Automated identification of cancerous smears using various competitive intelligent techniques

G. DOUNIAS[1], B. BJERREGAARD[2], J. JANTZEN[3], A.TSAKONAS[4], N. AMPAZIS[1], G. PANAGI[5] and E. PANOURGIAS[6]

[1]Department of Financial and Management Engineering, University of the Aegean, 31 Fostini Str., 82100 Chios, Greece; [2]Herlev University Hospital, DK-2730 Herlev; [3]Oersted-DTU Automation, Technical University of Denmark, DK-2800 Kongens Lyngby, Denmark; [4]Artificial Intelligence and Information Analysis Laboratory, Department of Informatics, Aristotle University of Thessaloniki, 54124 Thessaloniki, [5]Department of Radiology, General Hospital of Chios 'Skilitsion', Chios; [6]Department of Radiology, Euroclinic Hospital, 9 Athanasiadou Str., 11521 Athens, Greece

Received September 6, 2005; Accepted September 28, 2005

Abstract. In this study the performance of various intelligent methodologies is compared in the task of pap-smear diagnosis. The selected intelligent methodologies are briefly described and explained, and then, the acquired results are presented and discussed for their comprehensibility and usefulness to medical staff, either for fault diagnosis tasks, or for the construction of automated computer-assisted classification of smears. The intelligent methodologies used for the construction of pap-smear classifiers, are different clustering approaches, feature selection, neuro-fuzzy systems, inductive machine learning, genetic programming, and second order neural networks. Acquired results reveal the power of most intelligent techniques to obtain high quality solutions in this difficult problem of medical diagnosis. Some of the methods obtain almost perfect diagnostic accuracy in test data, but the outcome lacks comprehensibility. On the other hand, results scoring high in terms of comprehensibility are acquired from some methods, but with the drawback of achieving lower diagnostic accuracy. The experimental data used in this study were collected at a previous stage, for the purpose of combining intelligent diagnostic methodologies with other existing computer imaging technologies towards the construction of an automated smear cell classification device.

Correspondence to: Dr Georgios Dounias, Department of Financial and Management Engineering, University of the Aegean, 31 Fostini Str., 82100 Chios, Greece
E-mail: g.dounias@aegean.gr

Key words: computer assisted pap-smear diagnosis, computational intelligence, clustering, feature selection, ANFIS, neuro-fuzzy systems, inductive machine learning, second order neural networks, nearest neighbor classification, genetic programming

Introduction

This report is the result of extensive collaborative work between engineers and doctors of different fields of expertise originating from different EU countries. The implementation of an efficient computer-assisted methodology for automated classification of cell images is important because manual screening is a tedious and error prone task. Various intelligent methods for data analysis and knowledge extraction are briefly explained and their performance is compared in terms of effectiveness, accuracy and comprehensibility, in the medical domain of pap-smear diagnosis. Extensive details for some of the intelligent methods applied to pap-smear diagnosis, together with related comparisons of performance and findings, can be found in refs. 1-3.

According to Meisels and Morin (4), cervix is the anatomical region between the body of the uterus and the vagina and is covered by epithelium. In the vaginal part of the cervix there is stratified squamous epithelium, in the endocervical part of the cervix there is simple columnar epithelium and between them there is the transitional zone. Histologically, the cells of the squamous epithelium are arranged in four layers: the basal, the parabasal, the intermediate and the superficial layer. The columnar epithelium consists only of one layer. Using a small brush, a specimen is taken from the cervix and transferred on to a slide. The smear is stained using the Papanikolaou method thus making it possible to see characteristics of cells on the microscope and classify them according to the homonymous classification.

The purpose of Papanikolaou classification (3,4) is to diagnose pre-malignant changes before they progress to invasive carcinoma. In normal columnar epithelial cells, the nucleus is located at the bottom of the cytoplasm. When viewed from the top, the nucleus seems larger. When viewed from the side, the cytoplasm seems larger. Cells of the basal and parabasal layer have small nuclei and cytoplasm. Cells of intermediate and superficial layer have small nuclei and larger cytoplasm. Squamous dysplastic cells generally, have larger and darker nuclei and tend to cling together in clusters.

Table I. Classification of cells and sample details.

