

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
31 May 2007 (31.05.2007)

PCT

(10) International Publication Number
WO 2007/059629 A1

(51) International Patent Classification:

G01N 15/00 (2006.01) *G01N 33/483* (2006.01)
C12M 1/34 (2006.01) *G06F 19/00* (2006.01)
C12Q 1/00 (2006.01) *G06T 7/40* (2006.01)
G01N 15/02 (2006.01) *G01N 1/30* (2006.01)
G01N 21/84 (2006.01)

(21) International Application Number:

PCT/CA2006/001933

(22) International Filing Date:

27 November 2006 (27.11.2006)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

60/739,412 25 November 2005 (25.11.2005) US

(71) Applicant (for all designated States except US): **BC CANCER AGENCY** [CA/CA]; 600 West 10th Avenue, Vancouver, British Columbia V5Z 1G1 (CA).

(71) Applicants and

(72) Inventors: **MACAULAY, Calum, E.** [CA/CA]; 338 East 37th Avenue, Vancouver, British Columbia V5W 1E7 (CA). **GUILLAUD, Martial, Daniel** [CA/CA]; 1950 Matthews Avenue, Vancouver, British Columbia V6J 2T7 (CA).

(72) Inventor; and

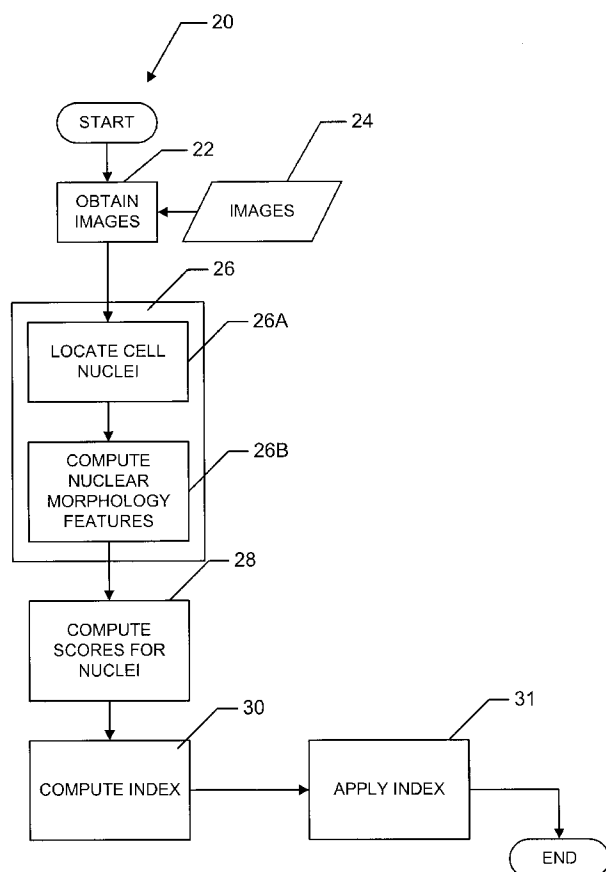
(75) Inventor/Applicant (for US only): **ROSIN, Miriam, Pearl** [CA/CA]; 160 Aspenwood Drive, Port Moody, British Columbia V3H 5A5 (CA).

(74) Agents: **MANNING, Gavin, N.** et al.; Oyen Wiggs Green & Mutala LLP, 480 - The Station, 601 West Cordova Street, Vancouver, British Columbia V6B 1G1 (CA).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LV, LY, MA, MD, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS,

[Continued on next page]

(54) Title: APPARATUS AND METHODS FOR AUTOMATED ASSESSMENT OF TISSUE PATHOLOGY



(57) Abstract: Methods and apparatus for assessing tissue pathology involve computing an index having a value based upon measures of a plurality of morphological features of cell nuclei in the tissue. The methods may be performed completely automatically or semi-automatically. The index value can be predictive of outcome. The index value may be determined by computing discriminant scores for the cell nuclei based upon values of the measures of morphological features and classifying the nuclei into bins based upon the discriminant score values. The index may be based upon proportions of the nuclei classified in different ones of the bins.

WO 2007/059629 A1



RU, SC, SD, SE, SG, SK, SL, SM, SV, SY, TJ, TM, TN,
TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

- (84) **Designated States** (*unless otherwise indicated, for every kind of regional protection available*): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Declaration under Rule 4.17:

— *of inventorship (Rule 4.17(iv))*

Published:

— *with international search report*
— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

**APPARATUS AND METHODS FOR AUTOMATED ASSESSMENT OF TISSUE
PATHOLOGY**

Reference to Related Application

5 **[0001]** This application claims priority from United States application No. 60/739,412
filed on 25 November 2005 and entitled NOVEL SYSTEMS AND METHODS FOR
ASSESSMENT OF ORAL PRE-MALIGNANT LESIONS. For purposes of the United
States of America, this application claims the benefit of the filing date of United States
patent application No. 60/739,412 filed on 25 November 2005 which is hereby
10 incorporated herein by reference.

Technical Field

[0002] This invention relates to apparatus and automated methods for assessing the
pathology of tissues. In some embodiments, the invention is applied to evaluating the
15 risk that lesions will develop into cancer.

Background

[0003] Cancer is a disease that claims a large number of lives each year. For many
types of cancer the best currently-known treatments are only partially effective. For
20 example, each year 500,000 people worldwide are diagnosed with head and neck
squamous cell carcinoma (HNSC). The 5-year survival rate has remained unchanged at
about 50 % for the past few decades. For many types of cancer, early detection and
treatment are critical to improving the survival rate.

25 **[0004]** Medical systems around the world lack the financial and other resources to
closely follow all patients who have lesions that could be pre-cancerous. A tool capable
of evaluating the risk that a lesion will develop into cancer would allow medical
resources to be directed in a more focussed manner. For example, the majority (~90
%) of oral pre-malignant lesions (OPLs) do not progress to invasive oral squamous cell
30 carcinoma. A method for determining which OPLs will progress to cancer would allow
at-risk patients to be followed closely to permit earlier intervention (e.g. surgery,
chemoprevention or radiation etc.) with improved survival outcomes. OPLs are
pre-neoplastic lesions of the oral cavity. OPLs include conditions described in the field
as oral epithelial dysplasia; mild dysplasia; leukoplakia; erythroplakia; hyperplasia,
35 moderate dysplasia; low-grade dysplasia; and, high grade dysplasia.

[0005] Current methods for detecting many types of cancer require skilled human
intervention. For example, many types of cancer screening are performed by a
pathologist or cytotechnologist who examines images of tissue samples for cells that
40 look abnormal to the trained eye.

[0006] Standard histopathological grading involves classifying tissues into various categories. for example, OPLs may be classified into the categories “Normal”; “Hyperplasia”; Mild Dysplasia”; “Moderate Dysplasia”; “Severe Dysplasia” and “Squamous Cell Carcinoma”. Standard histopathological grading does not correlate well with progression or patient outcome for low-grade lesions and therefore has low prognostic value. Because histological grading involves the application of human judgment, the results are not consistently reproducible. Further, these tissue assessment methods have the disadvantage that they are time consuming and expensive to implement and tend to be qualitative rather than quantitative.

[0007] Other techniques for the assessment of tissues (such as loss of heterozygosity (“LOH”) measurements, molecular biomarkers, imaging, may provide some prognostic information but are either too slow, costly (or both) to be readily implemented clinically. Further, some such methods can only be performed for lesions that are large enough to provide enough tissue for molecular analysis.

[0008] It has been found that LOH patterns can be used to classify OPLs that are not distinguishable by histopathological grading according to cancer risk. Rosin, M. P., et al. *Use of allelic loss to predict malignant risk for low-grade oral epithelial dysplasia*. Clin Cancer Res, 6: 357-362, 2000 used LOH to classify OPLs having minimal (mild/moderate) or no dysplasias into different risk categories. Compared with OPLs without LOH at 3p and 9p (low-risk), those with LOH at 3p and/or 9p but not in other arms (intermediate-risk) demonstrated a 3.8-fold increase in relative cancer risk. OPLs with LOH at 3p and/or 9p plus additional losses (at 4q, 8p, 11q, or 17p, high-risk) demonstrated 33-fold increases in relative cancer risk.

[0009] Despite the billions of dollars that have been spent worldwide on cancer research, there is currently, and has been for many years, an urgent need for practical ways to do one or more of:

- detect cancer;
- assess the malignant potential of lesions;
- assess the likely response of a particular cancer to treatment; and,
- assess the disease prognosis (e.g. survival or likelihood of progression to cancer within a time period) of affected patients.

There is a particular need for such ways in relation to oral cancers.

