

Digitization of Gynecology Using Artificial Intelligence: Cervical Mapping Corroborated With Clinical Data for Conization Necessity

Dorina Adelina Minciună¹, Demetra Gabriela Socolov¹, Attila Szócs², Doina Ivanov¹, Tudor Gîscă¹, Valentin Nechifor¹, Sándor Budai³, Ákos Bálint³, Răzvan Socolov¹

¹ "Grigore T. Popa" University of Medicine and Pharmacy, Iași, Romania

² Ascorb Research S.R.L., Târgu Mureș, Romania

³ Cattus Distribution S.R.L., Târgu Mureș, Romania

CORRESPONDENCE

Attila Szócs

Bd. 22 Decembrie 1989 nr. 15/20
Târgu Mureș, Romania
Tel: +40 758 360 083
Email: szocsatti@gmail.com

ARTICLE HISTORY

Received: September 3, 2023

Accepted: September 12, 2023

Dorina Adelina Minciună
Demetra Gabriela Socolov
Doina Ivanov
Tudor Gîscă
Valentin Nechifor
Sándor Budai
Ákos Bálint
Răzvan Socolov

ABSTRACT

Background: Cervical cancer is the fourth most common female malignancy worldwide. In developing countries, it is the most common subtype of cancer and the third leading cause of cancer mortality among women. Artificial intelligence has the potential to be of real use in the prevention and prompt diagnosis of cervical cancer. The aim of our study was to develop a medical platform consisting of an automated observation sheet containing colposcopy data, a software that would use a machine learning module based on clinical and image data for diagnosis and treatment, and a telemedicine module to enable collaboration between gynecologists. **Materials and methods:** Clinical and colposcopy image data from 136 patients were introduced into a machine learning module designed to generate an algorithm for proposing a preliminary diagnosis and treatment. The clinical and imaging data were corroborated to generate six options: 'Follow-up', 'Pharmacotherapy', 'Biopsy', 'Curettage', 'DTC', and 'Conization'. **Results:** Data generated by the machine learning module regarding treatment options were compared with the opinion of gynecologists and yielded an accuracy of 78% for 'Follow-up', 81% for 'Pharmacotherapy', 84% for 'Biopsy', 90% for 'Curettage', 96% for 'DTC', and 81% for 'Conization'. **Conclusions:** The developed software can be an important step towards the digitization of existing gynecology offices and the creation of intelligently automated gynecology offices related to prevention and treatment of cervical cancer. More data is needed to improve the accuracy of the developed software.

Keywords: cervical cancer, colposcopy, human papilloma virus, machine learning, artificial intelligence

INTRODUCTION

Cervical cancer (CC) is the fourth most common female malignancy worldwide.¹ In developing countries, it is the most common subtype of cancer and the third leading cause of cancer-related mortality among women. One of the most important discoveries in the etiological investigation of cancer in the last 25 years was that CC is caused by persistent infection by certain genotypes of human papillomavirus (HPV).¹⁻³ Accumulating scientific evidence from virological, molecular, clinical and epidemiological studies has unequivocally demonstrated that CC is in fact a consequence of long-term unresolved infection with certain HPV genotypes. Thus, we can state that CC is the result of a viral infection, and preventing and treating the infection at the right time is the most important strategy to consider in the primary prevention of CC and other diseases caused by HPV. More than 95% of cervical cancers (squamous cell carcinomas and adenocarcinomas) have been attributed to infection with high-risk HPV strains including HPV 16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, and 59 (these viruses are assigned by the IARC to Group 1). However, the most common (~70%) strains involved in cervical carcinogenesis are HPV 16 and 18.¹⁻³