Class no.	Cell diagnosis	Detailed description of the type of cells	
1	Normal	Columnar epithelial	50
2	Normal	Parabasal squamous epithelial	50
3	Normal	Intermediate squamous epithelial	50
4	Normal	Superficial squamous epithelial	50
5	Abnormal	Mild squamous non-keratinizing dysplasia	100
6	Abnormal	Moderate squamous non-keratinizing dysplasia	100
7	Abnormal	Severe squamous non-keratinizing dysplasia	100

In severe dysplasia, nuclei are large, with dark granules and usually deformed. Cytoplasm is small and dark compared to nuclei. Cervical carcinoma is the most common malignancy of the female genital system. Carcinogenesis is a long-standing procedure, which begins from normal epithelium that becomes dysplastic, evolves to carcinoma *in situ* and then to cancer. The long time interval between the stages, allows the possibility of an early diagnosis, with complete cure.

Materials and methods

The aim of the study was the attempt for pap-smear classification through different competitive computational intelligence techniques, in order to establish an efficient computer assisted methodology for automated classification of cell images (1-3). The collection of cell images is done automatically with the aid of software especially designed to recognise, under the electronic microscope, the regions of nucleus-cytoplasm-background. Our sample consists of 500 pap-smear cases (images of smears, available also on the web) described by 20 numerical diagnostic attributes (cell measurements), belonging to seven different classes (Table I). Four classes represent normal epithelial cells (columnar, parabasal, intermediate and superficial) and three classes represent squamous dysplasia divided as mild, moderate and severe. Specifically, the following measures were the diagnostic attributes: nucleus area, cytoplasm area, K/C index (nucleus area divided to cytoplasm + nucleus area), nucleus intensity, cytoplasm intensity, nucleus shortest diameter, nucleus longest diameter, nucleus elongation, nucleus roundness, cytoplasm shortest diameter, cytoplasm longest diameter, cytoplasm elongation, cytoplasm roundness, nucleus perimeter, cytoplasm perimeter, nucleus x coordinate, nucleus y coordinate, cytoplasm x coordinate, cytoplasm y coordinate, nucleus position in cytoplasm, nucleus maxima, nucleus minima, cytoplasm maxima, cytoplasm minima.

The intelligent methodologies used for the construction of pap-smear classifiers were: a) hard C-means, b) fuzzy C-means, c) Gustafson-Kessel clustering method, d) feature selection and clustering, e) ANFIS neuro-fuzzy classification, f) nearest neighbour classification, g) entropy information based inductive machine learning, h) genetic programming-derived crisp rule-based system, i) LMAM/OLMAM type second order neural networks.

Below we briefly explain the main idea of each intelligent approach applied on the examined pap-smear data and we give the basic literature review for the interested reader.

Hard C-means. Hard C-means (HCM) is described in ref. 5, as well as in refs. 1,2, and in fact, corresponds to the discovery of natural clusters in the data. Standard C-means clustering has relatively few parameters to tune, and therefore it can be an attractive method to use. Furthermore, the used data, have no limit on how many dimensions it can have, i.e. the number of features. But the downfall is that the method will often give a worse result, because it is not good at utilizing and modelling its data. Clustering (either hard or fuzzy, see below) can be supervised or unsupervised, if an expert guides or not the process of identifying the desired number of clusters to be formed.

Fuzzy C-means. The related methodology was first introduced by Bezdek (6), and a complete presentation of the approach can be found in ref. 7. Fuzzy c-means (FCM) in fact modifies HCM mentioned above, by allowing the data points belonging to all of the clusters, with memberships in the interval [0, 1]. The total membership to all clusters from a single data point should be equal to 1. Similar downfall as above, exists for the FCM. Linear feature scaling takes place in FCM.

Gustafson-Kessel clustering method. The Gustafson-Kessel clustering method was first described in 1979 (8). FCM produces spherical clusters and does not let a cluster change its shape dependent on the data. The Gustafson-Kessel method has the advantage of forming clusters of data, which adapt to hyper ellipsoidal shapes. In contrast to fuzzy C-means, the Gustafson-Kessel clustering method is capable of making its own non-linear scaling of features that is different from cluster to cluster.

