

# Bayesian Networks to Support the Management of Patients with ASCUS/LSIL Pap Tests

Panagiotis Bountris, Charalampos Tsirmpas, Maria Haritou, Abraham Pouliakis, Petros Karakitsos, and Dimitrios Koutsouris

**Abstract**—In the majority of cases, cervical cancer (CxCa) develops as a result of underestimated abnormalities in the Pap test. Nowadays, there are ancillary molecular biology techniques providing important information related to CxCa and the Human Papillomavirus (HPV) natural history, including HPV DNA test, HPV mRNA tests and immunocytochemistry tests. However, these techniques have their own performance, advantages and limitations, thus a combinatorial approach via computational intelligence methods could exploit the benefits of each method and produce more accurate results. In this paper we present a risk assessment model based on a Bayesian Network which, by combining the results of Pap test and ancillary tests, may identify women at true risk of developing cervical cancer and support the management of patients with ASCUS or LSIL cytology. The model, following the paradigm of other implemented systems, can be integrated into existing platforms and be available on mobile terminals for anytime/anyplace medical consultation.

**Keywords**—cervical cancer; cytology; human papillomavirus (HPV); bayesian networks; risk assessment

## I. INTRODUCTION

Cervical cancer (CxCa) is one of the most frequent female cancers worldwide. The necessary cause for CxCa is infection by human papillomavirus (HPV). Although this is a necessary cause, it leads to disease only in a fraction of the cases, as the majority of infections regress thanks to the human immune system. The study of this virus has led to the development of numerous medical tests related to HPV genetic material.

---

This study was partially funded by the Greek Ministry of Development (General Secretariat for Research and Technology-GSRT), Project “HPVGuard”, Cooperation 2011-2013 (code: 11SYN\_10\_250) <http://hpvguard.org>.

P. Bountris is with the Biomedical Engineering Laboratory, School of Electrical and Computer Engineering, National Technical University of Athens, Athens, Greece, e-mail: pbountris@biomed.ntua.gr

C. Tsirmpas is with the Biomedical Engineering Laboratory, School of Electrical and Computer Engineering, National Technical University of Athens, Athens, Greece, e-mail: htsirbas@biomed.ntua.gr

M. Haritou is with the Institute of Communication and Computer Systems, National Technical University of Athens, Athens, Greece, e-mail: mhari@biomed.ntua.gr

A. Pouliakis is with the Department of Cytopathology, “ATTIKON” University Hospital, University of Athens, Athens, Greece, e-mail: apouliak@med.uoa.gr

P. Karakitsos is with the Department of Cytopathology, “ATTIKON” University Hospital, University of Athens, Athens, Greece, e-mail: pkarak@med.uoa.gr

D. Koutsouris is with the Biomedical Engineering Laboratory, School of Electrical and Computer Engineering, National Technical University of Athens, Athens, Greece, e-mail: dkoutsou@biomed.ntua.gr

For 60 years, the primary, and most valuable, test for preventing CxCa is the cytological examination via test Papanikolaou (Pap test). The outcome of Pap test is standardized into various levels (ranked): (a) within normal limits (WNL), (b) atypical squamous cells of undetermined significance (ASCUS), (c) low-grade squamous intraepithelial lesion (LSIL), (d) atypical squamous cells probably high (ASCH), (e) high-grade squamous intraepithelial lesion (HSIL), and (f) carcinomas (mentioned as CxCa). Specifically for carcinomas there are two groups: squamous cell carcinoma (SCC) or adenocarcinoma (Adeno-Ca). Depending on the Pap test results, women are referred to subsequent treatment or follow up. However, the final diagnosis of each case is confirmed (if required) by subsequent biopsy and histological examination. The histological examination standardization is a multi-tiered Cervical Intraepithelial Neoplasia (CIN) grading system; the categories, ranked according to severity, are: (a) without evidence of malignancy (NEGATIVE), (b) cervical intraepithelial neoplasia grade I (CIN1), (c) grade II or III (CIN2 or CIN3), and (d) CxCa (SCC or Adeno-Ca). Histology is considered as the golden standard however it is a procedure requiring biopsy, thus is neither possible nor ethical to perform in all cases. In contrast, Pap test is a minimally invasive method with no adverse effects.