Summary

[0010] Various non-limiting embodiments of the invention are described herein. These embodiments are meant to be exemplary and illustrative, not limiting in scope. The various embodiments have different sets of features. As will be apparent to those of skill in the art, features of the example embodiments presented herein may be combined in other combinations and sub-combinations to yield other embodiments of the invention.

[0011] Methods and apparatus according to this invention perform quantitative morphometry on the nuclei of cells in tissue from patients.

[0012] One aspect of the invention provides methods for the automated assessment of tissue pathology. The methods comprise: obtaining images of a plurality of cell nuclei in a tissue sample; computing values for a plurality of morphometric features of the plurality of cell nuclei, the morphometric features including one or more texture features; from the values of the morphometric features, computing a value for an index characterizing tissue of the tissue sample; and, recording the index value.

[0013] Another aspect of the invention provides methods for the automated assessment of tissue pathology. The methods derive an index indicative of a likelihood of an outcome for a tissue. The methods comprise: obtaining images of a plurality of cell nuclei in sample of the tissue; computing values for a plurality of morphometric features of the plurality of cell nuclei; computing discriminant scores for the plurality of nuclei based upon the values of the morphometric features for the nuclei; classifying the nuclei into a bins of a series of bins based upon the discriminant scores for the nuclei; computing the index value based upon the relative numbers of the nuclei classified in the bins; and, recording the index value.

[0014] Another aspect of the invention provides methods for the automated assessment of oral lesions. The methods comprise: obtaining images of cell nuclei in a sample of tissue from the oral lesion; computing values for a plurality of morphometric features of the plurality of cell nuclei; from the values of the morphometric features, computing a value for an index indicative of a likely outcome for the lesion; recording the index value.

[0015] Another aspect of the invention provides apparatus for automatically assessing tissue pathology. The apparatus comprises a data processing unit connected to obtain images of cell nuclei in tissue samples. The data processing unit comprises: nucleus identification means for identifying the cell nuclei in the images; feature value

computation means for computing values for a plurality of morphometric features of a cell nucleus identified by the nucleus identification means, the morphometric features including one or more texture features; index calculation means for computing a value for an index characterizing the tissue from the values of the morphometric features
5 calculated by the feature value computation means; and, means for recording index values computed by the index calculation means.

[0016] Apparatus provided by another aspect of the invention comprises: a data processing unit connected to obtain images of cell nuclei in tissue samples. The data
10 processing unit comprises: nucleus identification means for identifying the cell nuclei in the images; feature value calculation means for computing values for a plurality of morphometric features of the plurality of cell nuclei; discriminant score computation means for computing discriminant scores for the plurality of nuclei based upon the values of the morphometric features for the nuclei; classification means for classifying
15 the nuclei into a bins of a series of bins based upon the discriminant scores for the nuclei; index computation means for computing the index value based upon the relative numbers of the nuclei classified in the bins; and, means for recording the index value.

[0017] Another aspect of the invention provides program products comprising media
20 bearing software instructions which, when executed by a data processor, cause the data processor to coordinate the performance of a method according to the invention.

[0018] Further aspects of the invention and features of various embodiments of the invention are illustrated in the accompanying drawings and described below.

25

Brief Description of the Drawings

[0019] Non-limiting exemplary embodiments are illustrated in the drawings.

[0020] Figure 1 is a flow chart illustrating a method for assessing the pathology of
30 tissue samples.

[0021] Figure 1A is a flow chart illustrating a method for computing a morphological index.

35 [0022] Figure 1B is a flow chart illustrating a method for training a system that performs the method of Figure 1.

[0023] Figure 2 shows distributions of discriminant scores for normal-like cells and cancer-like cells.

[0024] Figure 3 shows apparatus according to an embodiment of the invention.

[0025] Figure 3A shows apparatus according to another embodiment of the invention.

- 5 [0026] Figures 4 and 4A illustrate correlations between morphological index values for a number of oral preneoplastic lesions and histopathological grading for those lesions.

[0027] Figure 5 shows MI values for cases exhibiting hyperplasia and low-grade dysplasia that have been classified according to whether or not they progress to cancer.

10

[0028] Figure 6 shows the progression to cancer over time for a set of lesions having morphological index values above a threshold value and another set of lesions having morphological index values below the threshold.

15 Description

[0029] Throughout the following description specific details are set forth in order to provide a more thorough understanding to persons skilled in the art. However, well known elements may not have been shown or described in detail to avoid unnecessarily obscuring the disclosure. Accordingly, the description and drawings are to be regarded
20 in an illustrative, rather than a restrictive, sense.

20

[0030] Figure 1 shows a method 20 for determining the malignant potential (risk of progression) of a lesion and the probability of survival of the subject patient. Block 22 obtains images 24 of cells in a tissue sample from the patient. The tissue sample may
25 comprise, for example a biopsy taken from a lesion in the patient. In some embodiments the lesion is an OPL. Images 24 are digital images. Images 24 may be obtained directly from a digital system that images the tissues (such as a digital microphotography system). Images 24 could also be obtained by scanning microphotographs although this adds an additional step.

30

[0031] In a prototype embodiment, images 24 are obtained by imaging thin sections of tissue with a high-resolution imaging system comprising a 12-bit double correlated sampling digital camera designed for interactive analysis of tissue sections. The imaging system is a modified version of the Cyto-Savant automated quantitative system
35 described in Garner, D., et al. *The Cyto-Savant system*, p. 305-317. New York: Igaku-shoin, 1994. The illumination wavelength was 600 ± 5 nm, corresponding to the absorption peak of Thionin stain. The effective pixel sampling space within the plane of the sample was $0.34 \mu\text{m}^2$ and the effective pixel sampling area was $0.116 \mu\text{m}^2$.

35

[0032] Images 24 permit measurement of one or more of the shape, size and texture of DNA distribution within cells in the tissue sample. Various imaging techniques may be used to obtain images 24. Some non-limiting examples of imaging techniques that may be applied to obtain images 24 are:

- 5 • optical microscopy of stained tissue specimens;
- phase-contrast microscopy of stained or unstained tissue specimens;
- fluorescence microscopy;
- and the like.

Alternative high-resolution imaging techniques that permit measurement of the
10 morphological features of cell nuclei may also be used to obtain images 24.

[0033] In some embodiments the tissue samples are sectioned. This may be performed in any suitable manner. For example, a biopsy, which comprises a full-thickness epithelium in some embodiments, may be prepared by:

- 15 • fixing the biopsy in a preservative, such as formalin;
- embedding the biopsy in a suitable material such as paraffin,
- sectioning the embedded biopsy, for example, with a microtome; and,
- where necessary or desirable to enhance the acquisition of images 24, staining the embedded biopsy with a suitable stain. In some embodiments, the stain is a
20 stain that permits the quantitative determination of DNA, for example a Feulgen-Thionin stain. Other stains may be used in other embodiments.

The biopsy may be sectioned along cutting planes extending perpendicular to the surface of the skin in some embodiments.

25 [0034] In some embodiments, the method is carried out in parallel with histopathology grading by a trained person. In such embodiments, some sections of the biopsy may be prepared for histopathology (for example, by staining with a suitable stain such as haematoxylin-eosin) and other sections may be prepared for the acquisition of images 24 by staining with a suitable stain. In some embodiments, a MI is determined from
30 images of haematoxylin / eosin-stained sections (which could be the same sections prepared for histopathological investigation). Such embodiments have the advantages of avoiding the need to make additional sections of the biopsy and eliminating staining steps.

35 [0035] In some embodiments, sample-to-sample variations in staining intensity are corrected for by collecting leukocytes/lymphocytes from the sample. This is optional.

[0036] In block 26, image analysis is performed on a number of the cell nuclei that are discernable in images 24. Block 26A locates suitable cell nuclei. In block 26B, the

image analysis computes values for each of a plurality of nuclear morphology features of the cell nuclei. Typically the image analysis in block 26 computes the nuclear morphology features for at least 35 cell nuclei. Preferably the image analysis in block 26 computes the nuclear morphology features for 50 or more cell nuclei. In some
5 embodiments, the nuclear morphology features are computed for 50 to 100 cell nuclei.

[0037] In some embodiments, the cell nuclei are identified automatically in images 24. In other embodiments, the cell nuclei are selected automatically from intact and in-focus cell nuclei identified in images 24. In some embodiments, the selection of cell
10 nuclei from among the nuclei discernible in images 24 is assisted by a person who may specify by way of a suitable user interface one or more of:

- specific nuclei to be selected;
- regions within an image 24 from within which some or all of the nuclei ought to be selected.

15 In some embodiments, one or more regions of interest in image 24 may be assisted by location of the corresponding diagnostic area, as determined by a trained pathologist, in an adjacent serial section that has been stained for histopathology.

[0038] The nuclear morphology features may include features relating to one or more
20 of:

- the shape of the cell nucleus;
- the size of the cell nucleus; and,
- the texture of DNA distribution within the cell nucleus.