Feature selection and clustering. The approach is extensively described in refs. 1,2, while details on feature selection can be found in ref. 9. It constitutes a hybrid method in which feature selection is first attempted in order to reduce complexity of the search space. Then methods like the previously described Gustafson-Kessel clustering are applied on the reduced feature set. Feature selection is generally used to select subsets of features that give the best classification results. Some features can have some noise, others maybe do not help classification at all, and others maybe cannot be utilized by the chosen classifier. Sometimes even the available number of data is too small to justify the amount of features with a particular classifier. In some instances feature selection is also combined with hierarchical classification, a process usually guided by collaborating field experts.

ANFIS neuro-fuzzy classification. The method was proposed by Jang (10) and is also described in detail in ref. 11. A neural network is a method that effectively approximates a function,

Table II. Classification accuracy results, for separating: a) between normal and abnormal cells (2-class problem), and b) between all 7 diagnostic classes; accuracy refers to the best performance obtained on test (new) data.

Computational intelligence methodologies used for pap-smear classification	Classification accuracy for the 2-class problem (%)	Classification accuracy for the 7-class problem (%)
Hard C-means	94-96	72-80
Fuzzy C-means	96-97	72-77
Gustafson-Kessel clustering method	89-96	<75
Feature selection and clustering	90-97	<75
ANFIS neuro-fuzzy classification	96	<75
Nearest neighbour classification	96	<75
Inductive machine learning	<75	<75
Genetic programming	89	81
Second order neural networks	99	<75

but it is usually impossible to interpret the result in terms of natural language. Fuzzy logic on the other side is an attempt to model, quantify and describe in natural language, existing vagueness and ambiguity of every day phenomena. The fusion of neural networks and fuzzy logic in neuro-fuzzy models provide learning as well as readability. This combination often proves useful, because the models can be interpreted and supplemented by domain experts.

Nearest neighbour classification. The method has been introduced (12) and is also adequately described in ref. 13. The nearest neighbour classification methodology is a popular non-linear classifier. In fact, it is a data classification method based on k-NN density estimation. The nearest neighbour rule is the selection of class of the nearest training sample. In other words, given a set of classified records the method attempts to classify a new record, by examining those records nearest to the new one.

Entropy information based inductive machine learning. The core idea of the divide and conquer principle was described by Hunt *et al* (14) and then Quinlan (15) combined this idea with entropy information measures in an intelligent algorithmic scheme. Systematically repeated events are preferred to be selected as premise parts of generalized decision rules, whereas random occurrences existing in a dataset are identified and excluded, with the aid of entropy measures. Methods that further accelerate the algorithm performance like boosting (the most often used advancement) can be found in ref. 16. Boosting is a general method of producing very accurate prediction rules by combining rough and moderately inaccurate rules of thumb.

Genetic programming-derived crisp rule-based system. Genetic programming was first introduced by Koza (17,18) as an extension of genetic algorithms (19) whereas advances in genetic programming can be found in ref. 20 and the description of the specific methodology implemented in this study can be found in ref. 21. Any evolutionary computing approach incorporates variations of the Darwinian principle of evolution and the survival of the fittest. A hybrid approach has been applied on the pap-smear data, which combines standard genetic programming and heuristic hierarchical crisp rule-base construction. Then, genetic programming for the production of crisp rule based systems is attempted. Results denote advantages of the method, in terms of efficiency, accuracy and comprehensibility.

LMAM/OLMAM type second order neural networks. The method was first introduced in ref. 22 and part of the methodology can also be found in ref. 23, while ref. 24 gives details on its application to pap-smear diagnosis. Specifically, highly efficient second order neural network training algorithms are applied, namely the LMAM (Levenberg-Marquardt with adaptive momentum) and OLMAM (optimized Levenberg-Marquardt with adaptive momentum), for the construction of an efficient pap-smear test classifier. The algorithms are based on iterations of the form employed in the Levenberg-Marquardt (LM) method for non-linear least squares problems with the inclusion of an additional adaptive momentum term arising from the formulation of the training task as a constrained optimization problem. The classification results obtained from the application of the algorithms on a standard benchmark pap-smear data set reveal the power of the two methods to obtain excellent solutions in difficult classification problems whereas other standard computational intelligence techniques achieve inferior performances.