Given that the usual decision point for the clinicians to surgically treat is histology outcome CIN2 or worse (CIN2+), it is crucial to obtain an accurate prediction of the histological status (CIN status) based on cytological, and/or HPV related information. Achievement of this goal could lead to the optimal management of women with abnormal Pap test, because a part of women having LSIL may actually have CIN2+, women with HSIL may actually have CIN1 or even normal histology and women with ASCUS in cytology present more problems in their management.

Today, there is no consensus for the management of women with abnormal Pap test; women with ASCUS or LSIL are either immediately referred to colposcopy (Fig.1), or advised for a new Pap test after six months. Immediate colposcopy can overload colposcopy clinics and lead to over-treatment due to imperceptible findings. Repeated cytology has the risk of missing HSILs, women may not appear for examination, and there is prolonged and increased psychological burden. Thus, it is extremely important to reduce unnecessary Pap tests, colposcopies and treatments, and a method improving the diagnostic accuracy is crucial.

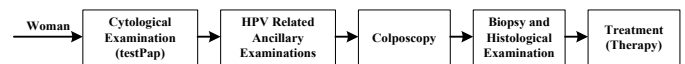


Figure 1. Order of examinations

Recently HPV related genetic tests have emerged. The classical test is the detection of HPV DNA existence. Specifically, there are more than 100 HPV sub-types, of which 16 are oncogenic (High Risk–HR) and two subtypes (16 and 18) are considered responsible for the majority of CxCa cases. Thus, HPV DNA tests detect various subtypes or families of HPV. More recent studies propose the detection of HPV mRNA [1, 2]. The mRNA typing via nucleic acid sequence based amplification (NASBA) technique or Flow Cytometry (FLOW) method has been used in the detection of cancer and precancerous lesions with encouraging results (i.e. increased positive predictive value and therefore reduction of unnecessary colposcopies [1, 2]). Another approach involves the immunocytochemical detection of genetic effects, for instance the overexpression of p16. This method combined with Pap test may increase the diagnostic accuracy [1].

From the published studies [3, 4], it can be concluded that the combination of Pap test with HPV DNA test, results in higher sensitivity compared to each individual method, therefore, these two methods complement each other. However, the specificity of the combination is lower than the performance of each individual method. Thus, HPV tests have not yet a clear role, moreover published papers report significantly different performance for the examined methods. Thus, the application of one method may provide a level of protection in one population, but the same method cannot determine reliably the risk of each individual woman.

On the above basis, we tried to employ techniques based on Artificial Intelligence (AI) and Pattern Recognition (PR) for the optimal, non-linear combination of HPV related tests being available nowadays [5-7]. Since 2010, our team has developed an intelligent clinical decision support system (CDSS), based on Artificial Neural Networks (ANNs), for personalised CxCa diagnosis and prognosis [7]. Moreover, we developed an information technology (IT) system integrating the developed ANNs available as a web-based CDSS for CxCa [8]. This IT system is capable to support many individual users, therefore it can be simultaneously used by researchers, physicians and medical laboratories worldwide.

In this paper we present a new risk assessment model based on a Bayesian Network (BN), which intelligently combines the results of Pap test and the HPV genetic tests in order to support the management of patients with ASCUS or LSIL, therefore can identify women being at real risk to develop CxCa. The BN model is integrated into the web-based CDSS, serving as a risk assessment tool for the improved management and triage of women with abnormal Pap test. This IT system is available at <http://hvp.biomed.ntua.gr>, and is available for mobile devices as well, consisting in this way a remote diagnosis and patient management system like the ones proposed in [9, 10].

## II. MATERIALS AND METHODS

### A. Clinical Data

For the system’s development and testing we used anonymized data collected randomly from women enrolled in a study that was run by the Department of Cytopathology of the Medical School of Athens’ University (“ATTIKON”

University Hospital). The study was approved by the Bioethics Committee and all participant women had signed an Informed Patient Consent (ICON) form allowing anonymous usage of their data. Data included HPV genetic examinations, cytological diagnoses, subsequent histological examination of biopsies, visit dates, patient age, etc. These have been stored into a database designed for the specific study.

From this database, 740 cases with full tests’ series were selected for examination (Table I). Cases with one or more missing or invalid/inadequate tests’ results were excluded from the study.