Some examples of shape-related nuclear morphometric features are:

- 25
- sphericity;
 - harmonics of the nuclear boundary (the *N*th harmonic of the nuclear boundary indicates the degree to which the shape of the nuclear boundary can be explained by *N* lobes;
 - other measures of smoothness of the nuclear boundary

30 Some examples of size-related nuclear morphometric features are:

- nucleus area;
- maximum linear dimension of nucleus;

Some examples of texture-related nuclear morphometric features are:

- 35
- Density;
 - Discrete texture,
 - Markovian texture,
 - Non-Markovian texture,
 - Fractal texture (e.g. fractal area - the relative spatial distributions of high and low optical density variations within the nucleus);

- OD skewedness - a measure of whether the nucleus is dark with light areas or light with dark areas;
 - Run-length texture (e.g. long run - a measure of how far across the nucleus one can travel before encountering a significant change in optical density).
- 5 In some embodiments, the nuclear morphometric features include at least one or more texture-related features. In some embodiments, the nuclear morphometric features include at least a size-related feature and a texture related feature. In some embodiments, the nuclear morphometric features include at least a shape-related feature and a texture related feature. In a prototype embodiment the values for the features
- 10 were calculated using Getafics™ tissue section analysis software. Typically the nuclear morphometric features include 4 or more features.

[0039] Methods for measuring various nuclear morphological features of cells are known to those of skill in the field automated analysis of images of tissues. For

15 example, some such methods are described in:

- Doudkine, A., et al.. *The comparison of single nuclei analysis and histometric texture feature analysis for detection of MAC in breast and bronchial biopsies* Proc. 8th Int'l Symp. on Diagnostic Quantitative Pathology, Amsterdam, The Netherlands, Sept. 14-16, p. 171 (1994);
 - 20 • Maticic, J.P., et al. *The use of nuclear texture features for the analysis of 989 biopsy proven squamous atypia* 24th European Congress of Cytology, Ljubljana, Slovenia, September 21-24 (1997);
 - Poulin, N., al. *Histometric texture analysis of DNA in thin sections from breast biopsies: Application to the detection of malignancy associated changes in carcinoma in situ* *Analyt. Quant. Cytol. Histol.* 17:291-299 (1995)
- 25

[0040] In block 28 a score is computed for each cell nucleus based upon the values of the nuclear morphometric features calculated for that nucleus. The score may be computed, for example, by computing a sum or a weighted sum of the values of the

30 nuclear morphometric features for the nucleus.

[0041] In block 30 the scores for individual nuclei are combined to yield an index, referred to herein as a morphological index or "MI". The value for the MI is indicative of one or more of:

- 35
- the prospects that the lesion will develop into cancer;
 - the prospects that cancer of the lesion will recur after treatment;
 - the prospects that the patient will survive.

The MI is applied in block 31 by, for example, one or more of:

- displaying the MI to a clinician or other user;

- recording the MI;
- generating a report including the MI;
- comparing the MI to a threshold;
- establishing a prognosis based on the MI;
- 5 • scheduling an appointment for treatment or other follow up based on the MI;
- scheduling further tests of the tissue in response to the MI satisfying some criteria;
- etc.

10 [0042] In some embodiments, the MI is calculated by a method 33 as shown in Figure 1A. In block 34 the scores for the different cell nuclei scored in block 28 are sorted into bins 35 according to their scores. In some embodiments there are 4 or more or 5 or more bins. For example, there may be between 4 and 20 bins in some embodiments. In an example embodiment there are 10 bins (identified as 35A to 35J in Figure 1A).

15 [0043] The number of nuclei that have been associated with each bin 35 is determined in block 36. The MI is computed in block 38. In a preferred embodiment, the MI is computed as a function of the numbers of nuclei having scores falling within the range for each bin. For example, the MI may be computed as a weighted sum of the number
20 of nuclei having scores belonging to each bin (or the percentage of the nuclei having scores in each bin) .

[0044] Where the score for each nucleus has been computed in such a manner that the score tends to be higher for cancerous (or, more generally, abnormal cells) and the
25 score tends to be lower for normal cells then the higher-valued bins may be weighted significantly more heavily than the lower-valued bins. This tends to produce an MI that discriminates well between normal cells and cells that are cancerous or are disposed to develop cancer.

30 [0045] In a prototype embodiment, the weights used to compute the MI for the 10 bins 35A to 35J are shown in Table 1. It can be observed that the weights increase linearly with bin number in this embodiment. In other embodiments, the weights may increase faster than linearly with bin number.

Table 1 - Weights for Bins			
Bin	Weight	Bin	Weight
35A	1	35F	6
35B	2	35G	7
35C	3	35H	8
35D	4	35I	9
35E	5	35J	10

[0046] The 10 regions or bins (35A to 35J) represent classes of nuclei with different morphometric characteristics according to their discriminant function scores. For example, bin 35A may cover a range of discriminant score values such that normal-like cells are associated with bin 35A while bin 35J may cover a range of discriminant score values such that cancer-like cells are associated with bin 35J. Bins 35B to 35I cover intermediate ranges of discriminant score. The range of discriminant scores associated with each bin 35 may be set with reference to the scores of nuclei from samples of "normal" tissues and samples of tissues affected by cancer, as described below. The range of discriminant scores for bin 35A may be set so that the scores for normal nuclei tend to fall into bin 35A. The range of scores for bin 35J may be set so that the scores for at least a given percentage of the nuclei from the tissues affected by cancer fall in bin 35J.

[0047] In the prototype embodiment, the MI is calculated according to:

$$MI = (1 \times A) + (2 \times B) + (3 \times C) + (4 \times D) + (5 \times E) + (6 \times F) + (7 \times G) + (8 \times H) + (9 \times I) + (10 \times J) \quad (1)$$

where: the letters A through J represent the percentage of nuclei in bins 35A to 35J respectively. An MI calculated according to Equation (1) thus represents the weighted sum of the 10 bins with the numbers (1-10) representing the weight given to the percentage of nuclei in each of the 10 bins. More weight is assigned to the bins associated with nuclei having more cancer-like features. In the prototype embodiment, the MI has possible values in the range of 1 to 10.

[0048] Figure 1B illustrates a method 40 that may be used for:

- generating the function used in computing the scores in block 28;
- selecting nuclear morphological features on which the scores are based; and,
- determining the ranges of scores to be allocated to bins 35.

[0049] In block 41A images of a number of cell nuclei in normal tissue are obtained. The images may be images of sectioned biopsies obtained as described above, for example. Cell nuclei may be selected from each image as described above. The images
5 of the normal cell nuclei may be described as a “first training set”.

[0050] In block 41B images of a number of cell nuclei in tissue known to be cancerous are obtained. The images may be images of sectioned biopsies obtained as described above, for example. Cell nuclei may be selected from each image as described above.
10 The images of the normal cell nuclei may be described as a “second training set”.

[0051] In block 42, the values of each of a set of nuclear morphological features are computed for each of the nuclei in each of the first and second training sets. In a prototype embodiment, the set of nuclear morphological features included 125 different
15 features. The set of features includes one or more texture-based features in some embodiments. The majority of the features are texture-based features in some embodiments. In preferred embodiments, the set of features includes one or more texture-based features, one or more size-based features and one or more shape-based features.

20 [0052] In the prototype embodiment the first training set comprised 5261 cell nuclei selected from 30 biopsies of normal tissue and the second training set comprised 5506 cell nuclei selected from 29 biopsies of tissue known to be affected by oral squamous cell carcinoma.

25 [0053] In block 44 forward linear discriminant analysis is performed to identify a subset of the features that can be used as arguments in a discriminant function that will separate the normal cells from the cancer-like cells. This represents a cell-by-cell classification procedure that assigns a discriminant score to each nucleus (e.g. 50 to
30 100 nuclei per biopsy) analysed for each OPL biopsy specimen indicating its likely position along the pathway of progression to cancer. Table 2 shows nuclear morphometric features used for forward discriminant analysis in the prototype embodiment.

	Table 2 - <i>Nuclear Morphometric Features Used for Forward Linear Discriminant Analysis for Prototype Embodiment Assessing Oral Preneoplastic Lesions</i>		
5	Function Name	Coefficient	Description
	Fractal_area1	-0.00016	a measure of the relative spatial distribution of high and low optical density variations in the nucleus – this measure is related to heterochromatin/euchromatin organization
	OD_Skewness	0.59	3 rd moment of the nuclear optical density distribution
	Long90_Run	0.1484	run-length features – this provides a measure of chromatin texture
	Max_radius	-0.1	maximum radius of the nucleus
10	Harm03_fft	-0.13	shape – measure of irregularity of the nuclear boundary – in this case the degree to which the boundary can be represented by three lobes.