Results

There were two kinds of experiments that generally took place with each computational intelligent methodology described above, using the datasets given in Table I. First, the methods were tested on their ability to handle the 2-class diagnosis problem, i.e. to classify adequately pap-smear cases as normal (those belonging to classes 1, 2, 3, 4) or abnormal (those belonging to classes 5, 6, 7). Second, the methods were applied on the 7-class diagnosis problem, i.e. trying to find the exact class to which a cell belongs. All the experiments took place following a standard 90-10% cross-validation approach, which will be considered from now on the default testing

performance condition for this study. Detailed results regarding the accuracy obtained by each intelligent methodology on both the 2-class and the 7-class problem are presented in Table II. The majority of the methodologies applied, performed very well in the 2-class problem, exceeding 90% of testing accuracy in most cases. On the contrary, the majority of the methods performed rather poorly in the 7-class problem achieving generally a testing accuracy <75% (with a few exceptions). Second order neural networks scored highest in the 2-class problem while genetic programming showed the best average generalization capability, in both kinds of problems examined. Below we provide detailed comments for each of the methods presented in the previous section, focusing also on the medical viewpoint of the results acquired.

Hard C-means. The hard C-means approach achieved satisfactory performance in the 2-class problem. The testing accuracy ranged approximately between 94-96%. In the 7-class problem, the performance was considerably lower, ranging from 72 to 80%. The supervised clustering procedure performed much better than unsupervised clustering for nearly all results, perhaps because clusters were not nicely separable in natural clusters. Supervised clustering allowed clusters to adapt to data of the same class, without interference from data of other classes. From the medical viewpoint, it was interesting that the results most of the time contained six features, four of which described nucleus characteristics (nucleus area, nucleus shortest and longest diameter, and minima in nucleus) and the other two were N/C ratio and cytoplasm longest diameter.

Fuzzy C-means. The fuzzy C-means approach performed well when it had to discriminate between the two major classes (1, 2, 3, 4) vs (5, 6, 7). Testing accuracy was high (96-97%) for both, unsupervised and supervised fuzzy C-means. Unfortunately, testing accuracy was very low (72-77%) when it had to discriminate among all seven classes. There was no considerable difference in the performance between unsupervised and supervised fuzzy C-means, as well as in the feature selection process, when this was preceding clustering. In general, the discrimination process involved seven features in all the experiments, namely cytoplasm area, N/C ratio, brightness of nucleus and cytoplasm, nucleus shortest and longest diameter and cytoplasm shortest diameter.

Gustafson-Kessel clustering method. The testing accuracy of the unsupervised Gustafson-Kessel clustering method, when discriminating between two classes (normal vs abnormal) was 89%, but the supervised method's testing accuracy was much higher at 96%. When the method had to discriminate among all seven classes the results were poor, like previously. The features most frequently involved in all the Gustafson-Kessel based experiments were nucleus area and elongation, cytoplasm shortest diameter, nucleus perimeter and maxima in nucleus.

Feature selection and clustering. When feature selection was combined with Gustafson-Kessel clustering, compared to the previous approach, it produced better results in the discrimination between the two classes. Especially the supervised model performed very well (97%) in comparison to the unsupervised model (90%). For the 7-class problem, the feature selection and clustering method performed worse. Obviously, this improvement in performance had to do with the application of feature selection prior to the algorithm, as expected. The two features that are principally involved in the discrimination process for our given dataset, either for the Gustafson-Kessel clustering method, or for the hierarchical clustering, both driven by feature pre-selection, are N/C ratio and nucleus longest diameter.

ANFIS neuro-fuzzy classification. Regarding the classification between normal and abnormal smear cells the algorithm achieved 96% testing accuracy. For the 7-class problem, the ANFIS method performed much worse. In terms of accuracy, the method seems to perform as well as the nearest neighbour method described below. Regarding the medical viewpoint of the results, experts played a key role. Specifically, a number of expert rules were initially given to the system, as general guidelines for medical diagnosis, combining the nucleus area and colour, the cytoplasm colour and the N/C ratio. For example, small and dark nucleus, light cytoplasm and small N/C, implies normal cells. Three similar general rules define different types of dysplasia. Later, modified expert rules were added to the system, building a network consisting of 34 adaptive parameters, 27 of which were premise and 7 were consequent parameters. This architecture obtained reduced false negative rate but increased false positive rate.