HPV infection is basically a sexually transmitted disease. As such, both men and women are involved in the epidemiological chain of infection. There are asymptomatic carriers, transmitters, and victims of HPV infection simultaneously.¹⁻⁵ In this sense, the risk factors associated with HPV infection are clearly related to the individuals' sexual behavior. Epidemiological and clinical studies incorporating high-sensitivity molecular biology techniques in biological specimens have detected oncogenic or high-risk HPV in almost all cases of CC. HPV DNA is detected in most (70–90%) precursor lesions or high-grade intraepithelial lesions (CIN II–III) and in a smaller fraction (20–50%) in low-grade lesions (CIN I) as well.¹⁻³ However, it is also worth mentioning that in the category of cytological lesions of uncertain nature (ASCUS and AGUS) the HPV infection rate is close to 50%.¹

Another important aspect from an epidemiological point of view is the association of Epstein–Barr virus (EBV) and HPV co-infection with the development of CC. Even if this correlation is not very obvious, there are numerous studies that draw attention to this aspect. EBV co-infection with HPV confers a four-fold increased risk of CC in EBV-positive women.¹ Similarly, precancerous cervical lesions are twice as common in EBV-positive women compared to EBV-negative cases.⁴ These data point to EBV as a potentially active cofactor in the pathogenesis

and progression of CC. The presence of EBV in the cervix may also accelerate the integration of HPV genome into the genome of cervical cells, increasing the genomic instability of infected cervical cells. In addition, chronic cervicitis may also facilitate EBV infection and its potential oncogenic effects. These viruses, alone or in collaboration, can induce oncogene activation and epithelial–mesenchymal transition, one of the key steps in tumor progression and metastasis.⁵ Based on the analyzed studies, we can draw the conclusion that this area of research represents a worldwide discussion.⁴⁻¹⁹

Colposcopy is often used as a follow-up procedure when a Pap smear (Pap test) indicates abnormal cervical cells. The colposcope allows the healthcare provider to visualize the cervix more clearly and identify areas of concern that may require further evaluation or treatment. Colposcopy can help in identifying and diagnosing various cervical abnormalities such as cervical dysplasia, HPV infection, and cervical inflammation.²⁰ Colposcopy is also used to monitor the effectiveness of treatment for cervical abnormalities. After treatment, regular colposcopy examinations can help ensure that the treated area is healing properly and that no further abnormalities are present. Colposcopy can also guide surgical procedures aimed at removing abnormal tissue. It can be clearly stated that colposcopy images are crucial in the diagnosis of CC.²⁰⁻²³

The main purpose of this study was the digitization of gynecology clinics by developing a software that would aid the recognition of malignant and premalignant lesions caused by CC. Our aim was to develop a digital platform that is useful for gynecologists from several points of view and contain the following: 1) an online medical record containing colposcopy data that can be filled out quickly and easily, acting as an automated observation sheet that generates highly accurate statistical data both for medical practice and for clinical studies; 2) a software for establishing preliminary diagnosis and treatment; 3) a telemedicine module.

MATERIALS AND METHODS

Development of an automated medical record platform with telemedicine module

In the first stage, an observation sheet regarding CC was created, containing the following data: age; reason of examination; antecedents for cytology, HPV infection, colposcopy, cervical biopsy, and conization; current colposcopy results; squamous-cylindrical junction: upper limit, transformation zone; abnormal aceto-white area: aceto-white epithelium, punctations, mosaic, leukoplakia, mi-

croinvasion, invasion, Schiller test, iodine negative areas, iodine negative areas corresponding to aceto-white epithelium, contour of iodine negative areas, atypical vessel, other aspects (polyp, vegetation, condyloma, leukoplasia, adenosis, inflammation, atrophy, stenosis, Nabothian cyst, ulceration); SWEDE score; diagnosis, recommended treatment. An automated medical record platform was created based on this observation sheet.

The software was developed using clinical and colposcopy data from the medical records of 136 patients admitted to the Clinica Avicena Profertis gynecology hospital in Iași, Romania. As the software is intended for the recognition of malignant and premalignant lesions necessary for a preliminary diagnosis and treatment, the clinical data were introduced into a machine learning module. All data were anonymized according to the legislation in force. A telemedicine module entitled 'Cervix map' was added to the platform, allowing collaboration and opinion exchange between doctors.