Each data series included the following tests:

- Pap test formulated according to the Bethesda classification (TBS2001 system).
- HPV DNA test, conducted by the CLART® HUMAN PAPILOMAVIRUS 2 test (GENOMICA) allowing simultaneous detection of 35 different high or low risk HPV genotypes.
- NASBA test (NucliSENS EasyQ® HPV v1.0) for identification of E6/E7 mRNA for HPV types: 16, 18, 31, 33 and 45.
- Flow Cytometry (PermiFlow®, Invirion Diagnostics, LLC, Oak Brook, IL) that allows the identification of E6/E7 mRNA expression of high-risk HPV using Flow Cytometry technique (FLOW), and
- Immunocytochemical expression of p16 using the CINtec® Cytology Kit.

Regarding the HPV DNA test, we considered only the high-risk (HR) HPV types as it is well known that the probability of a low risk subtype to cause cervical lesions is very small.

A cervical biopsy was performed if Pap test revealed ASCUS and above cytological categories (ASCUS+) or there was a visible lesion upon colposcopy. For those cases having a histological outcome, the histological diagnosis was used as the golden standard. Random biopsies were not obtained in clinically negative cases, which are defined as cases with negative cytology and negative colposcopy; in these cases it is not ethically allowed to obtain samples for histological examination. The three-tiered Cervical Intraepithelial Neoplasia (CIN) grading system was used for histological diagnosis.

TABLE I. CASES DISTRIBUTION (CYTOLOGY VERSUS HISTOLOGY)

| Histology           | Pap Test Result |       |      |      |      | Total       |
|---------------------|-----------------|-------|------|------|------|-------------|
|                     | WNL             | ASCUS | LSIL | HSIL | CxCa |             |
| Clinically Negative | 196             | 0     | 0    | 0    | 0    | 196 (26.5%) |
| Negative            | 35              | 60    | 22   | 5    | 0    | 122 (16.5%) |
| CIN1                | 31              | 66    | 142  | 22   | 0    | 261 (35.3%) |
| CIN2/3              | 3               | 13    | 27   | 93   | 0    | 136 (18.4%) |
| CxCa                | 0               | 1     | 2    | 7    | 15   | 25 (3.4%)   |
| <b>Total</b>        | 265             | 140   | 193  | 127  | 15   | 740         |

## B. Bayesian Networks and Model's Design

A Bayesian Network (BN) is considered to be a method for knowledge representation and support for reasoning under uncertainty. The scope of BN is to represent a joint distribution  $P$  over a set of random variables  $X = \{X_1, X_2, \dots, X_n\}$  [11]. A BN can be described as an acyclic directed graph (DAG). The DAG defines a factorization of a joint distribution over the random variables that are represented by its nodes. The factorization is estimated by the direct connections among the nodes. For each BN, a set of nodes (random variables)  $X^G = \{X_1, X_2, \dots, X_n\}$  and a pair  $B = (G, P)$  is defined, where  $G$  is a DAG,  $P$  factorizes over  $G$  and also it is specified as a set of conditional probability distributions (CPDs) associated with  $G$ 's nodes, i.e. CPD is defined as individual factors  $P(X_i | P_{a_{X_i}}^G)$ , where  $P_{a_{X_i}}^G$  denotes the parent of  $X_i$  in  $G$ .

The process of Bayesian network modelling consists of a number of computational algorithms. The structure and the parameter values of a Bayesian network can be provided by expert or by automatic learning methods. The literature has highlighted such methods both for searching the Bayesian structure and estimating parameter values from training data.

### 1) Bayesian learning algorithms

The methods of Bayesian learning can be classified into two categories based on the objective of the learning algorithm: the structure learning and the parameters' learning algorithms.

a) *Structure learning*: There are various algorithms for estimating the structure of a BN based on a training dataset with the most common to be categorized into two approaches: the constrained-based and the search & score based methods. The PC algorithm [12, 13] is a well-known constrained-based method for learning the structure of a BN. The search & score-based algorithms is a different approach to learning structure. It is an optimization-based search approach based on a scoring function and a search strategy. Such a method, which was also implemented in the current paper, is the K2 algorithm [14]. It is a very fast learning algorithm that initially begins with a given ordering of the attributes (nodes). Then, it processes each node by adding edges from the previously processed nodes to the current one. The criterion for adding an edge is the maximization of network's score [14]. When there is no further improvement it continues to the next node. In order to avoid overfitting, the number of parents for each node can be previously restricted to a maximum value. Although the ordering of the method based on the previous processed nodes ensures the construction of an acyclic graph, the result depends on the initial ordering, so it makes sense to run the algorithm several times with different random orderings.

