Using Classification and Regression Trees, Liquid-Based Cytology and Nuclear Morphometry for the Discrimination of Endometrial Lesions

Abraham Pouliakis, Ph.D.,[1]* Charalampia Margari, M.D.,[2]
Niki Margari, M.D., Ph.D.,[1] Charalampos Chrelias, M.D. ASS. PROF.,[3]
Dimitrios Zygouris, M.D.,[3] Christos Meristoudis, M.D.,[1]
Ioannis Panayiotides, M.D. ASS. PROF.,[4] and Petros Karakitsos, M.D., PROF.[1]

'The objective of this study is to investigate the potential of classification and regression trees (CARTs) in discriminating benign from malignant endometrial nuclei and lesions. The study was performed on 222 histologically confirmed liquid based cytological smears, specifically: 117 benign cases, 62 malignant cases and 43 hyperplasias with or without atypia. About 100 nuclei were measured from each case using an image analysis system; in total, we collected 22783 nuclei. The nuclei from 50% of the cases (the training set) were used to construct a CART model that was used for knowledge extraction. The nuclei from the remaining 50% of cases (test set) were used to evaluate the stability and performance of the CART on unknown data. Based on the results of the CART for nuclei classification, we propose two classification methods to discriminate benign from malignant cases. The CART model had an overall accuracy for the classification of endometrial nuclei equal to 85%, specificity 90.68%, and sensitivity 72.05%. Both methods for case classification had similar performance: overall accuracy in the range 94–95%, specificity 95%, and sensitivity 91–94%. The results of the proposed system outperform the standard cytological diagnosis of endometrial lesions. This study highlights interesting diagnostic features of endometrial nuclear morphology and provides a new classification approach for endometrial nuclei and cases. The proposed method can be a useful tool for the everyday practice of the cytological laboratory. Diagn. Cytopathol. 2014;42:582–591.

Key Words: liquid-based cytology; classification and regression trees; artificial intelligence; endometrial cytology; image processing

Introduction

The endometrium is histologically a complex tissue system; the morphological diagnosis of endometrial lesions is therefore fraught with difficulties in histological as well as in cytological material. In addition, endometrium morphology is affected by many parameters: sex steroids, growth factors, oncogene products and various peptides[1]; therefore,diagnostic procedures may be very susceptible to artifacts due to delayed fixation or inappropriate tissue sampling. Due to the fact that the natural history of endometrial hyperplasia has not yet been elucidated, many classification systems have been elaborated.[3] As a consequence, there is no standardized therapy ranging from hormonal therapy to hysterectomy.

During the last decades, various classification techniques have been used in medicine and especially in diagnostic cytology, involving either classical statistical models or

[1]Department of Cytopathology, University of Athens, "ATTIKON" University Hospital, Athens, Greece
[2]Oncology Department, "Evaggelismos" Hospital, Athens, Greece
[3]3rd Department of Obstetrics and Gynecology, University of Athens, "ATTIKON" University Hospital, Athens, Greece
[4]2nd Department of Pathology, University of Athens, "ATTIKON" University Hospital, Athens, Greece
Contract grant sponsor: Operational Program for Competitiveness, Program for the support of young researchers PENED; Contract grant number: 03EΔ 348.
*Correspondence to: A. Pouliakis, e-mail: apou1967@gmail.com
Received 27 December 2012; Accepted 29 October 2013
DOI: 10.1002/dc.23077
Published online 22 November 2013 in Wiley Online Library (wileyonlinelibrary.com).

more advanced techniques, such as neural networks.[4–9] More specially, concerning cytology, these techniques have been applied to various organs, among others stomach,[10–12] breast,[13–16] urinary system,[17–19] cervix,[20–24] and thyroid.[25–28] Classification and Regression Trees (CARTs) are very popular techniques, and have been often used in the past.[22,29–31] CARTs are attractive, being relatively straightforward to construct and they can be interpreted as a set of rules; which are easy to be understood and thus allow smooth integration with other medical decision support systems. Additionally, CARTs have a relatively fast learning curve, thus facilitating their adoption in the environment of a medical laboratory. Finally, in every step of their rules, CARTs include the probabilities for the "item" under classification/investigation to belong to the individual categories.

Concerning classification of endometrial lesions, and in order to overcome the problem of an accurate, reproducible and objective diagnosis, many morphological attempts have been carried out, with relatively promising results.[32–38] Computer aided diagnosis of endometrial lesions, is not new.[39,40] Most attempts are related to the discrimination of the endometrial hyperplasias, as their classification has numerous diagnostic difficulties. However, CARTs have not been applied on nuclear morphometric features, moreover, there are no studies related to LBC samples.

In our study, we aim to investigate the potential role of CARTs and image morphometry applied on LBC samples. The ultimate goals were the identification of the most useful morphometric features for the classification of individual nuclei and the assessment of the performance of the method for both nuclei and subsequently classification of endometrial cytological cases. Studied cases covered the complete range of endometrial lesion types, from normal endometrium to hyperplasias with or without atypia, as well as carcinoma.

Materials and Methods

Sample Collection, Preparation, and Diagnosis

The study involved 222 women and was carried out on smears taken by direct sampling of the endometrial cavity using the EndoGyn® Sampler (Biogyn S.n.c., Mirandola, Italy). Material was collected by rotating the device in the endometrial cavity many times; the EndoGyn sampler was then withdrawn and immersed into a vial containing 30–50 mL of a hemolytic, mucolytic, and proteinolytic purposes solution (e.g., CytoLyt®, Cytyc Corporation). Via the rotation of EndoGyn into CytoLyt, the collected material in the arms of the device was easily removed; additionally CytoLyt can preserve the cells under optimal conditions for up to 7 days, either at room temperature or at 4°C. Moreover, the use of CytoLyt ensures the absence of artifacts due to sample transportation and air-drying before the fixation of cells. After sample collection, the vial was transferred in the laboratory, the liquid was centrifugated at 2500 rounds per minute for 5 min, the obtained pellet containing the endometrial cells was transferred to another vial with fixative solution (PreservCyt®, Cytyc Corporation). PreservCyt mildly fixes the cells within 10–15 min, and may preserve the material for at least 3 weeks at room temperature. Additionally, PreservCyt provides a safe method for transportation, since it contains antiinfective agents. The vial was then inserted into the automated slide processor ThinPrep® TP2000 (Cytyc Corporation). TP2000 prepares a single layer endometrial smear in 90 seconds; the slide area is measuring 2 cm in greatest dimension. One slide was prepared from each case and stained with the Papanikolaou stain using an automated staining machine (Varistain® 24-3 Thermo Electron Corporation [formerly Shandon], Runcorn, UK); slides were used both for cytological diagnosis and image morphometry.