Development of the software

The developed software is based on a machine learning module. The development included two stages: 1) feature extraction from clinical data; 2) feature extraction from colposcopy images.

Feature extraction from clinical data

In the first stage, clinical data extracted from the medical records were introduced into a machine learning module for training purposes. The module was designed to generate an algorithm for proposing a preliminary diagnosis and treatment. For this purpose, we trained decision tree and random forest machine learning models. Based on the data from the observation sheets, the module provides six treatment options: 'Follow-up', 'Pharmacotherapy', 'Biopsy', 'Curettage', 'DTC' (diathermocoagulation), and 'Conization/LLETZ' (large loop excision of the transformation zone). In the first step, the module generates a decision tree for each of the six options. Figure 1 presents the decision tree generated for conization/LLETZ.

In the case of random forest models, the weight of each input parameter can be visualized. The results generated by this method are based on the importance of features for each of the six options mentioned above. Random forest models generated for follow-up and pharmacotherapy are presented in Figure 2 and for DTC and conization in Figure 3.

Feature extraction from colposcopy images

At this stage, our goal was to collect a large set of colposcopy images. We collected a set of 550 colposcopy images from

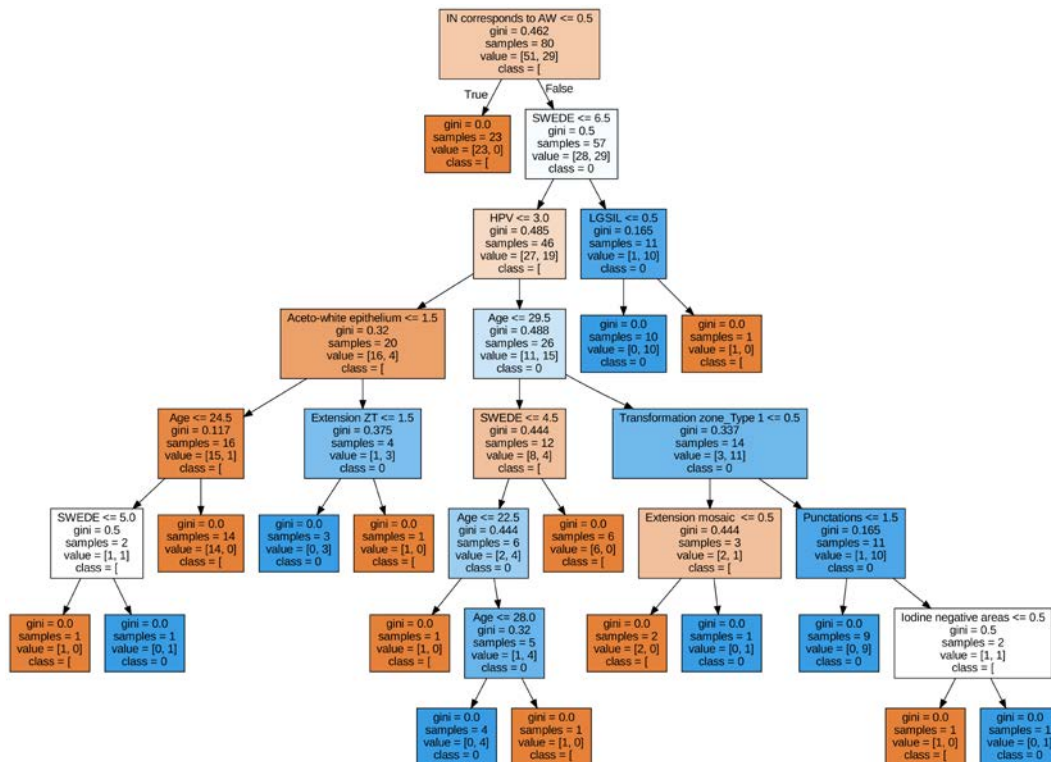


FIGURE 1. Decision tree for conization/LLETZ

