49.25%	58.97%	50%	44.44%
Specificity	65.26%	54.76%	98.70%	98.69%
Accuracy	53.7%			
95% CI	(0.4811, 0.5923)			

Sensitivity values in the case of unstained cervical samples are low for the samples in the III, IV and V Papanicolaou group (Table 1), while in the case of stained cervical samples the sensitivity percentages are somewhat higher (Table 2). Specificity values are relatively high for both stained and unstained samples, except for the unstained samples in the II Papanicolaou group.

Previous studies on cervical samples have reported a high degree of accuracy in cervical cancer detection by Optomagnetic imaging spectroscopy. Tested on the set of 69 negative cases (II Papanicolaou group) and 66 cancer cases (V Papanicolaou group), binary classification using Naïve Bayes classifier gave the accuracy of 85.18% (sensitivity was 97.92% and specificity was 78.16%) [9]. Previous study on the unstained cervical samples that tested sample classification by OMIS into healthy/cancerous group with Naïve Bayes classifier gave even higher accuracy of 96% (sensitivity was 75%, and specificity was 97%) compared to the stained samples [5]. In this study, binary classification based on the Random Forest model where one class was made of the samples reported as II and III Papanicolaou group and the other was made of the samples reported as IV and V Papanicolaou group gave the accuracy of 92.33% (sensitivity of 37.14%, specificity of 98.97%) for the stained samples and the accuracy of 96.97% (sensitivity of 56.67%, specificity of 100%) for the unstained samples, with the area under the ROC (Receiver Operating Characteristic) of 0.68 for stained and AUC of 0.83 for unstained samples. The high percentages of accuracy obtained for the binary classification problem are promising, indicating the possibility of OMIS to differentiate negative cases from the cancer cases. The sensitivity values in the four class classification as well as in the binary classification of the cervical samples in this study are lower than sensitivity obtained in previous studies, especially in the case of unstained samples, due to the uneven distribution of sample number in the classes. In the binary classification problem, first class consisted of 1455 stained samples and second class consisted of 175 stained samples (304 unstained samples and 32 unstained samples), and in four class classification problem classes consisted of 672, 784, 82 and 94 stained samples and 254, 50, 16 and 16 unstained samples. To overcome this problem, more cancer samples need to be collected. The sensitivity and specificity of the Papanicolaou test in detecting high grade intraepithelial neoplasia are 55.4 and 96.8%, thus the values of the sensitivity and specificity of the OMIS in detecting abnormal cervical cells are in the range of the so far most efficient screening test [10].

4. CONCLUSION

Optomagnetic imaging spectroscopy has proven to be a fast, easy to use and an accurate method for the detection of abnormal cervical samples. The main advantage of the Optomagnetic imaging spectroscopy is that it can be implemented in low resource settings, because it doesn't require special infrastructure. In previous studies, binary classification of cervical samples, both stained and unstained, based on their optomagnetic spectra showed a high degree of accuracy (85.18% for stained and 96% for unstained samples). This way, Optomagnetic imaging spectroscopy can be used to refer patients with reported result that classify the sample into non-healthy group to further testing. In this study, classification in four classes was tested with the Random Forest classification model in order to develop an algorithm that would classify a cervical sample based on its optomagnetic spectra into one of the four Papanicolaou groups. Higher percentages of sensitivity were obtained for the stained cervical samples compared to the unstained samples. Classification of the stained cervical smears into four groups (II, III, IV and V Papanicolaou groups) tested with the Random Forest classification model demonstrated interclass sensitivity of 49.25%, 58.97%, 50%, 44.44% for II, III, IV and V Papanicolaou group respectively, and specificity of 65.26%, 54.76%, 98.70% and 98.69% for II, III, IV and V Papanicolaou group respectively. A higher number of negative cases compared to the number of cancer cases is evident, thus the next step would be to collect more samples reported as IV and V Papanicolaou group and evaluate tested classification models on the new set of data in order to improve the results obtained in this study.

5. ACKNOWLEDGEMENT

This research has been funded by the Ministry of Education, Science and Technological Development of the Republic of Serbia, through Project III 41006.