Cytological diagnosis was made by three cytopathologists, with more than 10 years of experience in endometrial cytology, according to the criteria outlined by Fox.[41] The cytological diagnosis was ultimately confirmed by histological examination of endometrial curettage and/or hysterectomy specimens, according to the International Society of Gynaecological Pathology Classification.[42] From all available cases, of our files, we selected 222 cases, in which both diagnoses (cytological and histological) were in agreement; so as to ensure that the rest of the experiment was based on a solid background and nuclear morphometry was based on nuclei with known and confirmed nature. Correlation of cytological and histological diagnoses for the cases of our study is shown in Table I.

Image Analysis System, Measurements, and Their Biological Meaning

The measuring procedure involves only endometrial nuclei and has two steps: (a) selection of nuclei and (b) measurement of nuclei. The average time to select 100 nuclei from each specimen is dependent from the image quality and the ability to use semiautomatic or manual selection of each individual nucleus. Specifically for manual nuclei selection, the user draws the nucleus boundary by the use of the mouse, the average time for 100 nuclei is 38 min. For semiautomatic nuclei selection the user simply clicks on the nucleus, subsequently specialized software algorithms take care and automatically create the nucleus boundary; in this case, the average time for 100 nuclei semiautomatically selected is 6 min. The computer time to extract the measurements from a single nucleus is less than a second, since it is completely automated and

Table I. Correlation of Histological with Cytological Findings

	Histological category									
Cytological group	*Adeno Ca*	*Atrophy*	*Endometrioid Ca*	*Endometrioid Ca grade I*	*Hyperplasia with atypia*	*Hyperplasia without atypia*	*Proliferative phase*	*Secretory phase*	*Serous Ca*	*Total*
Benign		26					56	35		127
Non atypical hyperplastic cells						37				37
Atypical hyperplastic cells					6					6
Malignant	2		48	8					4	62
Total	2	26	48	8	6	37	56	35	4	222

operates unattended and massively on all the nuclei selected from each case.

From available cytological slides, we measured 22,783 nuclei from epithelial cells (about 100 nuclei/case). The nuclei were either isolated or from cell groups that fulfill the cytological criteria of the morphological diagnosis. The measuring process created a dataset of the 22,783 feature vectors, each one was related to a single nucleus, thus each feature vector is representative of one nucleus and each vector component represent a single measured nuclear feature.

The basis of the image analysis system was a computer with Pentium IV processor, the computer was equipped with a frame grabber and a digital camera, model SONY DFW-X700 (Sony Corporation, Tokyo, Japan). The camera was attached on top of the microscope (Leica Microsystems, Wetzlar, Germany) via a X-mount adapter. The images were captured with a 40× objective and subsequently digitized to 1024 × 768 pixels using 24 bits depth (8 bits for each individual color: red green and blue).

Two image analysis software packages were used. Specifically: PathSight version 4.3 (Medical Solutions PLC, UK) was used to capture images from the microscope, and subsequently the images were analyzed with Image-Pro Plus VERSION 4.5 (Media Cybernetics, Bethesda, MD, USA). Image-Pro Plus was used for: background correction, isolation of the nuclei and automated extraction of the measurements corresponding to the measured nuclear features. The Background correction is an important process as removes noise that are caused by the optics of the microscope, for instance, remaining dust even after careful clearing of the objectives. Additionally before the measurement a light calibration procedure was applied, thus all images were captured under similar lighting conditions.

From each nucleus, we extracted two types of features: geometric and densitometric.[43–47] The first group includes nucleus area, major and minor axis, aspect ratio, minimum, maximum and average caliper, minimum, maximum and average nuclear radius, radius ratio, perimeter, roundness and fractal dimension. The group of densitometric features includes integrated optical density, minimum, maximum, mean value and standard deviation of the optical density, separate color measures expressed as mean value of the red, green and blue color components and finally textural features expressed via the margination and heterogeneity. The geometric features are related to characteristics that are produced from the nucleus boundary, in contrast, densitometric features are based on the pixel content of each nucleus; these characteristics describe the nucleus color and texture and reflect DNA characteristics. Specifically, geometric characteristics are related to the nucleus shape, which are deformated due to the neoplastic changes in nucleoskeleton and cytoskeleton. Therefore, geometric characteristics represent a method to measure nuclear alterations from carcinogenesis side effects. Densitometric characteristics are related to DNA, specifically: (a) the DNA quantity as measured via nuclear color characteristics and optical density and (b) DNA status as this is reflected from the distribution of DNA, the former is reflected by the two texture measurements (margination and heterogeneity) and, to some degree, from the standard deviation of the optical density.

Nuclei Classification, Methodology, and Tools

The aim of the CART is to classify individual nuclei in one of two groups, based on histological diagnosis: either "Benign" (histological categories: prolipherative and secretory phase, atrophy, and hyperplasias without atypia) or "Malignant" (histological categories: adenocarcinoma, endometrioid Ca, endometrioid Ca-grade I, serous Ca and hyperplasia with atypia).

We selected CARTs[48] because they are an explanatory technique, able to reveal data structure, identify important characteristics, and develop rules. CARTs can be used quickly and repeatedly. A classification tree is constructed from many repeated splits related to the target variable, these splits are producing rules of the form: If X is less than value A and Y is within a range $[C,D]$, then the sample is classified as benign with probability P. These splits end when no further split can be done, either because all

Table II. Distribution of Cases and Nuclei into the Training and Test Sets

	Number of data in the training and test set		
	Training set	*Test set*	*Total*
Histological category		*Cases*	
Adeno Ca	1	1	2
Atrophy	13	13	26
Endometrioid Ca	24	24	48
Endometrioid Ca?grade I	4	4	8
Hyperplasia with atypia	3	3	6
Hyperplasia without atypia	18	19	37
Proliferative phase	28	28	56
Secretory phase	18	17	35
Serous Ca	2	2	4
Total	**111**	**111**	**222**
		Nuclei	
Adeno Ca	78	90	168
Atrophy	1285	1269	2554
Endometrioid Ca	2527	2482	5009
Endometrioid Ca?grade I	394	384	778
Hyperplasia with atypia	326	333	659
Hyperplasia without atypia	1800	1994	3794
Proliferative phase	2945	2849	5794
Secretory phase	1953	1707	3660
Serous Ca	179	188	367
Total	**11487**	**11296**	**22783**

of their observations belong to the same group, or because the number of observations at the same node is small (according to a predefined value). These represent the terminal nodes of the tree.