b) *Parameters' estimation*: The second category of BN learning methods is allied to the task of estimating the parametric values of BN. The Bayesian parameter learning consists of estimating the values of parameters  $\theta$  corresponding to DAG structure  $G$  and distributions  $P$  from a complete database of cases  $D = \{D^1, \dots, D^N\}$  composed by  $N$  independent and identically distributed observations from a distribution  $\Pr(D|\theta)$  [11]. Let  $\theta = \{\theta_i\}$ , where  $\theta_i = \{\theta_{jk}\}$  and  $\theta_{jk} = \{\theta_{jk}\}$ , such as  $\theta_{jk} = \Pr(X_i = k | pa(X_i) = j)$  for each  $i$ ,

$j, k$ . The parameters that best represent a dataset is known as  $\theta_{MLE}$ , where MLE denotes the Maximum Likelihood Estimation and it is defined in the equation (1).

$$\theta_{MLE} = \underset{\theta \in \Theta}{\operatorname{argmax}} \Pr(D|\theta) \quad (1)$$

Suppose in the dataset  $D$  over variable  $X$  with  $k$  values, there are  $M_i$  instances and  $\theta_i = \Pr(D = x_i)$ . Then, the multinomial likelihood is:

$$L(\theta|D) = \Pr(D|\theta) = \prod_{i=1}^k \theta_i^{M_i} \quad (2)$$

So by assuming as prior probability a Dirichlet distribution, the  $\theta_{MLE}^{ijk}$  of a BN with discrete distribution that contains  $X_1, \dots, X_N$  random variables with  $X_i: 1, 2, \dots, r_i$ , number of configurations of  $pa$  (parents) of  $X_i: 1, 2, \dots, q_i$ , is defined in the following equation:

$$\theta_{MLE}^{ijk} = \frac{a_{ijk} + m_{ijk}}{\sum_j a_{ijk} + m_{ijk}} \quad (3)$$

where  $a_{ijk}, m_{ijk}$  are parameters from Dirichlet distribution  $\operatorname{Dir}(a_1 + m_1, \dots, a_R + m_R)$  [11].

### 2) Bayesian inference engine

An inference Engine computes the marginal probability distribution of members of a set of variables given evidence. There are inference engines that utilize exact methods and approximate methods. Additionally, the approximate methods can be categorized in either deterministic or stochastic.

In our case, the BNs are discrete and simple so we adopted the Junction Tree (JT) [15] as an inference engine. More specifically JT is a well-known inference method based on message passing algorithm for executing inference on BNs and other graphical models. JT performs an exact inference by building a tree representation of the BN. It creates a tree of cliques, and carries out a message-passing procedure on this tree. The basic idea is to represent probability distribution corresponding to any graph as a product of clique potentials.

### 3) Bayesian network modelling

The objective of BN modelling is to build a BN decision model as close as possible to the diagnosis criteria. Initially, the dataset is divided into training and test sets, where the training set is used for the structure and parameters learning of the BN model, and the test set is used for performance evaluation. In order to extract the optimal BN model, an iterative process of: a) BN structure learning based on K2 algorithm and parameters estimation by the MLE, and b) evaluation of each trained BN using well-defined performance metrics (validation process), has been adopted. The 70% of the training set has been used for the learning procedure and the rest 30% has been used for the validation of each trained BN so as to select the best-performing model. After the extraction of the optimal BN, the test set is used to assess its predictive performance on data which have not been used in any way in the designing process.

### III. RESULTS

In order to implement the previous described BN modelling process, as well as the subsequent classification, we utilized the WEKA suite [16], a well-known machine learning software.

The BN was built in such way so as to be used mainly as a classifier that gives the posterior probability distribution of the classification node given the values of the other attributes. More specifically, we set the ‘‘Histology’’ node as the classification node, thus the BN is used to classify each case into the 4 classes corresponding to the cervical histology (Negative, CIN1, CIN2/3 and CxCa), taking into consideration the results of Pap test, HR HPV, FLOW, p16 and NASBA tests. Thus, the model takes as inputs results from the examinations of each case and outputs its classification group, providing in this way a prediction regarding the actual cervical status of each woman (CIN status).