The score for a nucleus may be computed by multiplying the value of each nuclear morphometric feature for the nucleus by the corresponding coefficient and then summing the results. For example, the score for the i^{th} nucleus may be expressed as:

$$Score_i = \sum_{j=1}^{j=N} C_j F_{ij} \quad (2)$$

where: $Score_i$ is the score for the i^{th} nucleus; C_j is the coefficient for the j^{th} nuclear morphometric feature; N is the number of nuclear morphometric features; and, F_{ij} is the value for the j^{th} nuclear morphometric feature for the i^{th} nucleus.

20 [0054] In block 46 the ranges of scores to be associated with each bin 35 is determined so that most normal nuclei from the first training set tend to have scores that fall in bin 35A and a desired fraction of nuclei from the second training set have scores that fall into the highest bin (in this prototype case, bin 35J).

25 [0055] Figure 2 shows distributions of the discriminant scores for nuclei in the first training set (identified by reference numeral 50) and the second training set (identified by reference numeral 52). Also shown in Figure 2 are the discriminant score ranges corresponding to bins 35A to 35J in the prototype embodiment. It can be seen that

distribution **50** for the first training set has the form of a normal distribution while distribution **52** for the second training set has a much broader range of discriminant score values (including a significant number of higher discriminant score values) and also has a shape different from that of a normal distribution. The methods described
5 above are sensitive to the different shapes of distributions **50** and **52**

[0056] One surprising result is that a system can be trained using only normal tissue samples and tissue samples affected by cancer and the system can produce MI values that provide quantitative indications of:

- 10
- the risk that a lesion will progress to cancer (i.e. the risk of malignant progression);
 - the overall disease prognosis (survival) of the subject;
 - the likely response of a lesion to treatment; and/or
 - the likelihood that cancer will occur after treatment.

15 An advantage of the embodiment described herein is that the MI may be determined on a lesion-by-lesion basis. This, in turn, can provide one or more of the indications above on a lesion-by-lesion basis. As mentioned above, pathological grading does not itself provide a good indication of the outcome of lesions, especially low-grade lesions such as OPLs. It is not obvious that MI values for different lesions or different patients
20 could have a good correlation to outcomes for those lesions or patients.

[0057] A MI determined as described herein may be used on its own or in combination with other diagnostic methods. For example, a MI may be used together with results of: one or more of:

- 25
- loss of heterozygosity (LOH);
 - histopathological grading;
 - toluidine blue staining;
 - molecular markers; and
 - various imaging techniques;

30 to provide a prognosis for a subject. In some embodiments, one or more of the additional diagnostic methods is invoked automatically in response to a MI value exceeding a threshold value.

[0058] The value of a MI can be used by a clinician as a basis for determining whether
35 immediate intervention (e.g. local surgery, radiation) is advised or whether watchful waiting is an acceptable alternative to immediate treatment. In some embodiments of

the invention, a MI is determined for a patient (or for one or more lesions of a patient). The value of the MI is used as a basis for determining whether to schedule the patient for further tests or treatment. In some embodiments, the MI is automatically compared to a threshold value. Further testing of the patient (or of one or more tissue samples
5 from the patient) or treatment of the patient may be scheduled automatically if the MI reaches or exceeds the threshold value.

[0059] In the embodiments described above, a high value for the MI correlates to a greater risk of malignant progression and a reduced long-term survival rate.
10 Conversely, a lower value for the MI correlates to a smaller risk of malignant progression and an increased long-term survival rate.

[0060] The prototype embodiment described herein has a number of advantages over current histopathological grading methods performed by skilled humans. Apart from
15 reduced reliance on human resources, which can be important, these advantages include:

- Computation of a MI is not subjective;
- The MI provides a quantitative measure;
- The MI can take into account relationships among different high-risk
20 parameters; and
- The MI can be based upon a wider range of tissue phenotypic changes and perform analysis based upon finer details than can a person inspecting an image of a tissue sample visually. For example, an increased chromatin level is known to be a high-risk feature. Traditional pathology judges chromatin level by one
25 criterion, hyperchromatism. By contrast, a MI may be based upon multiple chromatin features or nuclear phenotypes including measures of features such as whether the increased DNA is distributed around the edge of the nucleus or clustered in the centre; whether the nucleus is dark with light areas or light with dark areas and whether the increased chromatin is evenly distributed
30 (euchromatin) or clumped locally (heterochromatin).

[0061] In some embodiments, a MI is based upon factors in addition to nuclear morphological features. For example, a MI or the discriminant scores on which the MI is based may be based in part upon one or more of:

- 35 • cell location (e.g. a layer of tissue in which the cell is found) ;

- cell organization;

in addition to one or more nuclear morphological features.

5 [0062] Figure 3 is a block diagram of apparatus **60** according to an embodiment of the invention. Apparatus **60** comprises a high-resolution imaging subsystem **62** that acquires digital images **64**. A data processor **66** executes software instructions **68** in a memory **67** accessible to data processor **66**. Software instructions **68** cause data processor **66** to identify cell nuclei in images **24**. Data processor **66** is connected to a display **69** and a user interface **70**. Software instructions **68** optionally cause data
10 processor **66** to display an image **64** on display **69** and to receive input from a human user by way of interface **70** that directs data processor **66** to select certain cell nuclei or cell nuclei from certain areas within image **64**.

15 [0063] On each selected cell nucleus, the software instructions cause data processor **66** to compute values for a number of morphology features of the cell nuclei. In the illustrated embodiment, the features are calculated in accordance with software subroutines **73A** to **73F** in memory **67**. The software instructions **68** cause data processor **66** to compute a discriminant score for each cell nucleus based upon a set of stored coefficients **74** and the values computed for the morphological features. Software
20 instructions **68** then cause data processor **66** to compute a MI based upon the discriminant scores.

[0064] In some embodiments, processor **66** performs one or more of:

- displaying the MI on display **69**;
- 25 • printing prints a report that includes the MI on a hard copy output device **75**;
- storing the MI in memory **67**; and,
- taking a conditional action based upon a value of the MI.

The conditional action may include, for example, one or more of:

- flagging a file for urgent attention;
- 30 • generating an electronic message indicating that follow up is required;
- scheduling a patient associated with image **24** (or a tissue sample from the patient) for further tests;
- scheduling an appointment for a patient associated with image **24** for treatment or other follow up; and,

- the like.

Scheduling may be performed by a computer-based scheduling system. The scheduling system may receive input identifying a patient or tissue sample that requires scheduling and may create a job ticket requesting further testing or the like. The scheduling system
5 may maintain records indicating the availability of resources (such as physicians, treatment or diagnostic apparatus, or the like) and schedule the patient for an appointment at a time that a required resource is available. In some embodiments, a priority for an appointment may be based upon a value for the MI such that MI values indicating higher risk of progression to cancer would be allocated appointments with
10 higher priority than for lower MI values.

[0065] Image processing and calculation is described above as being performed in a data processor operating under software control. In alternative embodiments of the invention some or all of the image-processing, or computation steps, or both may be
15 performed in specialized image processing or calculating hardware which may include, for example, logic circuits, image processing circuits, suitably-configured field-programmable gate arrays, or the like.

[0066] Figure 3A shows apparatus **80** according to another embodiment of the
20 invention. Apparatus **80** acts on images **82**. A nucleus identification means **84** identifies cell nuclei in images **82**. A feature value computation means **86** computes values for a plurality of morphometric features of a cell nucleus identified by nucleus identification means **84**. The morphometric features may include shape features, size features, texture features or mixtures of shape features, size features and texture features.
25 Apparatus **80** has an index calculation means **88** which computes a value for an index characterizing the tissue from the values of the morphometric features calculated by the feature value computation means **86**. Apparatus **80** also includes output means **90** for recording, printing, or otherwise exploiting index values computed by the index calculation means **88**.

30

[0067] Any of the blocks in Figure 3A may be implemented in software executing on one or more data processors, in hardware, or in some combination thereof.

Example

[0068] The inventors have conducted a study in which they have correlated nuclear phenotype changes as measured by an MI determined as described above with: pathological grading, molecular changes and outcome of OPLs. The results of the study
5 showed that the Computer Imaging-detected nuclear phenotype changes correlated strongly with histology grading, genetic damage (LOH) and risk for cancer progression.

[0069] MI scores were based upon the nuclear morphometric features and coefficients
10 given in Table 2 for tissue samples from a number of OPLs in a number of patients. 94% of 'normal-like' cells and 77% of the 'cancer-like' cells were correctly classified using these 5 discriminant features. These features were used to generate the cell-by-cell discriminant scores which were amalgamated across all of the cells in each sample to generate the MI for each sample. The samples were divided into a training
15 set, a study set and a test set.