No preprocessing of data was performed before training the classification tree, i.e. all measurements were used exactly as extracted by the image measurement system, because preprocessing of data is not a prerequisite to construct CARTs. Fifty percent of the available cases were randomly selected in order to train the CART model. These cases were selected in respect to the histological classes, thus each histological category was equally represented in this dataset. The related feature vectors of these cases (i.e., 11,541 vectors) formed the training set. The remaining 11,296 vectors, from the unused cases, were used to evaluate the performance of the trained classifier (test set). The distribution of the cases and the measured nuclei into the training and test sets are presented in Table II. As mentioned the CART model was created by using the data of the training set, during this process there were determined the important variables and the decision rules, in the sequel the model was evaluated by applying the data of the test set using the rules obtained during training.

Case Classification Techniques

For the classification of individual cases, a relatively simple method, based on the nuclei classification results from the CART was used. Actually, two variants of the same method were evaluated. For the first variant, two features were used: the number of nuclei classified as benign and the number of nuclei classified as malignant (this method is called the numeric classifier). For the second variant, the percentages of nuclei classified as benign and as malignant were used (this is subsequently called the percentages classifier); for both variants, the numbers and the percentages are related to the nuclei of the specific case under investigation.

Table III. Statistics of the Tree Classifier for the Training and Test Set and for the Complete Dataset

Statistical measure	*Training set (%)*	*Test set (%)*	*Complete dataset (%)*
Sensitivity	73.49	70.59	72.05
Specificity	90.30	91.07	90.68
PPV	76.84	77.64	77.23
NPV	88.60	87.58	88.09
FPR	9.70	8.93	9.32
FNR	26.51	29.41	27.95
OA	85.17	84.82	85.00

Statistics Tools and Techniques

The CART system was created using IBM SPSS Statistics 19 for Windows (SPSS, Chicago, IL). To assess the performance, various statistical measures were extracted: specificity, sensitivity, positive and negative predictive value (PPV and NPV), false positive and false negative rates (FPR and NPR), and overall accuracy (OA). The results of the classifiers are provided for the nucleus classification level and for the cases classification level. The stability related to the performance of the classifiers for the training against the test set was evaluated by x^2 comparison of proportions and additionally using Kappa statistics, in order to assess the inter-rater agreement between the numeric and percentages classifiers and for each one of these classifiers against the histological category; used as the golden standard.

Results

Statistical measures of the CART model results appear in Table III. For CART construction, the CRT[31] growing method was used, additionally CART was restricted to have more or equal to 200 nuclei for the parent nodes and at least 20 nuclei for the child nodes; the model was allowed to have up to eight levels of depth, and finally, it was pruned to avoid overfitting. The constructed CART model appears in Figures 1–4. In each box, the number of nuclei belonging in each category (Benign or Malignant) and the related percentages, representing the probability of the nucleus to be benign or malignant (P_B and P_M, respectively) are shown.

It is relatively easy for a cytopathologist to understand and use the CART: for example, if a nucleus has an area

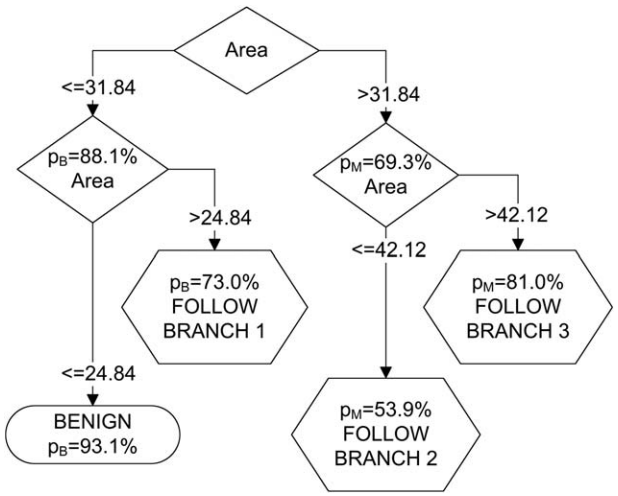

Fig. 1. Main branch of the CART. The probabilities are for the combined dataset (training and test set). P_B, probability for a nucleus to be benign; P_M, probability for a nucleus to be malignant.

less than or equal to 31.84 μm^2, then (Fig. 1) $P_B = 88.1\%$. If this probability is satisfactory, then the cytopathologist may stop the classification procedure at this stage; otherwise, the next step is to examine again the nucleus area (see Fig. 1): if it is smaller than 24.84 μm^2, then the probability of this nucleus being benign accounts to $P_B = 93.1\%$.

CART enabled the correct classification of 85.17% of the nuclei composing the training set and 84.82% of the nuclei composing the test set. The overall accuracy on the complete dataset was 85.00%. The test for proportions between the overall accuracy for the training and test set indicates that the classifier has similar behavior for both sets, because there is no statistically significant difference between training and test set ($x^2 = 0.520$ $P = 0.4708 > 0.05$). These results indicate that the performance of the CART is similar and stable for both training and the test sets.

Despite all available features have been provided to the CART during training, only part of these was statistically important and was eventually involved in the tree classification rules. The most important characteristic was the nuclear area, since it is the discriminating characteristic in the first and second level of the CART. Other important features are the nuclear mean optical density of the color components and nuclear textural characteristics, more specifically: heterogeneity and margination. It is worth noting that apart from nuclear area; no other geometric characteristic was included in the tree structure during the construction process. This finding is in accordance to the diagnostic process followed by cytopathologists, as the most important characteristic that is examined during a cytological examination is nuclei area as well.