The available dataset (Table I) was randomly divided into 2 sets: the training set (512 cases) which was used to build the model and perform the required learning and validation as it has been described in the previous section, and the test set (228 cases) which was used for performance evaluation of the optimal BN model. During the random division of the dataset, stratification has been taken into consideration so as the classes’ distribution in each set is approximately the same as that in the initial dataset. For example, from the 228 cases of the test set, the 35% (80 cases) are CIN1, as it is the percentage of CIN1 in the initial dataset (see Table I). Thus, each set contains representative samples of the same larger population.

Several BN models have been built, trained and evaluated according to the previously described iterative learning process. The optimal BN, extracted with the previously described procedure, is illustrated in Fig. 2. The constructed Bayesian model graph describes the direct dependencies among the nodes (random variables). Each edge donates a dependency with a specific direction. For example the result of the NASBA test depend on the observations of the HR HPV, FLOW and p16 test, and alike, the Pap test result has a direct dependency on p16.

Table II presents the confusion matrix obtained by testing the final BN model on the test set.

In order to evaluate the performance of the proposed model compared to the tests involved in this study, we calculated the sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV) and overall accuracy (OA) of the methods on the basis of detecting CIN2+, i.e. high-grade cervical intraepithelial neoplasia and CxCa (Table III). The cut-off of CIN2+ was used in order to have comparable results between the BN and the other medical tests. According to this threshold, the cases with histological diagnosis of CIN1 and below were considered negative and the cases with histological diagnosis of CIN2 and above were considered positive. As shown in Table III, different positivity thresholds have been taken into consideration for the Pap test. For the BN, positivity was defined as a classification result of CIN2/3 or CxCa. In comparison to the medical tests involved in this study, the BN produced the most balanced results in terms of sensitivity, specificity, PPV and NPV, and the best overall accuracy.

Moreover, the BN’s propagation of evidence, which actually consists of updating the probability distributions of the variables according to the newly available evidence, allows us to make valuable observations and conclusions. By observing a set of variables of interest following various scenarios in terms of different given evidences, we are able to test different cases of patients who have undergone various combinations of tests. Some examples are presented in Table IV.

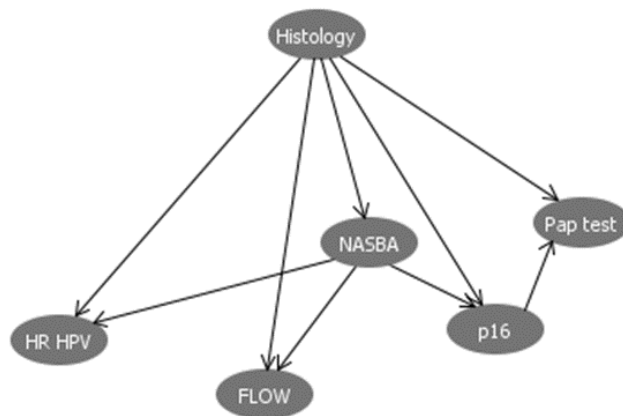


Figure 2. CxCa Bayesian Network

TABLE II. CONFUSION MATRIX OF BN ON THE TEST SET

| Histology                           | Bayesian Network Result |      |        |      |
|-------------------------------------|-------------------------|------|--------|------|
|                                     | Negative                | CIN1 | CIN2/3 | CxCa |
| Negative/<br>Clinically<br>Negative | 89                      | 7    | 1      | 0    |
| CIN1                                | 14                      | 58   | 8      | 0    |
| CIN2/3                              | 1                       | 6    | 36     | 0    |
| CxCa                                | 0                       | 1    | 1      | 6    |