[0070] The samples included 83 well-characterized OPLs (a single biopsy per patient) with histological diagnoses of hyperplasia, mild or moderate dysplasia. The lesions were from patients without a prior history of head and neck cancer. Twenty-nine had
20 progressed to cancer and the remaining had not. A prerequisite for inclusion in the study was the availability of an unstained slide that was serial to slides used for previous LOH analysis. Forty-four cases qualified, including 15 progressing (6 hyperplasia, 5 mild dysplasia, and 4 moderate dysplasia) and 29 non-progressing cases (15 hyperplasia, 8 mild dysplasia, and 6 moderate dysplasia). There was no difference
25 between the progressing and the non-progressing lesions in terms of gender, age distribution, and smoking history (all $P > 0.05$). On average, the non-progressing cases had been monitored for over twice the duration of progressing cases (96 months vs. 37 months) to ensure that progression did not occur.

[0071] A set of samples from the ends of the phenotypic spectra of oral tissue in
30 neoplastic development were selected as the training set. The training set included 30 oral biopsies (e.g., amalgam tattoo or melanotic macule) with areas of relatively normal oral mucosa and 29 SCCs.

[0072] All samples were formalin fixed and paraffin embedded. The histological
35 diagnoses of the samples were reviewed by two oral pathologists. Serial sections,

4- μ m in thickness, were cut from each sample and placed on glass slides, one stained with Hematoxylin and Eosin (H & E), and the other with Feulgen-Thionin.

Representative areas of the histological diagnosis for each sample were circled on the H & E slide by an oral pathologist, and corresponding areas on the

5 Feulgen-Thionin-stained slide were examined by an automated imaging system.

[0073] The MI was correlated with each of: results of histopathological grading (training, study and test sets); molecular pattern (LOH results - for the study set); and outcome (from the study set).

10

[0074] Figures 4 and 4A show correlations between the MI and histopathology grade. In Figures 4 and 4A, the error bars represent 5th and 95th percentiles, the boxes represent the central 50th percentile the horizontal bar represents the histopathology grade median, and the dark square represents the median of MI values. In Figure 4A,
15 N stands for "normal"; Hp stands for "hyperplasia"; LG Dys stands for "low-grade dysplasia"; HG Dys stands for "high-grade Dysplasia"; and SCC stands for squamous cell carcinoma".

[0075] Table 3 shows association between MI values and molecular risk patterns as
20 evaluated by loss of heterozygosity (LOH) or allelic loss (see Rosin, M. P., et al. *Use of allelic loss to predict malignant risk for low-grade oral epithelial dysplasia*. Clin Cancer Res, 6: 357-362, 2000 for a description of LOH analysis. MI values were consistently higher for samples having LOH at all 7 chromosome regions examined. This increase was significant for 3p (P = 0.03), 4q (P = 0.002), and 9p (P = 0.001).
25 Elevated MI values were strongly associated with the presence of high-risk LOH patterns: multiple losses (P = 0.0001), and LOH at 3p &/or 9p plus LOH at any of the arms 4q, 8p, 11q, 13q and 17p (P = 0.0001).

Table 3 - Comparison of MI to LOH Results					
	With LOH		Without LOH		P-value
	# of Cases	MI	# of Cases	MI	
3p	10	5.4 $\pm$ 1.6	28	4.3 $\pm$ 1.1	0.03
4q	5	6.0 $\pm$ 1.3	36	4.2 $\pm$ 1.1	0.002
8p	8	5.1 $\pm$ 1.5	33	4.4 $\pm$ 1.2	0.18
9p	17	5.3 $\pm$ 1.6	26	4.0 $\pm$ 0.7	0.001
11q	7	5.3 $\pm$ 1.7	36	4.4 $\pm$ 1.1	0.06
13q	3	5.0 $\pm$ 1.3	37	4.4 $\pm$ 1.12	0.38
17p	8	5.1 $\pm$ 1.9	35	4.4 $\pm$ 1.1	0.14
Any loss	23	5.0 $\pm$ 1.5	20	3.9 $\pm$ 0.5	0.004
At least 1 LOH	15	5.5 $\pm$ 1.6	28	4.0 $\pm$ 0.7	0.0001
At least 2 LOH	11	5.6 $\pm$ 1.7	32	4.2 $\pm$ 0.9	0.0009
High Risk*	15	5.5 $\pm$ 1.6	28	4.0 $\pm$ 0.7	0.0001

*LOH at 3p and/or 9p plus at least 1 other arm (4q, 8p, 11q, 13q or 17p)

[0076] Figure 5 shows MI values for cases exhibiting hyperplasia and low-grade dysplasia that have been classified according to whether or not they progress to cancer. Progressing lesions showed significantly higher MI values than non-progressing lesions (mean MI 5.72 $\pm$ 1.36 for progressing lesions vs. 3.94 $\pm$ 0.66 for non-progressing lesions; P < 0.0001). The difference was still statistically significant even when the hyperplasia and dysplasia were examined separately (smaller sample sets): mean MI 4.74 $\pm$ 0.70 for progressing hyperplasia vs. 3.85 $\pm$ 0.50 for non-progressing

hyperplasia ($P = 0.004$); and mean NPS 6.38 ± 1.31 for progressing low-grade dysplasia vs. 4.03 ± 0.81 for non-progressing low-grade dysplasia ($P < 0.0001$).

5 [0077] These experimental results show that an appropriately-defined MI has values that differ sufficiently depending upon the risk that a lesion will progress to cancer. This makes it possible to create a bright-line test such that lesions having MI below a threshold value can be considered low-risk while lesions having MI values above the threshold value can be considered higher-risk. For the MI used in the prototype embodiment being used to study OPLs, a cutoff value of 4.5 was found to provide good
10 separation between non-progressing lesions and the progressing lesions. In the training set, which consisted of biopsies from lesions graded as either normal or SCC, only 5 of 30 (17%) normal samples had a MI value ≥ 4.5 (high -MI) and 28 of the 29 (97%) of SCC had MI values ≥ 4.5 .

15 [0078] When the study set samples (hyperplasia and low-grade dysplasia) were categorized using this threshold value, 17 of the 44 test cases had a high-MI. Molecularly, 16 of these 17 cases exhibited LOH. Of these 16 cases, 11 (69%) showed LOH at 3p &/or 9p, a loss believed to be essential although not sufficient for cancer progression. In contrast, only 7 of the 27 (26%) cases with MI < 4.5 (low-MI) showed
20 such a loss ($P = 0.001$).

[0079] Prognostically, 13 of the 17 (76%) cases with high-MI progressed to invasive SCC. In contrast, only 2 of the 27 (7%) cases with low-MI progressed to cancer in the study period ($P < 0.0001$). In other words, 86% of non-progressing cases (25/29) and
25 86% of progressing cases (13/15) were correctly classified by comparing MI values to a threshold.

[0080] Time-to-progression curves were plotted as function of the MI classification into one of the two groups (high-MI vs. low-MI). A significant difference was observed
30 between the low-MI group and the high-MI group ($P < 0.00015$). Time to progression was significantly different for these 2 groups. The proportion of lesions progressing to cancer at 5 years was 71% (12/17) for lesions with high-MI and only 22% for lesions with low-MI. There was a 10-fold ($RR = 10.3$ [CI: 2.86-59.8]) increase in the relative risk of progression to cancer for oral lesions with a high-MI in
35 comparison to those with a low-MI.

[0081] These results demonstrate the utility of MI values for predicting cancer outcomes (as distinct from grading lesions). The outcomes may comprise, for example, one or more of:

- the risk that a lesion will progress to cancer within some defined period if the lesion is not treated;
- the likelihood that a lesion will not progress to cancer within some defined period;
- the risk that a particular lesion will progress to cancer (or that cancer will recur) after having been treated with a particular treatment or set of treatments; and,
- the likely response of a particular lesion to a specified treatment.

These outcomes may be predicted for low-or higher-grade lesions.

[0082] MI values can be used to grade lesions. The MI values tend to increase with increasing severity of histological grading ($P < 0.0001$). MI tends to differ significantly between non-dysplastic lesions and dysplastic lesions; and between dysplastic lesions and invasive cancer. As shown in Figure 4A, low-grade dysplasia shows the largest variability in MI values among the different histological grades. This indicates that lesions classified as low-grade dysplasia by standard histopathological grading include a mixture of lesions having different potentials for progressing.

20

[0083] Figure 6 shows the progression to cancer over time for a set of lesions having morphological index values above a threshold value and another set of lesions having morphological index values below the threshold.