Subsequently, we examined the potential of nucleus area as a single diagnostic characteristic. For this purpose,

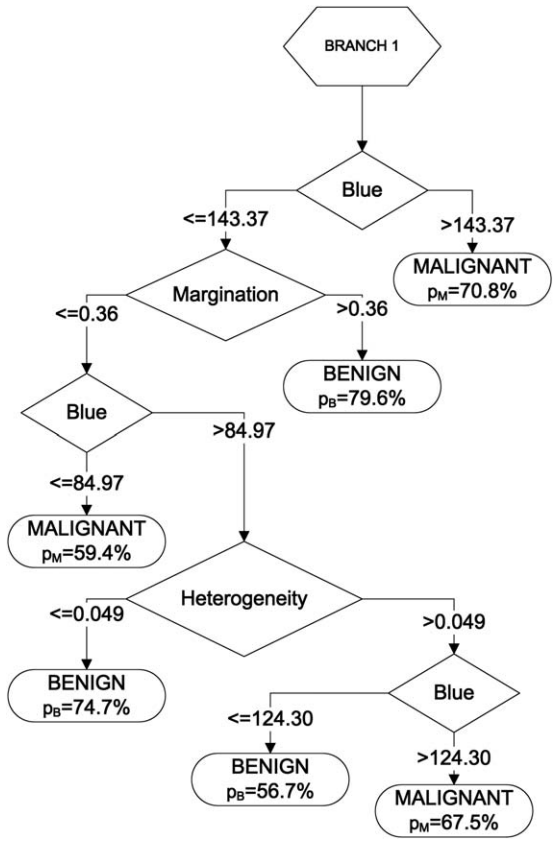

Fig. 2. BRANCH 1 of the CART. This branch is mainly related to the benign nuclei.

we constructed a diagram with the percentages of benign and malignant nuclei for various nuclear sizes (Fig. 5). From this diagram, we concluded that as the nuclear area increases, the probability of a nucleus to be malignant increases as well, in a nonlinear way; similarly, the probability of a nucleus to be benign increases as the nuclear area becomes smaller. The area under curve (AUC) of the ROC diagram for nuclear area was 0.876 with a very small standard error (SE = 0.002, 95% confidence interval (95% CI): lower bound = 0.872, upper bound = 0.881), a fact that confirms the CART finding; that area alone is the most important classification factor. However, there is still an overlap of the benign and malignant categories as indicated in Figure 5.

The performance metrics of the two cases classification methods are depicted in Table IV. The percentages classifier assigned correctly 211 of the 222 cases and missed seven benign cases (three in the training set and four in the test set) and four malignant cases (two in the training set and two in the test set). The numeric classifier classified correctly 209 cases and missed seven benign cases (two in the training set and five in the test set) and six malignant cases (three in the training set and three in the test set). For the case classification method based on the

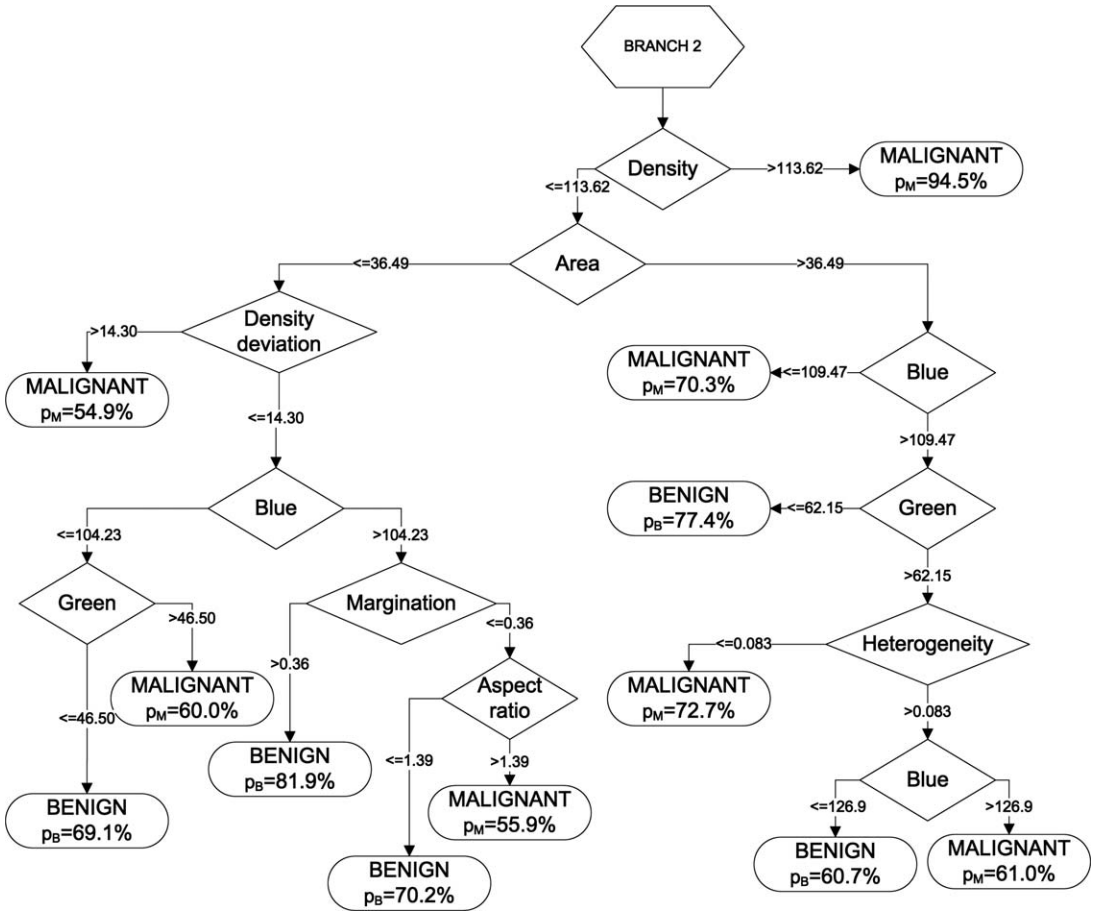

Fig. 3. BRANCH 2 of the CART. This branch is mainly related to the gray zone between benign and malignant nuclei.

number of benign nuclei, since these are classified by the CART (the numeric classifier), the threshold was determined from the training set cases: specifically, the threshold started from 0 and increased up to 150 using a step of 1; for each step, the percentage of the cases that were correctly classified was determined. As optimum threshold was selected the value that gave the best overall accuracy for the training set, in our case this value was 54. Therefore, a proposed algorithm could be: if the case has more than 54 nuclei classified as benign by the CART then the case can be considered as benign otherwise as malignant. The overall accuracy of the numeric classifier was 95.50% on the training set and 92.79% on the test set, the comparison of the overall accuracy for the two sets proves that the "behavior" of this technique is similar on both sets and therefore stable, as there is no statistical significant difference in the two percentages ($x^2 = 0.330$, $P = 0.5659 > 0.05$).

In a similar manner, for the percentages classifier, it was applied a similar approach for threshold determination. The optimum threshold for this method was determined to be 63.5%, thus the rule for this method is: if a case has more than 63.5% of nuclei classified as benign by the CART, then this case may be considered as benign otherwise as malignant. The overall accuracy of this method on the training set was 95.50% and on the test set 94.59%, for this classifier the "behavior" was stable as well, because there is no statistical difference between the two percentages ($x^2 = 9.76 \times 10^{-6}$, $P = 0.9975 > 0.05$).