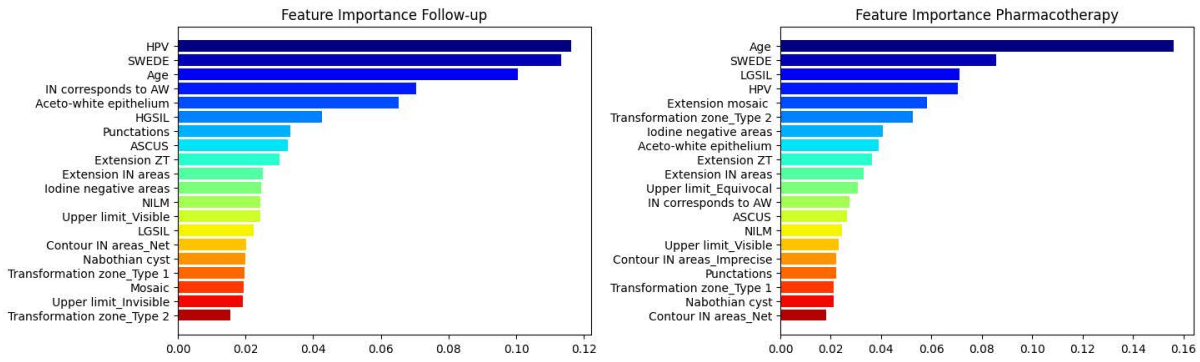


FIGURE 2. Random forest model for follow-up and pharmacotherapy

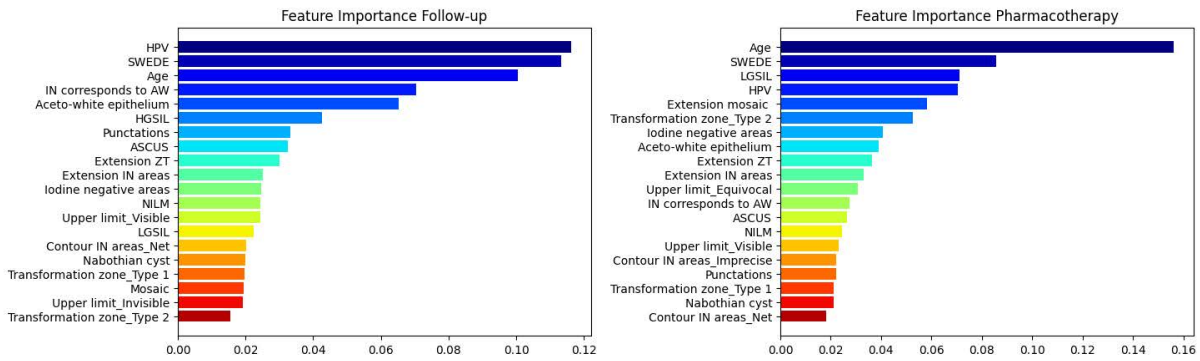


FIGURE 3. Random forest model for DTC and conization

the 136 medical records included in the study. In order to be able to differentiate between healthy tissues and pathological changes of the cervical tissue, the images were manually segmented and annotated by specialists. The segmentation used the following colors codes: blue for the squamous-cylindrical junction; purple for the aceto-white area; red for atypical vessels, punctations, and mosaic; yellow for Naboth cysts; black for cuffed gland opening; and white for leukoplasia. Examples of segmented images are presented in Figure 4. The segmented and annotated images were added to the machine learning module, and the features were corroborated with the clinical data obtained from the medical records.

Ethics

The study was conducted in accordance with the principles stated in the Declaration of Helsinki and was approved by the ethics committees of the institutions the authors are affiliated with.

RESULTS

Data generated by the machine learning module regarding treatment options were compared with the opinion of

gynecologists based on the segmented and annotated colposcopy images. Statistical analysis yielded the following results regarding the accuracy of the developed algorithm in suggesting a preliminary treatment: 78% for ‘Follow-up’, 81% for ‘Pharmacotherapy’, 84% for ‘Biopsy’, 90% for ‘Curettage’, 96% for ‘DTC’, and 81% for ‘Conization’.