[0084] Certain implementations of the invention comprise computer processors which execute software instructions which cause the processors to perform a method of the invention. For example, one or more processors in a computer system may implement the methods of the invention by executing software instructions in a program memory accessible to the processors. The invention may also be provided in the form of a program product. The program product may comprise any medium which carries a set of computer-readable signals comprising instructions which, when executed by a data processor, cause the data processor to execute a method of the invention. Program products according to the invention may be in any of a wide variety of forms. The program product may comprise, for example, media such as magnetic data storage media including floppy diskettes, hard disk drives, optical data storage media including

CD ROMs, DVDs, electronic data storage media including ROMs, flash RAM, or the like. The software instructions on the medium may optionally be compressed, encrypted, or both compressed and encrypted.

- 5 **[0085]** Where a component (e.g. a software module, processor, assembly, device, circuit, etc.) is referred to above, unless otherwise indicated, reference to that component (including a reference to a "means") should be interpreted as including as equivalents of that component any component which performs the function of the described component (i.e., that is functionally equivalent), including components which
10 are not structurally equivalent to the disclosed structure which performs the function in the illustrated exemplary embodiments of the invention.

[0086] While a number of exemplary aspects and embodiments have been discussed above, those of skill in the art will recognize certain modifications, permutations,
15 additions and sub-combinations thereof. By way of non-limiting example:

- the specific features used for generating discriminant scores or the functions applied to obtain values for those features may be varied;
- the number of features on which the discriminant scores are based may be varied;
- 20 • the number of nuclei for which features are measured for computation of a MI may be varied;
- the number of bins **35** may be varied;
- the weighting factors used in computation of the MI may be varied;
- 25 • individual values for nuclear morphological features or computations of discriminant scores in some embodiments may be performed manually as opposed to by the automated apparatus described above;
- alternative functions for computing discriminant scores or for computing the MI from the discriminant scores may be applied;
- 30 • methods and apparatus according to the invention may be applied to assessing a variety of tissues including oral tissues such as OPLs as described above, bronchial tissues, tissues which can host other epithelial-derived cancers such as the head and neck, cervix, ovary and lung;

It is therefore intended that the following appended claims and claims hereafter introduced are interpreted to include all such modifications, permutations, additions
35 and sub-combinations as are within their true spirit and scope.

WHAT IS CLAIMED IS:

1. A method for assessing tissue pathology comprising:
 - obtaining images of a plurality of cell nuclei in a tissue sample;
 - 5 computing values for a plurality of morphometric features of the plurality of cell nuclei, the morphometric features including one or more texture features; and,
 - from the values of the morphometric features, computing a value for an index characterizing tissue of the tissue sample; and,
 - 10 recording the index value.
2. A method according to claim 1 wherein computing the value for the index comprises:
 - computing discriminant scores for the plurality of nuclei based upon the
 - 15 values of the morphometric features for the nuclei;
 - classifying the nuclei into a bins of a series of bins based upon the discriminant scores for the nuclei; and,
 - computing the index value based upon the relative numbers of the nuclei classified in the bins.
 - 20
3. A method according to claim 2 wherein computing the index value based upon the relative numbers of the nuclei classified in the bins comprises computing a weighted sum of the fraction of the nuclei classified in each of the bins.
- 25 4. A method according to claim 3 wherein computing the weighted sum comprises multiplying the fraction of the nuclei classified in each of the bins by a weighting coefficient associated with the bin wherein the weighting coefficients increase linearly with bin number.
- 30 5. A method according to any one of claims 2 to 4 wherein classifying the nuclei comprises sorting the nuclei into at least 5 bins.

6. A method according to claim 5 comprising associating a range of values for the discriminant score with a first one of the bins that corresponds to nuclei of normal-like cells.
- 5 7. A method according to claim 5 or 6 comprising associating a range of values for the discriminant score with a last one of the bins that corresponds to nuclei of cancer-like cells.
8. A method according to any one of claims 1 to 7 wherein the texture features
10 comprise a fractal area.
9. A method according to any one of claims 1 to 8 wherein the texture features comprise a measure of a how far across the nucleus one can travel before encountering a significant change in optical density.
- 15 10. A method according to any one of claims 1 to 4 wherein the texture-related features comprise one or more features selected from the group consisting of: density; discrete texture; Markovian texture; non-Markovian texture; fractal texture; OD skewedness; and, run-length texture.
- 20 11. A method according to any one of claims 1 to 10 wherein the morphometric features include one or more features having values that measure sizes of the nuclei.
- 25 12. A method according to claim 11 wherein the one or more features having values that measure sizes of the nuclei include a feature having a value representing a maximum diameter of the nuclei.
- 30 13. A method according to any one of claims 1 to 12 wherein the morphometric features include one or more features having values that measure shapes of boundaries of the nuclei.

14. A method according to any one of claims 1 to 13 wherein computing the value for the index is based upon the values for the morphometric features for at least 35 nuclei.
- 5 15. A method according to any one of claims 1 to 14 comprising comparing the index to a threshold.
16. A method according to claim 15 comprising automatically scheduling a further test on the tissue based upon a result of the comparison.
- 10 17. A method according to claim 15 comprising automatically scheduling an appointment for a patient corresponding to the tissue based upon a result of the comparison.
- 15 18. A method according to claim 15 wherein automatically scheduling the appointment comprises setting a priority for the appointment according to a value of the index.
- 20 19. A method according to any one of claims 1 to 18 wherein obtaining the images comprises imaging thin sections of tissue with a digital microscope imaging system.
20. A method according to claim 19 wherein the thin sections are stained with a stain.
- 25 21. A method according to claim 20 wherein the stain is quantitative for DNA.
22. A method according to claim 21 wherein the stain comprises a Feulgen-Thionin stain.
- 30 23. A method according to claim 20 wherein the stain comprises a haematoxylin / eosin stain.

24. A method according to any one of claims 1 to 23 comprising computing the discriminant scores according to a discriminant function obtained by performing discriminant analysis on a first training set comprising normal-like cells and a second training set comprising cancer-like cells.
- 5
25. A method according to any one of claims 1 to 23 wherein the discriminant scores are based upon at least: a fractal area; a measure of skewedness, and a dimension of the nucleus.
- 10
26. A method according to claim 25 wherein a coefficient associated with the measure of skewedness is opposite in sign to coefficients associated with the fractal area and the dimension of the nucleus.
- 15
27. A method according to claim 1 wherein computing the value for the index is performed according to a function having the property that there exists a first range of values for the index for which there is a confidence in excess of 50% that the tissue belongs to a progressing phenotype if the index for the tissue is in the first range.
- 20
28. A method according to claim 27 wherein computing the value for the index is performed according to a function having the property that there exists a second range of values for the index that is non-overlapping with the first range for which there is a confidence in excess of 50% that the tissue belongs to a non-progressing phenotype if the index for the tissue is in the second range.
- 25
29. A method according to claim 1 wherein computing the value for the index is performed according to a function having the property that there exists a range of values for the index for which there is a confidence in excess of 50% that the tissue belongs to a non-progressing phenotype if the index for the tissue is in the range.
- 30
30. A method according to any one of claims 1 to 29 comprising expressing a likelihood of an outcome based upon the value of the index.

31. A method according to claim 30 wherein the outcome comprises a risk of progression to cancer.
- 5 32. A method according to claim 30 wherein the outcome comprises a likelihood of recurrence of cancer after a specified treatment.
33. A method according to claim 30 wherein the outcome comprises a likely response of a lesion to a specified treatment.
- 10 34. A method according to claim 30 wherein the outcome comprises a prognosis for a lesion.
35. A method according to any one of claims 1 to 34 wherein the tissue sample comprises a sample of tissue of an oral lesion.
- 15 36. A method according to claim 35 wherein the oral lesion comprises an oral preneoplastic lesion.
37. A method according to claim 35 wherein the oral lesion comprises a low-grade oral lesion.
- 20 38. A method according to claim 35 wherein the oral lesion comprises a high-grade oral lesion.
- 25 39. A method according to any one of claims 1 to 34 wherein the tissue sample comprises a sample of a tissue of a patient's head, neck, lung, cervix, or ovary.
- 30 40. A method according to any one of claims 1 to 39 comprising computing the value for the index for a plurality of tissue samples corresponding respectively to a plurality of lesions of a patient and prioritizing treatments of the plurality of lesions based upon the index values.

41. Apparatus for automatically assessing tissue pathology, the apparatus comprising:
- a data processing unit connected to obtain images of cell nuclei in tissue samples, the data processing unit comprising:
- 5 nucleus identification means for identifying the cell nuclei in the images;
- feature value computation means for computing values for a plurality of morphometric features of a cell nucleus identified by the nucleus identification means, the morphometric features including one or
- 10 more texture features; and,
- index calculation means for computing a value for an index characterizing the tissue from the values of the morphometric features calculated by the feature value computation means; and,
- means for recording index values computed by the index
- 15 calculation means.
42. Apparatus according to claim 41 comprising an appointment scheduling system configured to schedule appointments in cases where the index values meet specified criteria.
- 20
43. Apparatus according to claim 41 or 42 wherein the index calculation means comprises:
- means for computing discriminant scores for the plurality of nuclei based upon the values of the morphometric features for the nuclei;
- 25 means for classifying the nuclei into a bins of a series of bins based upon the discriminant scores for the nuclei; and,
- means for computing the index value based upon the relative numbers of the nuclei classified in the bins.
- 30 44. Apparatus according to claim 43 wherein the means for computing the index value is configured to compute a weighted sum of the fraction of the nuclei classified in each of the bins.