In order to evaluate if the two case classification methods had differences, we performed the test for proportions; more specifically on the overall accuracy for the training set, test set, and both sets. The results indicate that both case classification methods had not statistically significant performance for all types of datasets: training set, test set, and combined sets (results of test for proportions, respectively: $x^2 = 0.105$, $P = 0.7461 > 0.05$, $x^2 = 0.0759$, $P = 0.7829 > 0.05$, $x^2 = 0.0458$, $P = 0.8305 > 0.05$, respectively). Additionally the Kappa statistics indicate that there was excellent concordance between the two methods ($\kappa = 91.7\%$, SE = 0.0289, 95% CI = 0.86–0.973) as well as between numeric classifier and the histological category ($\kappa = 86.3\%$, SE = 0.0369, 95% CI = 0.79–0.935) and the percentages classifier against histology ($\kappa = 88.5\%$, SE = 0.0338, 95% CI = 0.819–0.951).

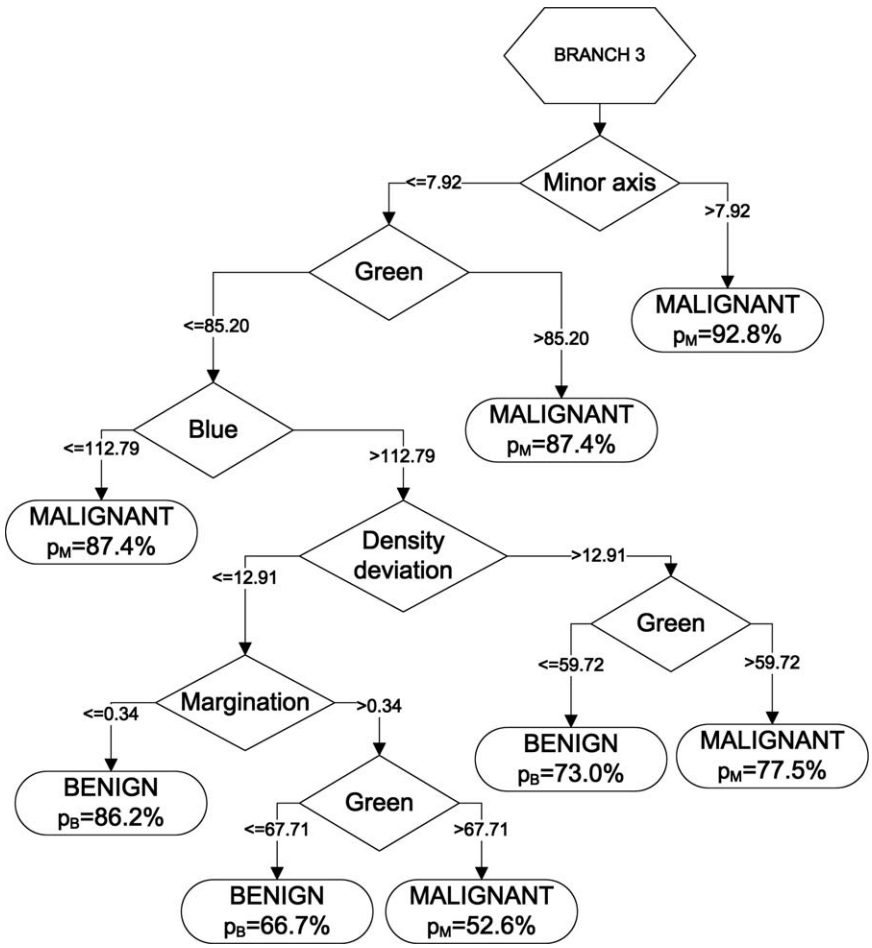

Fig. 4. BRANCH 3 of the CART. This branch is mainly related to the malignant nuclei.

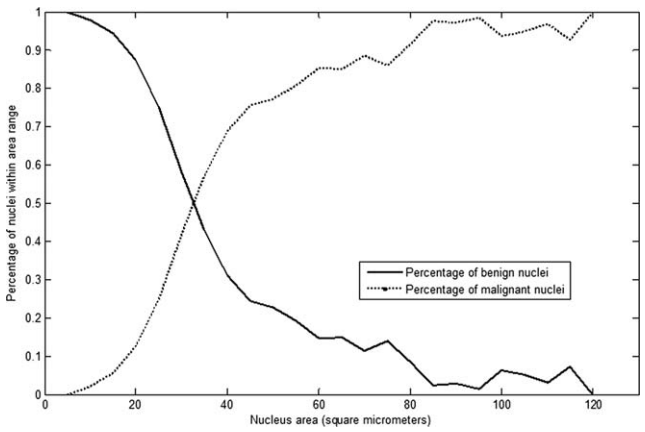

Fig. 5. Percentages of benign and malignant nuclei in relation to the nucleus area (aka probability for a nucleus to be benign or malignant in relation to nuclear size).

Discussion

Lack of objective and reproducible diagnostic criteria[49–51] for endometrial hyperplasias is the fundamental reason for the ambiguous knowledge of their biological behavior, the increased number of classification systems and the inconsistent management of endometrial lesions.[3] Consequently, endometrial cytology is not widely accepted as a valid method for the investigation of endometrial lesions. Criticism against endometrial cytology basically stems from the difficulty to evaluate the classical cytomorphological criteria and from the need to integrate cytoarchitectural characteristics into the diagnostic methodology.[52]

However, in previous studies, morphometric, textural and stereological data extracted by quantitative cytology and histology techniques and the subsequent classification by monoparametric or multivariate statistical models, showed that both the discrimination among complex, complex atypical hyperplasias and carcinomas[40] and the prediction of the patient clinical course are feasible to some degree. Additionally, during the last decade, an increased number of publications[52–66] pointed that cytology could be used as an alternative or ancillary method to histological examination of Dilatation and Curettage (D&C) specimens.