TABLE III. DIAGNOSTIC PERFORMANCE OF PAP TEST, BIOMARKERS AND THE BAYESIAN NETWORK IN DETECTING CIN2+

|                                 | Sensitivity | Specificity | PPV  | NPV  | OA   |
|---------------------------------|-------------|-------------|------|------|------|
| Pap test<br>(cut-off<br>ASCUS+) | 98.0        | 49.2        | 35.7 | 98.8 | 60.0 |
| Pap test<br>(cut-off<br>LSIL+)  | 94.1        | 63.8        | 42.9 | 97.4 | 70.6 |
| Pap test<br>(cut-off<br>HSIL+)  | 74.5        | 92.7        | 74.5 | 92.7 | 88.6 |
| HR-HRV<br>DNA                   | 98.0        | 70.6        | 49.0 | 99.2 | 76.6 |
| NASBA                           | 90.2        | 89.8        | 71.9 | 97.0 | 89.9 |
| FLOW                            | 96.1        | 81.4        | 59.8 | 98.6 | 84.6 |
| p16                             | 70.6        | 92.1        | 72.0 | 91.6 | 87.3 |
| Bayesian<br>Network             | 84.3        | 94.9        | 82.7 | 95.5 | 92.5 |

Performance measures have been calculated on the test set. Histology endpoint is CIN2+ for all cases. CIN2+: Cervical intraepithelial neoplasia grade 2 or worse. HR-HPV: High-risk human papillomavirus. NASBA: Nucleic Acid Sequence Based Amplification for the identification of E6/E7 mRNA of the HPV types: 16, 18, 31, 33 and 45. FLOW: Flow cytometric E6/E7 HPV mRNA assay. PPV: Positive Predictive Value. NPV: Negative Predictive Value. OA: Overall Accuracy.

TABLE IV. MODEL OBSERVATIONS BASED ON DIFFERENT SCENARIOS

| Scenario | Pap Test | HR HPV DNA | NASBA    | Flow Cytometry | p16      | BN Results (Posterior Probabilities %)  |
|----------|----------|------------|----------|----------------|----------|---|
| 1        | ASCUS    | Positive   | Unknown  | Positive       | Positive | Negative=4, CIN1=22, CIN2/3=66, CxCa=8  |
| 2        | ASCUS    | Negative   | Unknown  | Unknown        | Unknown  | Negative=63, CIN1=34, CIN2/3=3, CxCa=1  |
| 3        | ASCUS    | Positive   | Positive | Negative       | Negative | Negative=14, CIN1=64, CIN2/3=16, CxCa=1 |
| 4        | LSIL     | Negative   | Unknown  | Positive       | Unknown  | Negative=7, CIN1=73, CIN2/3=17, CxCa=3  |
| 5        | LSIL     | Positive   | Negative | Positive       | Negative | Negative=1, CIN1=70, CIN2/3=28, CxCa=1  |
| 6        | LSIL     | Positive   | Positive | Positive       | Positive | Negative=2, CIN1=41, CIN2/3=51, CxCa=6  |
| 7        | HSIL     | Negative   | Negative | Negative       | Unknown  | Negative=31, CIN1=49, CIN2/3=16, CxCa=4 |

#### IV. DISCUSSION AND CONCLUSIONS

The application of the proposed risk assessment model gave promising results, suggesting that such an approach may improve the accuracy of diagnosis and support the triage of ASCUS and LSIL. According to the results, the developed BN produced the most balanced results in terms of sensitivity, specificity, PPV and NPV, and demonstrated the highest overall accuracy, compared to cytology and the biomarkers used in the study. The results should be further assessed in larger datasets in order to confirm the reproducibility of these findings. Additionally, future studies should investigate the integration in the BN model of other demographic and medical history data such as age, births or vaccination against HPV among others.

This risk assessment model was built as a web app (developed in the Java language using WEKA [16]) and was integrated into the previous constructed and presented web-based information system [8], serving as a decision support system to physicians and medical researchers for the management of new cases or the follow up of existing cases with abnormal Pap tests.

As depicted in Fig.2, the BN provided various relations of the performed examinations; these are consistent with the HPV natural history. For example the relation of HPV and NASBA pinpoints that firstly there should be an infection (detected by HPV) and subsequently the virus integration (detected by NASBA). Additionally the progression to CIN2+ (detected by Histology) is not feasible without the virus integration (NASBA). The models' capability to estimate the risk of a woman to harbour CIN or CxCa (see Table IV) based on the combinations of the various tests' outcomes is a second important aspect of the BNs, mainly for two reasons: 1) it is possible to conclude on cases which have unknown examination results, as it may happen in the everyday lab routine for various reasons, and 2) it is possible to manage women with ambiguous results, especially for women having an ASCUS test Pap.