45. Apparatus according to claim 44 wherein computing the weighted sum comprises multiplying the fraction of the nuclei classified in each of the bins by a weighting coefficient associated with the bin wherein the weighting coefficients increase linearly with bin number and the apparatus comprises a memory storing the weighting coefficients.
46. Apparatus according to any one of claims 41 to 45 wherein the texture features comprise a fractal area.
47. Apparatus according to any one of claims 41 to 46 wherein the texture features comprise a measure of a how far across the nucleus one can travel before encountering a significant change in optical density.
48. Apparatus according to any one of claims 41 to 45 wherein the texture-related features comprise one or more features selected from the group consisting of: density; discrete texture; Markovian texture; non-Markovian texture; fractal texture; OD skewedness; and, run-length texture.
49. Apparatus according to any one of claims 41 to 48 wherein the morphometric features include one or more features having values that measure sizes of the nuclei.
50. Apparatus according to any one of claims 41 to 49 wherein the morphometric features include one or more features having values that measure shapes of boundaries of the nuclei.
51. A method for deriving an index indicative of a likelihood of an outcome for a tissue, the method comprising:
- obtaining images of a plurality of cell nuclei in sample of the tissue;
 - computing values for a plurality of morphometric features of the plurality of cell nuclei;
 - computing discriminant scores for the plurality of nuclei based upon the values of the morphometric features for the nuclei;

classifying the nuclei into a bins of a series of bins based upon the discriminant scores for the nuclei;

computing the index value based upon the relative numbers of the nuclei classified in the bins; and,

5 recording the index value.

52. A method according to claim 51 wherein computing the index value based upon the relative numbers of the nuclei classified in the bins comprises computing a weighted sum of the fraction of the nuclei classified in each of the bins.

10

53. A method according to claim 52 wherein computing the weighted sum comprises multiplying the fraction of the nuclei classified in each of the bins by a weighting coefficient associated with the bin wherein the weighting coefficients increase linearly with bin number.

15

54. A method according to any one of claims 51 to 53 wherein classifying the nuclei comprises sorting the nuclei into at least five bins.

55. A method according to claim 54 comprising associating a range of values for the discriminant score with a first one of the bins that corresponds to nuclei of normal-like cells.

20

56. A method according to any one of claims 53 to 55 comprising associating a range of values for the discriminant score with a last one of the bins that corresponds to nuclei of cancer-like cells.

25

57. A method according to any one of claims 51 to 56 comprising comparing the index to a threshold.

30 58. A method according to claim 57 comprising automatically scheduling a further test on the tissue based upon a result of the comparison.

59. A method according to claim 57 comprising automatically scheduling an appointment for a patient corresponding to the tissue based upon a result of the comparison.
- 5 60. A method according to any one of claims 51 to 59 wherein obtaining the images comprises imaging thin sections of tissue with a digital microscope imaging system.
- 10 61. A method according to claim 60 wherein the thin sections are stained with a stain that is quantitative for DNA.
62. A method according to claim 61 wherein the stain comprises a Feulgen-Thionin stain.
- 15 63. A method according to any one of claims 51 to 62 comprising taking a conditional action based upon a value of the index.
64. A method according to claim 63 wherein the conditional action comprises automatically flagging a file for urgent attention.
- 20 65. A method according to claim 63 wherein the conditional action comprises automatically ordering further tests on the sample of tissue.
- 25 66. A method according to claim 63 wherein the conditional action comprises automatically generating an electronic message indicating that follow up is required.
- 30 67. A method according to claim 63 wherein the conditional action comprises automatically scheduling an appointment for a patient associated with the tissue sample.

68. A method according to claim 67 wherein scheduling the appointment comprises automatically setting a priority for the appointment based upon the value of the index.
- 5 69. A method according to any one of claims 51 to 68 comprising computing the discriminant scores according to a discriminant function obtained by performing discriminant analysis on a first training set comprising normal-like cells and a second training set comprising cancer-like cells.
- 10 70. A method according to any one of claims 51 to 69 wherein the discriminant scores are based upon at least: a fractal area; a measure of skewedness, and a dimension of the nucleus.
- 15 71. A method according to claim 70 wherein a coefficient associated with the measure of skewedness is opposite in sign to coefficients associated with the fractal area and the dimension of the nucleus.
72. A method according to any one of claims 51 to 71 comprising associating an outcome with the value of the index.
- 20 73. A method according claim 72 wherein the outcome comprises a risk of progression to cancer.
- 25 74. A method according to claim 72 wherein the outcome comprises a likelihood of recurrence of cancer after a specified treatment.
75. A method according to claim 72 wherein the outcome comprises a likely response of a lesion to a specified treatment.
- 30 76. A method according to claim 72 wherein the outcome comprises a prognosis for a lesion.

77. A method according to any one of claims 51 to 76 comprising computing the value for the index for a plurality of tissue samples corresponding respectively to a plurality of lesions of a patient and prioritizing treatments of the plurality of lesions based upon the index values.
- 5
78. A method according to claim 77 wherein the plurality of lesions comprise oral lesions.
79. A method according to claim 78 wherein the oral lesions comprise oral preneoplastic lesions.
- 10
80. A method according to claim 78 wherein the lesions comprise low-grade lesions.
- 15
81. Apparatus for automatically assessing tissue pathology, the apparatus comprising:
- a data processing unit connected to obtain images of cell nuclei in tissue samples, the data processing unit comprising:
- nucleus identification means for identifying the cell nuclei in the images;
- 20
- feature value calculation means for computing values for a plurality of morphometric features of the plurality of cell nuclei;
- discriminant score computation means for computing discriminant scores for the plurality of nuclei based upon the values of the morphometric features for the nuclei;
- 25
- classification means for classifying the nuclei into a bins of a series of bins based upon the discriminant scores for the nuclei;
- index computation means for computing the index value based upon the relative numbers of the nuclei classified in the bins; and,
- means for recording the index value.
- 30
82. Apparatus according to claim 81 comprising an appointment scheduling system configured to schedule appointments in cases where the index values meet specified criteria.

83. Apparatus according to claim 81 or 82 wherein the index computation means comprises:
- means for computing discriminant scores for the plurality of nuclei based upon the values of the morphometric features for the nuclei;
- 5 means for classifying the nuclei into a bins of a series of bins based upon the discriminant scores for the nuclei; and,
- means for computing the index value based upon the relative numbers of the nuclei classified in the bins.
- 10 84. Apparatus according to claim 83 wherein the means for computing the index value is configured to compute a weighted sum of the fraction of the nuclei classified in each of the bins.
85. Apparatus according to claim 84 wherein computing the weighted sum
15 comprises multiplying the fraction of the nuclei classified in each of the bins by a weighting coefficient associated with the bin wherein the weighting coefficients increase linearly with bin number and the apparatus comprises a memory storing the weighting coefficients.
- 20 86. Apparatus according to any one of claims 81 to 85 wherein the texture features comprise a fractal area.
87. Apparatus according to any one of claims 81 to 86 wherein the texture features
25 comprise a measure of a how far across the nucleus one can travel before encountering a significant change in optical density.
88. Apparatus according to any one of claims 81 to 85 wherein the texture-related
30 features comprise one or more features selected from the group consisting of: density; discrete texture; Markovian texture; non-Markovian texture; fractal texture; OD skewedness; and, run-length texture.
89. Apparatus according to any one of claims 81 to 88 wherein the morphometric features include one or more features having values that measure sizes of the nuclei.

90. Apparatus according to any one of claims 81 to 89 wherein the morphometric features include one or more features having values that measure shapes of boundaries of the nuclei.
- 5 91. A method for assessing a likely outcome for an oral lesion, the method comprising:
- obtaining images of cell nuclei in a sample of tissue from the oral lesion;
- computing values for a plurality of morphometric features of the plurality of cell nuclei; and,
- 10 from the values of the morphometric features, computing a value for an index indicative of a likely outcome for the lesion; and,
- recording the index value.
- 15 92. A method according to claim 91 wherein the oral lesion comprises an oral preneoplastic lesion.
93. A method according to claim 91 or 92 comprising taking a conditional action based upon a value of the index.
- 20 94. A method according to claim 93 wherein the conditional action comprises automatically flagging a file for urgent attention.
95. A method according to claim 93 wherein the conditional action comprises automatically ordering further tests on the sample of tissue.
- 25 96. A method according to claim 93 wherein the conditional action comprises automatically generating an electronic message indicating that follow up is required.
- 30 97. A method according to claim 93 wherein the conditional action comprises automatically scheduling an appointment for a patient associated with the tissue sample.