Table IV. Statistical Characteristics for the Two Classification Methods

	Numeric classifier			*Percentage classifier*		
	Training set (%)	*Test set (%)*	*Combined (%)*	*Training set (%)*	*Test set (%)*	*Combined (%)*
Sensitivity	91.18	91.18	91.18	94.12	94.12	94.12
Specificity	97.40	93.51	95.45	96.10	94.81	95.45
PPV	93.94	86.11	89.86	91.43	88.89	90.14
NPV	96.15	96.00	96.08	97.37	97.33	97.35
FPR	2.60	6.49	4.55	3.90	5.19	4.55
FNR	8.82	8.82	8.82	5.88	5.88	5.88
OA	95.50	92.79	94.14	95.50	94.59	95.05

In this study, we tried to reduce the subjectivity introduced by humans and to introduce a more objective methodology. Specifically, we used objective and reproducible nuclear measurements, thus each nucleus is represented by a set of numbers. Additionally each nucleus was subsequently evaluated using an objective methodology, CARTs. The clinical application aims to the discrimination of endometrial lesions requiring immediate referral to D&C from those that could respond to pharmaceutical treatment or just monitored through follow up. According to our results, the proposed methodology gave encouraging results in the discrimination between the nuclei from cases of hyperplasia with cytological atypia and carcinoma, from those of benign endometrial cases and cases without cytological atypia. The two proposed case classification algorithms had excellent discrimination capabilities, in comparison to the standard cytological examination, as it is nowadays performed (i.e., without use of measurements).

According to our results and experience[53,54] the nuclear morphological characteristics may have an important role in endometrial cytology, regardless of the technique (conventional or LBC) used. However, LBC contributes toward objectivity requirements. The advantages of LBC are related to common diagnostic criteria, due to the standardized procedure related to fixation, transfer and cytological preparation of the material,[67,68] can be easily deployed in a cytopathology laboratory,[68] and enables application of ancillary techniques.[68,69] Via LBC, the cellular overlap is related only to the real three-dimensional structures and not due to the slide preparation method[70]; there is improved mucolysis and reduction of blood and artifacts and there is reported better diagnostic accuracy than conventional slides.[71] The most important characteristic is the nuclear area, a feature that is the primary evidence examined by cytopathologists during screening. Subsequently, another important factor is the chromatin distribution, which is reflected by the appearance of heterogeneity, margination, and the standard deviation of the nuclear optical density into the individual tree branches (Figs. 1–4). Eventually, the CART model identified a set of rules that are based on the same characteristics that are examined by cytopathologists during endometrium samples examination.

In conclusion, the proposed methodology introduces an objective method for endometrium nuclei and cases classification; thus, subjectivity due to the interpretative factor could be reduced. The criteria identified as important, are similar to the criteria used by skilled cytopathologists. The application of LBC allows additional objectification of the process. Specifically (a) the biological material is immediately immersed in the vial, thus there is no oxidization, therefore, reduction of artifacts (b) the background is clear due to lysis of mucus and erythrocytes, and (c) there is lack of artificial cellular and/or nuclear overlapping. Therefore, LBC is a technology enabling the application of CARTs. The proposed methodology could be of use in the everyday practice of the cytopathology laboratory as an adjunctive diagnostic test.

Acknowledgment

The authors thank the reviewers for their comments that helped to improve the manuscript.

References

1. Shyamala G, Ferenczy A. 1981. The effect of sodium molybdate on the cytoplasmic estrogen and progesterone receptors in human endometrial tissues. Diagn Gynecol Obstet 3:277–282.
2. Murphy LJ, Murphy LC, Friesen HG. 1987. Estrogen induces insulin-like growth factor-I expression in the rat uterus. Mol Endocrinol 1:445–450.
3. Ferenczy A, Bergeron C. Endometrial hyperplasia. In: Lowe D, Fox H, editors. Advances in gynaecological pathology. Edinburgh: Churchill Livingstone; 1992. p 207–234.
4. Karakitsos P, Cochand-Priollet B, Guillausseau PJ, Pouliakis A. 1996. Potential of the back propagation neural network in the morphologic examination of thyroid lesions. Anal Quant Cytol Histol 18:494–500.
5. Marchevsky AM, Tsou JA, Laird-Offringa IA. 2004. Classification of individual lung cancer cell lines based on DNA methylation markers: Use of linear discriminant analysis and artificial neural networks. J Mol Diagn 6:28–36.
6. Markopoulos C, Karakitsos P, Botsoli-Stergiou E, Pouliakis A, Ioakim-Liossi A, Kyrkou K, Gogas J. 1997. Application of the learning vector quantizer to the classification of breast lesions. Anal Quant Cytol Histol 19:453–460.
7. Pantazopoulos D, Karakitsos P, Iokim-Liossi A, Pouliakis A, Botsoli-Stergiou E, Dimopoulos C. 1998. Back propagation neural

network in the discrimination of benign from malignant lower urinary tract lesions. J Urol 159:1619–1623.