Nearest neighbour classification. Nearest neighbour performed rather high in terms of testing accuracy in the 2-class problem, reaching 96%, while its comprehensibility was low. The method performed worse in the 7-class problem. It used a reduced set of 12 features and was able to classify cells with a small false-negative rate. The nearest neighbour method, combined with simulated annealing for feature selection, insisted on the combination of six attributes for discriminating classes, namely, nucleus brightness, cytoplasm brightness, nucleus shortest diameter, nucleus roundness, cytoplasm - shortest diameter and nucleus position.

Entropy information based inductive machine learning. The advantage of this methodology is generally considered the high comprehensibility of the produced results. Accuracy on the testing set was very low, approximately 70%, for discriminating among all seven classes. For the 2-class problem, performance was generally better. One of the most comprehensible rules was extracted from the application of inductive machine learning to the 7-class problem. The most interesting rule acquired, covering approximately 10% of the total sample cases, was the following: K/C ≤0.044 and KernelLong >8.23 and CytoLong >52.39 and KernePeri >27.56 then class 3 (0.979); where, K/C denotes the ratio between nucleus area and cytoplasm area, KernelLong denotes the longest diameter of nucleus, CytoLong is the longest diameter of the cytoplasm, and KernelPeri is the perimeter of the nucleus. As observed, the above rule achieves high probability of correctly classifying new cases, approaching 98%.

Genetic programming-derived crisp rule-based system. GP-derived results were of medium to high accuracy for both, the 2- and the 7-class problems, achieving 89% and 81% respectively. Overall comprehensibility was medium to low, but a few interesting results occurred from the medical viewpoint. The GP approach produced a sequence of inter-dependent rules, part of which clearly denoted that the quantity KA/CS (nucleus area divided by cytoplasm shortest diameter) could perhaps discriminate among normal cells of class 4 and abnormal cells of classes 5, 6, 7, with a perfect accuracy of 100%. It would be interesting to further investigate statistically if this measure is useful for discriminating normal from abnormal cells.

LMAM/OLMAM type second order neural networks. Several experiments took place with both LMAM and OLMAM methods, either directly applied on the entire data set, or after feature selection (best performance occurred with a selection of 9 features). The methods obtained the highest diagnostic accuracy approaching 99% on new (test) data, but the outcome was totally lacking comprehensibility, behaving as black-box architecture.

Discussion

Most of the methods performed well when they had to discriminate between two classes, i.e. between normal cells (1, 2, 3, 4) and abnormal cells (5, 6, 7), whereas their performance was reduced when they had to discriminate among all seven cell types of Table I. In respect to the classification performance, it is hard to separate abnormal cells from each other and columnar cells are sometimes classified as severe dysplastic cells, i.e. if they are viewed from above. The differences among: a) normal and abnormal cells, b) various types of normal cells or, c) different grades of dysplastic cells, may be extremely obscured. In every day practice, medical experts may face exactly the same problems, in the cell discrimination process and thus, in the decision for the final staging of the patient, which originate mainly from the variety of form, of human cells. In order to come to an accurate conclusion, experts use established criteria mentioned in medical literature, two of which, where generated as extracted rules by inductive machine learning and genetic programming. In the result of the intelligent methods, the frequent appearance of various nucleus characteristics and N/C ratio reflects the real world medical experts' procedure which corresponds to the fact that, the appearance of the nucleus area plays a decisive role in the characterization of a cell as benign or malignant. Some methods may also play the role of an accurate second opinion tool, in the future, while others may 'provocate' the current medical knowledge for further knowledge extraction.

According to FIGO classification of cervical cancer, in stage 0 (carcinoma *in situ*) and stage I (carcinoma restricted to cervix) 5-year healing rates range between 75-99%. In the other stages, healing percentage diminishes, thus proving the importance of confirming an early diagnosis. The medical task of classifying and daily diagnosing several pap-smear images is a time-consuming process, at present done manually. A faster computer assisted technique able to perform classification and diagnosis, somewhat automatically and competitively to human experts, would represent a great advancement for cytopathologists in the future and is already under development.

The experimental data used in this study were collected at a previous stage, for the purpose of combining intelligent diagnostic methodologies with other existing computer imaging technologies towards the construction of an automated smear cell classification device. Specifically, after the training and automatic selection of smears from the microscope, the captured images of cells (data) are analysed through CHAMP, a software programme of DIMAC Co. (25), according to digital image processing principles and classified (through a C.I. method). An automated screening system (named Aphrodite) is able to discriminate among healthy slides and sort out 50% of them, thus diminishing the work load of cytopathologists (25). The system is now being tested in four different university hospitals, waiting for approval. Additional trials are about to take place, using a much larger and complete pap-smear database. The currently existing pap-smear database used in this study for experimentation is freely accessible in the web site given in ref. 26.