Modern mobile terminals along with the advancements in mobile communications allow physicians to execute the proposed risk assessment model at anyplace/anytime (via the web page <http://hvp.biomed.ntua.gr>), even in places that there is no fixed networking infrastructure available, allowing instant case management and decision making.

#### REFERENCES

- [1] C. Kottaridi, S. Tsiodras, A. Spathis, *et al.*, "Clinical performance of human papillomavirus E6, E7 mRNA flow cytometric assay compared to human papillomavirus DNA typing," *Anal Quant Cytol Histol*, vol. 33, pp. 305-10, Dec 2011.
- [2] A. Trope, K. Sjoborg, A. Eskild, *et al.*, "Performance of human papillomavirus DNA and mRNA testing strategies for women with and without cervical neoplasia," *J Clin Microbiol*, vol. 47, pp. 2458-64, Aug 2009.
- [3] J. Cuzick, M. Arbyn, R. Sankaranarayanan, *et al.*, "Overview of human papillomavirus-based and other novel options for cervical cancer screening in developed and developing countries," *Vaccine*, vol. 26 Suppl 10, pp. K29-41, Aug 19 2008.
- [4] P. Naucler, W. Ryd, S. Tornberg, *et al.*, "Efficacy of HPV DNA testing with cytology triage and/or repeat HPV DNA testing in primary cervical cancer screening," *J Natl Cancer Inst*, vol. 101, pp. 88-99, Jan 21 2009.
- [5] P. Karakitsos, A. Pouliakis, C. Meristoudis, *et al.*, "A preliminary study of the potential of tree classifiers in triage of high-grade squamous intraepithelial lesions," *Anal Quant Cytol Histol*, vol. 33, pp. 132-40, Jun 2011.
- [6] P. Karakitsos, C. Chrelias, A. Pouliakis, *et al.*, "Identification of women for referral to colposcopy by neural networks: a preliminary study based on LBC and molecular biomarkers," *J Biomed Biotechnol*, vol. 2012, p. 303192, 2012.
- [7] P. Bountris, M. Haritou, A. Pouliakis, *et al.*, "An intelligent clinical decision support system for patient-specific predictions to improve cervical intraepithelial neoplasia detection," *Biomed Res Int*, vol. 2014, p. 341483, 2014.
- [8] P. Bountris, G. Kotronoulas, T. Tagaris, *et al.*, "CxCaDSS: A web-based clinical decision support system for cervical cancer," in *Proc. 6<sup>th</sup> Euro Conf of IFMBE, MBEC 2014, Dubrovnik, Croatia*, Vol. 45, pp. 757-760.
- [9] D. Iliopoulou, K. Giokas, S. Mougiakakou, *et al.*, "A Telematic System for Diabetes Management, Reporting and Patient Advice," *The Journal on Information Technology in Healthcare*, vol. 3, pp. 307-313, 2005.
- [10] S. Mougiakakou, J. Stoitsis, D. Iliopoulou, *et al.*, "A Communication Platform for Tele-monitoring and Tele-management of Type 1 Diabetes," *Conf Proc IEEE Eng Med Biol Soc*, vol. 3, pp. 2207-10, 2005.
- [11] D. Koller and N. Friedman, *Probabilistic graphical models : principles and techniques*. Adaptive computation and machine learning. Cambridge, MA: MIT Press, 2009.
- [12] P. Spirtes and C. Glymour, "An Algorithm for Fast Recovery of Sparse Causal Graphs," *Social Science Computer Review*, vol. 9, pp. 62-72, April 1, 1991 1991.
- [13] P. Spirtes, C. N. Glymour, and R. Scheines, *Causation, prediction, and search*. Adaptive computation and machine learning, 2nd ed. Cambridge, Mass.: MIT Press, 2000.
- [14] G. Cooper and E. Herskovits, "A Bayesian method for the induction of probabilistic networks from data," *Machine Learning*, vol. 9, pp. 309-347, 1992/10/01 1992.
- [15] T. D. Nielsen and F. V. Jensen, *Bayesian networks and decision graphs*: Springer, 2009.
- [16] (2014). *Weka 3: Data Mining Software in Java*. Available: <http://www.cs.waikato.ac.nz/ml/weka/>