8. Astion ML, Wilding P. 1992. Application of neural networks to the interpretation of laboratory data in cancer diagnosis. Clin Chem 38: 34–38.
9. Cochand-Priollet B, Koutroumbas K, Megalopoulou TM, Pouliakis A, Sivolapenko G, Karakitsos P. 2006. Discriminating benign from malignant thyroid lesions using artificial intelligence and statistical selection of morphometric features. Oncol Rep 15:1023–1026.
10. Chien CW, Lee YC, Ma T, Lee TS, Lin YC, Wang W, Lee WJ. 2008. The application of artificial neural networks and decision tree model in predicting post-operative complication for gastric cancer patients. Hepatogastroenterology 55:1140–1145.
11. Yamamura Y, Nakajima T, Ohta K, Nashimoto A, Arai K, Hiratsuka M, Sasako M, Kodera Y, Goto M. 2002. Determining prognostic factors for gastric cancer using the regression tree method. Gastric Cancer 5:201–207.
12. Karakitsos P, Megalopoulou TM, Pouliakis A, Tzivras M, Archimandritis A, Kyroudes A. 2004. Application of discriminant analysis and quantitative cytologic examination to gastric lesions. Anal Quant Cytol Histol 26:314–322.
13. Ljung BM, Chew KL, Moore DH, 2nd, King EB. 2004. Cytology of ductal lavage fluid of the breast. Diagn Cytopathol 30:143–150.
14. Wolberg WH, Tanner MA, Loh WY, Vanichsetakul N. 1987. Statistical approach to fine needle aspiration diagnosis of breast masses. Acta Cytol 31:737–741.
15. Markopoulos C, Karakitsos P, Botsoli-Stergiou E, Pouliakis A, Gogas J, Ioakim-Liossi A, Kyrkou K. 1997. Application of back propagation to the diagnosis of breast lesions by fine needle aspiration. Breast 6:293–298.
16. Dey P, Logasundaram R, Joshi K. 2013. Artificial neural network in diagnosis of lobular carcinoma of breast in fine-needle aspiration cytology. Diagn Cytopathol 41:102–106.
17. Schaffer AA, Simon R, Desper R, Richter J, Sauter G. 2001. Tree models for dependent copy number changes in bladder cancer. Int J Oncol 18:349–354.
18. Karakitsos P, Pouliakis A, Kordalis G, Georgoulakis J, Kittas C, Kyroudes A. 2005. Potential of radial basis function neural networks in discriminating benign from malignant lesions of the lower urinary tract. Anal Quant Cytol Histol 27:35–42.
19. Vriesema JL, van der Poel HG, Debruyne FM, Schalken JA, Kok LP, Boon ME. 2000. Neural network-based digitized cell image diagnosis of bladder wash cytology. Diagn Cytopathol 23:171–179.
20. Giovagnoli MR, Cenci M, Olla SV, Vecchione A. 2002. Cervical false negative cases detected by neural network-based technology. Critical review of cytologic errors. Acta Cytol 46:1105–1109.
21. Kok MR, van Der Schouw YT, Boon ME, Grobbee DE, Kok LP, Schreiner-Kok PG, van der Graaf Y, Doornewaard H, van den Tweel JG. 2001. Neural network-based screening (NNS) in cervical cytology: No need for the light microscope? Diagn Cytopathol 24: 426–434.
22. Karakitsos P, Pouliakis A, Meristoudis C, Margari N, Kassanos D, Kyrgiou M, Panayiotides JG, Paraskevaidis E. 2011. A preliminary study of the potential of tree classifiers in triage of high-grade squamous intraepithelial lesions. Anal Quant Cytol Histol 33:132–140.
23. Kok MR, Habers MA, Schreiner-Kok PG, Boon ME. 1998. New paradigm for ASCUS diagnosis using neural networks. Diagn Cytopathol 19:361–366.
24. Boon ME, Kok LP. 1993. Neural network processing can provide means to catch errors that slip through human screening of pap smears. Diagn Cytopathol 9:411–416.
25. Varlatzidou A, Pouliakis A, Stamataki M, Meristoudis C, Margari N, Peros G, Panayiotides JG, Karakitsos P. 2011. Cascaded learning vector quantizer neural networks for the discrimination of thyroid lesions. Anal Quant Cytol Histol 33:323–334.
26. Rorive S, N Dh, Fossion C, Delpierre I, Arbaguia N, Avni F, Decaestecker C, Salmon I. 2010. Ultrasound-guided fine-needle aspiration of thyroid nodules: Stratification of malignancy risk using follicular proliferation grading, clinical and ultrasonographic features. Eur J Endocrinol 162:1107–1115.
27. Haymart MR, Cayo M, Chen H. 2009. Papillary thyroid microcarcinomas: Big decisions for a small tumor. Ann Surg Oncol 16: 3132–3139.
28. Karakitsos P, Cochand-Priollet B, Pouliakis A, Guillausseau PJ, Ioakim-Liossi A. 1999. Learning vector quantizer in the investigation of thyroid lesions. Anal Quant Cytol Histol 21:201–208.
29. Atkinson EN, Mitchell MF, Ramanujam N, Richards-Kortum R. 1995. Statistical techniques for diagnosing CIN using fluorescence spectroscopy: SVD and CART. J Cell Biochem Suppl 23:125–130.
30. Deshpande V, Kapila K, Sai KS, Verma K. 1997. Follicular neoplasms of the thyroid. Decision tree approach using morphologic and morphometric parameters. Acta Cytol 41:369–376.
31. Breiman L, Friedman J, Stone CJ, Olshen RA. Classification and regression trees. Belmont, CA: Wadsworth International Group; 1984. X, 358 p.
32. Ausems EW, van der Kamp JK, Baak JP. 1985. Nuclear morphometry in the determination of the prognosis of marked atypical endometrial hyperplasia. Int J Gynecol Pathol 4:180–185.
33. Baak JP, Nauta JJ, Wisse-Brekelmans EC, Bezemer PD. 1988. Architectural and nuclear morphometrical features together are more important prognosticators in endometrial hyperplasias than nuclear morphometrical features alone. J Pathol 154:335–341.
34. Colgan TJ, Norris HJ, Foster W, Kurman RJ, Fox CH. 1983. Predicting the outcome of endometrial hyperplasia by quantitative analysis of nuclear features using a linear discriminant function. Int J Gynecol Pathol 1:347–352.
35. Fu YS, Ferenczy A, Huang I, Gelfand MM. 1988. Digital imaging analysis of normal, hyperplastic and malignant endometrial cells in endometrial brushing samples. Anal Quant Cytol Histol 10: 139–149.
36. Norris HJ, Becker RL, Mikel UV. 1989. A comparative morphometric and cytophotometric study of endometrial hyperplasia, atypical hyperplasia, and endometrial carcinoma. Hum Pathol 20: 219–223.
37. Tezuka F, Chiba R, Takahashi T. 1994. Morphometric and multivariate statistical detection of cancer cells in endometrial cytology. Anal Quant Cytol Histol 16:332–338.
38. Tezuka F, Chiba R, Takahashi T. 1995. Quantitative cytopathology of endometrial lesions. J Cell Biochem Suppl 23:147–150.
39. Baak JP, Kurver PH, Overdiep SH, Delemarre JF, Boon ME, Lindeman J, Diegenbach PC. 1981. Quantitative, microscopical, computer-aided diagnosis of endometrial hyperplasia or carcinoma in individual patients. Histopathology 5:689–695.
40. Karakitsos P, Kyroudes A, Pouliakis A, Stergiou EB, Voulgaris Z, Kittas C. 2002. Potential of the learning vector quantizer in the cell classification of endometrial lesions in postmenopausal women. Anal Quant Cytol Histol 24:30–38.
41. Fox H. 1984. The endometrial hyperplasias. Obstet Gynecol Annu 13:197–209.
42. Kurman RJ, Ellenson LH, Ronnett BM. Blaustein's pathology of the female genital tract. New York, NY: Springer; 2011 xv, 1246 p.
43. Baxes GA. Digital image processing: Principles and applications. New York: Wiley; 1994. xviii, 452 p.
44. Bibbo M, Bartels PH, Galera-Davidson H, Dytch HE, Wied GL. 1986. Markers for malignancy in the nuclear texture of histologically normal tissue from patients with thyroid tumors. Anal Quant Cytol Histol 8:168–176.
45. Jain AK. Fundamentals of digital image processing. Englewood Cliffs, NJ: Prentice Hall; 1989. xxi, 569 p.