References

1. Byriel J: Neuro-Fuzzy Classification of Cells in Cervical Smears. MSc Thesis, Technical University of Denmark, Department of Automation, http://fuzzy.iau.dtu.dk/download/byriel99.pdf, 1999.
2. Martin E: Pap-Smear Classification. MSc Thesis, Technical University of Denmark, Oersted-DTU, Department of Automation, 2003.
3. Koss L: The application of PAPNET to diagnostic cytology. In: Artificial Neural Networks in Biomedicine. Lisboa P, Ifeachor E and Szczepaniak P (eds). Springer, London, pp51-68, 2000.
4. Meisels A and Morin C: Cytopathology of the Uterus. 2nd edition. ASCP Press, Chicago, IL, 1997.
5. Jain AK and Dubes RC: Algorithms for Clustering Data. Prentice-Hall, NJ, 1988.
6. Bezdek JC: Pattern Recognition with Fuzzy Objective Function Algorithms. Plenum, New York, 1981.
7. Höppner F, Klawonn F, Kruse R and Runkler T: Fuzzy Cluster Analysis. Wiley, New York, 1999.
8. Gustafson EE and Kessel WC: Fuzzy Clustering with a Fuzzy Covariance Matrix. IEEE CDC, San Diego, CA, pp761-766, 1979.
9. Liu H and Motoda H: Feature Selection for Knowledge Discovery and Data Mining. The Kluwer International Series in Engineering and Computer Science, Boston, MA, 1998.
10. Jang JSR: ANFIS: Adaptive-network-based fuzzy inference systems. IEEE Transactions on Systems. Man Cybernetics 23: 665-685, 1993.
11. Jantzen J: Neurofuzzy Modelling. Tech report no. 98-H-874 (nfmod). Technical University of Denmark Oersted-DTU, URL http://fuzzy.iau.dtu.dk/download/nfmod.pdf, 1998.
12. Cover TM and Hart PE: Nearest neighbour pattern classification. IEEE Trans Information Theory 13: 21-27, 1967.
13. Duda RO, Hart PE and Stork DG: Pattern Classification. 2nd edition. John Wiley & Sons, New York, 2001.
14. Hunt EB, Marin J and Stone PJ: Experiments in induction. Academic Press, New York, 1966.
15. Quinlan JR: C4.5: Programs for Machine Learning. Morgan Kaufmann, San Mateo, 1993.
16. Schapire R: The strategy of weak learnability. Machine Learning 2: 197-227, 1990.
17. Koza JR: Genetic Programming - On the Programming of Computers by Means of Natural Selection. MIT Press, Cambridge, MA, 1992.
18. Koza JR: Genetic Programming II - Automatic Discovery of Reusable Programs. MIT Press, Cambridge, MA, 1994.
19. Goldberg DE: Genetic Algorithms in Search, Optimization and Machine Learning. Addison-Wesley, 1989.

20. Angeline PJ and Jinnear KE Jr: Advances in Genetic Programming. MIT Press, 1996.
21. Tsakonas A, Dounias G, Jantzen J, Axer H, Bjerregaard B and von Keyserlingk DG: Evolving rule based systems in two medical domains using genetic programming. Artif Intell Med 32: 195-216, 2003.
22. Ampazis N and Perantonis SJ: Two highly efficient second order algorithms for training feedforward networks. IEEE Trans Neural Networks 13: 1064-1074, 2002.
23. Hagan MT and Menhaj M: Training feedforward networks with the Marquardt algorithm. IEEE Trans Neural Networks 5: 989-993, 1994.
24. Ampazis N, Dounias G and Jantzen J: Efficient Second Order Neural Network Training Algorithms for the Construction of a Pap-Smear Classifier. Lecture Notes in Artificial Intelligence - LNAI 3025. Springer-Verlag, Berlin, Heidelberg, pp230-245, 2004.
25. Automated Screening System Aphrodite, http://www.dimac-imaging.com/
26. The Pap-Smear Database, http://fuzzy.iau.dtu.dk/smear/