DISCUSSION

Colposcopy is crucial for the diagnosis of CC, and prevention and prompt diagnosis are of major importance. In this field, artificial intelligence proves to be of real use – the developed automated medical record platform with telemedicine could be important for physicians. In order to develop a lesion recognition software that is capable of establishing a preliminary diagnosis and treatment is important to corroborate clinical data with colposcopy image data. Of note, in our dataset the number of features was close to the number of records. As a consequence, machine learning models might find correlations between the output (treatment) and certain features that are irrelevant from a medical point of view. These correlations could be coincidences that exist only in this dataset, meaning that much more data is needed to increase the reliability of our models.

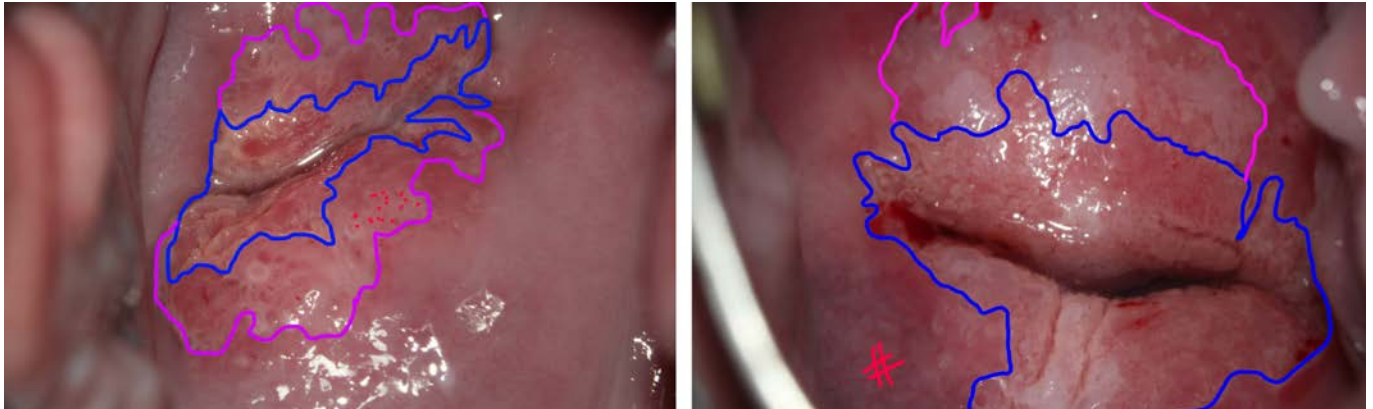


FIGURE 4. Segmented colposcopy images

CONCLUSIONS

The developed software can generate an algorithm for establishing a preliminary diagnosis and treatment. To increase the prediction accuracy of this software, it is necessary to collect a larger volume of data, taking into account in particular the corroboration of clinical data and colposcopy images. Our intention is to further develop this software. We consider our software to be an important step towards the digitization of existing gynecology offices and the development of intelligently automated gynecology offices related to the prevention and treatment of CC. We intend to collect more data and improve the accuracy of this software.

ACKNOWLEDGEMENT

This work was supported by Financing contract no. 75/221_ap2/21.08.2020, in the Competitiveness Operational Program 2014–2020, Priority Axis 2 – “Information and communication technology (TIC) for a competitive digital economy”, SMIS Code 130106.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