46. Pitas I. Digital image processing algorithms and applications. New York: Wiley; 2000. 419 p.
47. Sonka M, Hlavac V, Boyle R. Image processing, analysis, and machine vision. London, New York: Chapman & Hall Computing; 1993. xix, 555 p.
48. Sá JPMd. Applied statistics: Using SPSS, Statistica, MATLAB, and R. Berlin; New York: Springer; 2007. xxiv, 505 p.
49. Tezuka F, Namiki T, Higashiiwai H. 1992. Observer variability in endometrial cytology using kappa statistics. J Clin Pathol 45:292–294.
50. Reagan JW. 1980. Can screening for endometrial cancer be justified? Acta Cytol 24:87–89.
51. Nagai S, Sawada T. 1983. [Cytological findings on endometrial carcinoma and endometrial hyperplasia]. (abstract) Nihon Sanka Fujinka Gakkai Zasshi 35:2283–2290.
52. Yanoh K, Norimatsu Y, Hirai Y, Takeshima N, Kamimori A, Nakamura Y, Shimizu K, Kobayashi TK, Murata T, Shiraishi T. 2009. New diagnostic reporting format for endometrial cytology based on cytoarchitectural criteria. Cytopathology 20:388–394.
53. Papaefthimiou M, Symiakaki H, Mentzelopoulou P, Tsiveleka A, Kyroudes A, Voulgaris Z, Tzonou A, Karakitsos P. 2005. Study on the morphology and reproducibility of the diagnosis of endometrial lesions utilizing liquid-based cytology. Cancer 105:56–64.
54. Papaefthimiou M, Symiakaki H, Mentzelopoulou P, Giahnaki AE, Voulgaris Z, Diakomanolis E, Kyroudes A, Karakitsos P. 2005. The role of liquid-based cytology associated with curettage in the investigation of endometrial lesions from postmenopausal women. Cytopathology 16:32–39.
55. Norimatsu Y, Shimizu K, Kobayashi TK, Moriya T, Tsukayama C, Miyake Y, Ohno E. 2006. Cellular features of endometrial hyperplasia and well differentiated adenocarcinoma using the Endocyte sampler: Diagnostic criteria based on the cytoarchitecture of tissue fragments. Cancer 108:77–85.
56. Watanabe J, Nishimura Y, Tsunoda S, Kawaguchi M, Okayasu I, Kuramoto H. 2009. Liquid-based preparation for endometrial cytology—Usefulness for predicting the prognosis of endometrial carcinoma preoperatively. Cancer 117:254–263.
57. Kyroudi A, Paefthimiou M, Symiakaki H, Mentzelopoulou P, Voulgaris Z, Karakitsos P. 2006. Increasing diagnostic accuracy with a cell block preparation from thin-layer endometrial cytology: A feasibility study. Acta Cytol 50:63–69.
58. Norimatsu Y, Kouda H, Kobayashi TK, Moriya T, Yanoh K, Tsukayama C, Miyake Y, Ohno E. 2008. Utility of thin-layer preparations in the endometrial cytology: Evaluation of benign endometrial lesions. Ann Diagn Pathol 12:103–111.
59. Norimatsu Y, Kouda H, Kobayashi TK, Shimizu K, Yanoh K, Tsukayama C, Miyake Y, Ohno E. 2009. Utility of liquid-based cytology in endometrial pathology: Diagnosis of endometrial carcinoma. Cytopathology 20:395–402.
60. Norimatsu Y, Shigematsu Y, Sakamoto S, Ohsaki H, Yanoh K, Kawanishi N, Kobayashi TK. 2012. Nuclear features in endometrial cytology: Comparison of endometrial glandular and stromal breakdown and endometrioid adenocarcinoma grade 1. Diagn Cytopathol 12:1077–1082.
61. Firat P, Mocan G, Kapucuoglu N. 2002. Liquid-based endometrial cytology: Endometrial sample collection by using Tao brush. Diagn Cytopathol 27:393–394.
62. Garcia F, Barker B, Davis J, Shelton T, Harrigill K, Schalk N, Meyer J, Hatch K. 2003. Thin-layer cytology and histopathology in the evaluation of abnormal uterine bleeding. J Reprod Med 48: 882–888.
63. Fambrini M, Buccoliero AM, Bargelli G, Cioni R, Piciocchi L, Pieralli A, Andersson KL, Scarselli G, Taddei G, Marchionni M. 2008. Clinical utility of liquid-based cytology for the characterization and management of endometrial polyps in postmenopausal age. Int J Gynecol Cancer 18:306–311.
64. Buccoliero AM, Castiglione F, Gheri CF, Garbini F, Fambrini M, Bargelli G, Pappalardo S, Scarselli G, Marchionni M, Taddei GL. 2007. Liquid-based endometrial cytology: Its possible value in postmenopausal asymptomatic women. Int J Gynecol Cancer 17:182–187.
65. Kipp BR, Medeiros F, Campion MB, Distad TJ, Peterson LM, Keeney GL, Halling KC, Clayton AC. 2008. Direct uterine sampling with the Tao brush sampler using a liquid-based preparation method for the detection of endometrial cancer and atypical hyperplasia: A feasibility study. Cancer 114:228–235.
66. Sams SB, Currens HS, Raab SS. 2012. Liquid-based Papanicolaou tests in endometrial carcinoma diagnosis. Performance, error root cause analysis, and quality improvement. Am J Clin Pathol 137: 248–254.
67. Michael CW, McConnel J, Pecott J, Afify AM, Al-Khafaji B. 2001. Comparison of ThinPrep and TriPath PREP liquid-based preparations in nongynecologic specimens: A pilot study. Diagn Cytopathol 25:177–184.
68. Rossi ED, Zannoni GF, Lombardi CP, Vellone VG, Moncelsi S, Papi G, Pontecorvi A, Fadda G. 2012. Morphological and immunocytochemical diagnosis of thyroiditis: Comparison between conventional and liquid-based cytology. Diagn Cytopathol 40:404–409.
69. Cochand-Priollet B, Cartier I, de Cremoux P, Le Gales C, Ziol M, Molinie V, Petitjean A, Dosda A, Merea E, Biaggi A, Gouget I, Arkwright S, Vacher-Lavenu MC, Vielh P, Coste J. 2005. Cost-effectiveness of liquid-based cytology with or without hybrid-capture II HPV test compared with conventional Pap smears: A study by the French Society of Clinical Cytology. Diagn Cytopathol 33:338–343.
70. Selvaggi SM. 2005. Background features of endometrial carcinoma on ThinPrep cytology. Diagn Cytopathol 33:162–165.
71. Guidos BJ, Selvaggi SM. 2000. Detection of endometrial adenocarcinoma with the ThinPrep Pap test. Diagn Cytopathol 23:260–265.