6. REFERENCES

- [1] WHO, (n.d.). <https://www.who.int/cancer/prevention/diagnosis-screening/cervical-cancer/en/>.
- [2] R. Donken, A. J. King, J.A. Bogaards, P. J. Woestenbergh, C. Meijer, H. E. de Melker, *High*

effectiveness of the bivalent human papillomavirus (HPV) vaccine against incident and persistent HPV infections up to 6 years after vaccination in young Dutch women, J. Infect. Dis., Vol. 217 (2018) 1579–1589.

[3] R. Konno, H. Konishi, C. Sauvaget, Y. Ohashi, T. Kakizoe, *Effectiveness of HPV vaccination against high grade cervical lesions in Japan*, Vaccine, Vol. 36 (2018) 7913–7915.

[4] T. Palmer, L. Wallace, K. G. Pollock, K. Cuschieri, C. Robertson, K. Kavanagh, M. Cruickshank, *Prevalence of cervical disease at age 20 after immunisation with bivalent HPV vaccine at age 12-13 in Scotland: retrospective population study*, Bmj., Vol. 365 (2019) |1161|.

[5] B. Jeftić, M. Papić-Obradović, J. Munćan, L. Matija, D. Koruga, *Optomagnetic Imaging Spectroscopy Application in Cervical Dysplasia and Cancer Detection: Comparison of Stained and Unstained Papanicolaou Smears*, J. Med. Biol. Eng., Vol. 37 (2017) 936–943.

[6] A. Dragičević, L. Matija, Z. Krivokapić, I. Dimitrijević, M. Baroš, D. Koruga, *Classification of Healthy and Cancer States of Colon Epithelial*

Tissues Using Opto-magnetic Imaging Spectroscopy, J. Med. Biol. Eng., Vol. 39 (2019) 367–380.

[7] B. Lisul, D. Jelovac, M. Petrović, Z. Tepavčević, D. Koruga, D. Grga, *Predictive Value of Opto-magnetic Imaging Spectroscopy in Discriminating Oral Squamous Cell Carcinoma from Non-tumor Tissue in Surgical Margins*, J. Med. Biol. Eng., Vol. (2019) 1–11.

[8] D. Koruga, A. Tomić, *System and method for analysis of light-matter interaction based on spectral convolution*, US Patent Pub. No.: 2009/0245603, Pub. Date: Oct. 1, (2009).

[9] L. Matija, B. Jeftić, G. Nikolić, A. Dragičević, I. Mileusnić, J. Munćan, D. Koruga, *Nanophysical approach to diagnosis of epithelial tissues using Opto-magnetic imaging spectroscopy*, Nanomedicine, Vol. 21 (2014) 21.

[10] M. Karimi-zarchi, L. Zambagh, A. Shafii, S. Taghipour-Zahir, S. Teimoori, P. Yazdian-Anari, *Comparison of pap smear and colposcopy in screening for cervical cancer in patients with secondary immunodeficiency*, Electron. Physician., Vol. 7 (2015) 1542.



НОВИ ПРИСТУП ОТКРИВАЊА АБНОРМАЛНИХ ЋЕЛИЈА ГРЛИЋА МАТЕРИЦЕ

Сажетак: Оптомагнетна имицинг спектроскопија показала је висок проценат тачности у класификацији биолошких ткива, посебно када се ради о узорцима грлића материце, усне дупље и дебелог црева. Ова метода омогућава детекцију абнормалног ткива и ћелија и стога може да се користи као дијагностички тест у оквиру скрининг програма. У претходним студијама анализирани су цервикални узорци и узорци добијени техником цитологије на течној бази помоћу оптомагнетне имицинг спектроскопије у циљу детекције карцинома грлића материце и показано је да ова метода може да раздвоји здраво ткиво од канцера. До сада су тестирани само бинарни класификациони модели за класификацију цервикалних узорака на бази оптомагнетних спектра узорака. У овом раду тестирана је класификација Папаниколау размаза у четири класе (II, III, IV и V Папаниколау групе) коришћењем класификационог модела на бази Рандом Форест класификатора који је показао интеркласну сензитивност од 49,25%, 58,97%, 50%, 44,44% за II, III, IV и V Папаниколау групу респективно, и специфичност од 65,26%, 54,76%, 98,70% и 98,69% за II, III, IV и V Папаниколау групу респективно.

Кључне речи: оптомагнетна имицинг спектроскопија, карцином грлића материце, скрининг тест.