REFERENCES

- Xavier Castellsagué X. Natural history and epidemiology of HPV infection and cervical cancer. *Gynecol Oncol.* 2008; 110:S4–S7.
- Pimple S, Mishra G. Cancer cervix: Epidemiology and disease burden. *Cytojournal.* 2022;19:21.
- World Health Organization. Cervical cancer. <https://www.who.int/news-room/fact-sheets/detail/cervical-cancer>
- Chao A, Tsai Ch, Hsueh S. et al. Does Epstein-Barr virus play a role in lymphoepithelioma-like carcinoma of the uterine cervix? *Int J Gynecol Pathol.* 2009;28:279–285.
- Vranic S, Cyprian FS, Akhtar S. et al. The Role of Epstein–Barr virus in Cervical Cancer: A Brief Update. *Front Oncol.* 2018;8:113.
- Cyprian FS, Al-Farsi H, Vranic S. et al. Epstein-Barr Virus and Human Papillomaviruses Interactions and Their Roles in the Initiation of Epithelial-Mesenchymal Transition and Cancer Progression. *Front Oncol.* 2018;8:111.
- Pereira de Lima MA, Neto PJN, Lima LM, et al. Association between Epstein-Barr virus (EBV) and cervical carcinoma: A meta-analysis. *Gynecol Oncol.* 2018;148:317–328.
- Oliviera LHS, Santos LS, Nogueira FG. Epstein Barr virus detection in cervical samples of women living with human immunodeficiency virus. *Rev Inst Med Trop.* 2011;53:231–234.
- Purtilo DT, Okano M and Grierson HL. Immune deficiency as a risk factor in Epstein-Barr virus-induced malignant diseases. *Environ Health Perspect.* 1990;88:225–230.
- Feng M, Duan R, Gao Y, et al. Role of Epstein-Barr Virus and Human Papillomavirus Coinfection in Cervical Intraepithelial Neoplasia in Chinese Women Living With HIV. *Front Cell Infect Microbiol.* 2021;11:703259.
- Hachisuga T, Ookuma Y, Fukuda K, et al. Detection of Epstein-Barr virus DNA from a lymphoma-like lesion of the uterine cervix. *Gynecol Oncol.* 1992;46:69–73.
- Hørding U, Daugaard, S, Bock JE. Human papillomavirus, Epstein-Barr virus, and cervical carcinoma in Greenland. *Int J Gynecol Cancer.* 1992;2:314–317.
- Ida K, Tokuda H, Kanaoka T. et al. Epstein-Barr virus activating principle in husbands' semen of cervical cancer patients. *Am J Reprod Immunol.* 1991;26:89–92.
- Jezzoni JC, Gaffey MJ, Weiss LM. The Role of Epstein-Barr Virus in Lymphoepithelioma-like Carcinomas. *Am J Clin Pathol.* 1995;103:308–315.
- Saloua Kahla S, Oueslati S, Achour M. Correlation between EBV co-infection and HPV16 genome integrity in Tunisian cervical cancer patients. *Braz J Microbiol.* 2012;43:744–753.
- Khenchouche A, Sadouki N, Boudriche A. Human papillomavirus and Epstein-Barr virus co-infection in cervical carcinoma in Algerian women. *Virology.* 2013;10:340.
- Kim NR, Lin Z, Kim KR. et al. Epstein-Barr virus and p16INK4A methylation in squamous cell carcinoma and precancerous lesions of the cervix uteri. *J Korean Med Sci.* 2005;20:636–642.
- Kitano Y, Fujisaki S, Nakamura N. et al. Immunological disorder against the Epstein-Barr virus infection and prognosis in patients with cervical carcinoma. *Gynecol Oncol.* 1995;57:150–157.
- Landers RJ, O'Leary JJ, Crowley M. et al. Epstein-Barr virus in normal, pre-malignant, and malignant lesions of the uterine cervix. *J Clin Pathol.* 1993;46:931–935.
- Prendiville W, Sankaranarayanan R. Colposcopy and Treatment of Cervical Precancer. Lyon (FR): International Agency for Research on Cancer; 2017.
- Valls J, Baena A, Venegas, G. et al. Performance of standardised colposcopy to detect cervical precancer and cancer for triage of women testing positive for human papillomavirus: results from the ESTAMPA multicentric screening study. *Lancet Glob Health.* 2023;11:e350–60.
- Xue P, Ng MTA, Qiao Y. The challenges of colposcopy for cervical cancer screening in LMICs and solutions by artificial intelligence. *BMC.* 2020;169.
- Lanham S, Herbert A, Basarab A. et al. Detection of cervical infections in colposcopy clinic patients. *J Clin Microbiol.* 2001;39:2946–2950.