98. A method according to claim 97 wherein scheduling the appointment comprises automatically setting a priority for the appointment based upon the value of the index.
- 5 99. A method according any one of claims 91 to 98 wherein the outcome comprises a risk of progression to cancer.
100. A method according any one of claims 91 to 98 wherein the outcome comprises a likelihood of recurrence of cancer after a specified treatment.
- 10
101. A method according any one of claims 91 to 98 wherein the outcome comprises a likely response of a lesion to a specified treatment.
102. A method according any one of claims 91 to 98 wherein the outcome comprises a prognosis for a lesion.
- 15
103. Apparatus comprising any new and inventive element, combination of elements or sub-combination of elements described herein.
- 20 104. Methods comprising any new and inventive steps, acts, combinations of steps and/or acts or sub-combinations of steps and/or acts described herein.

1/10

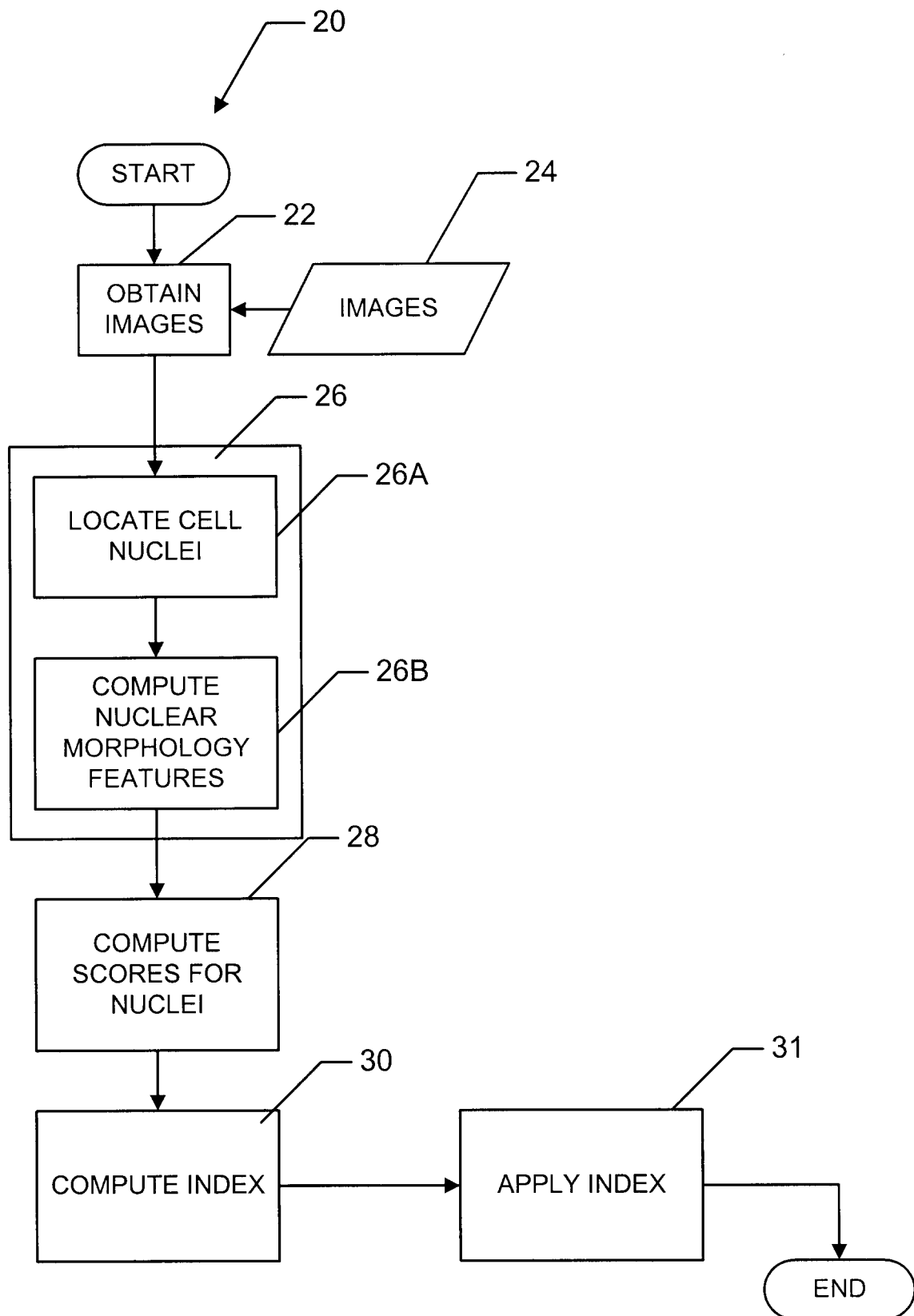


FIGURE 1

2/10

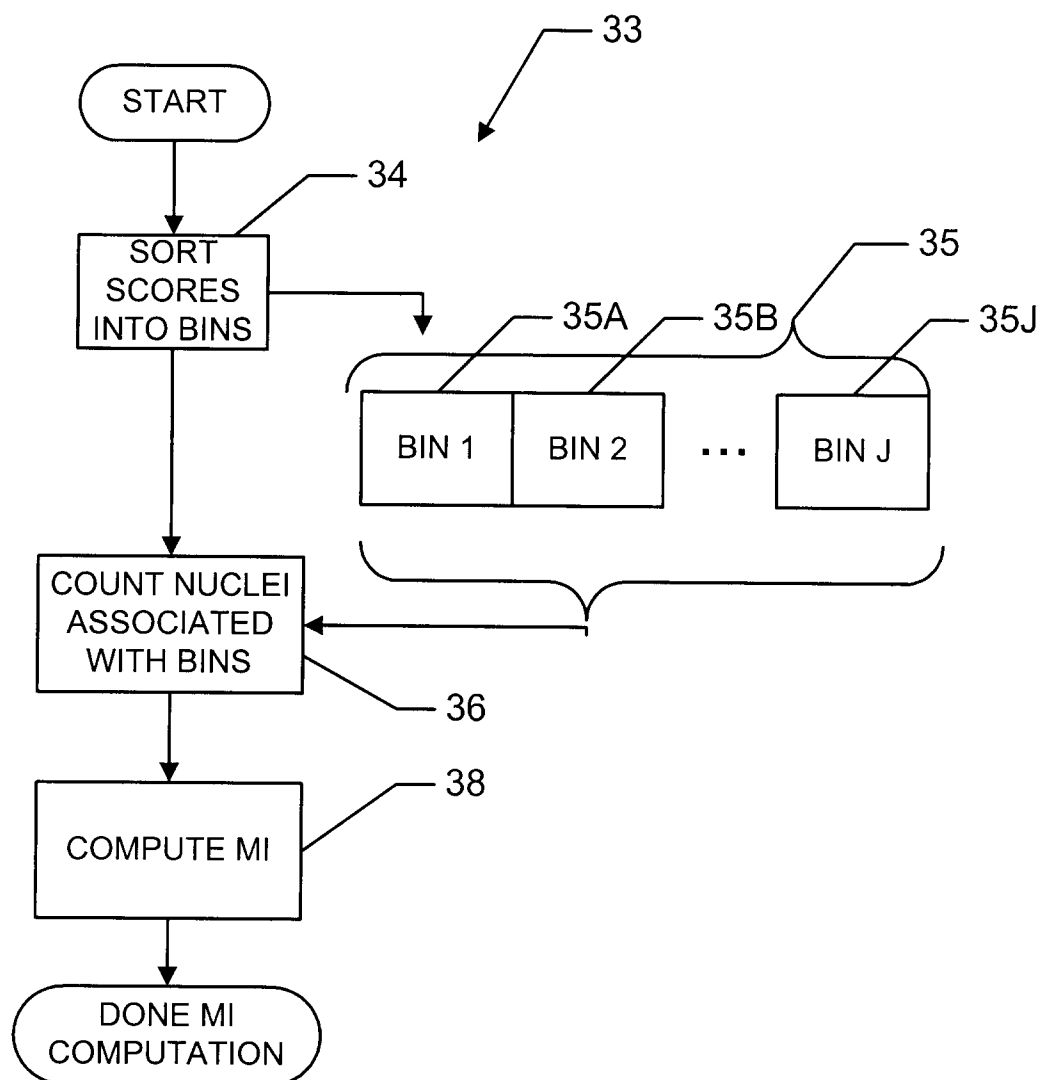


FIGURE 1A

3/10

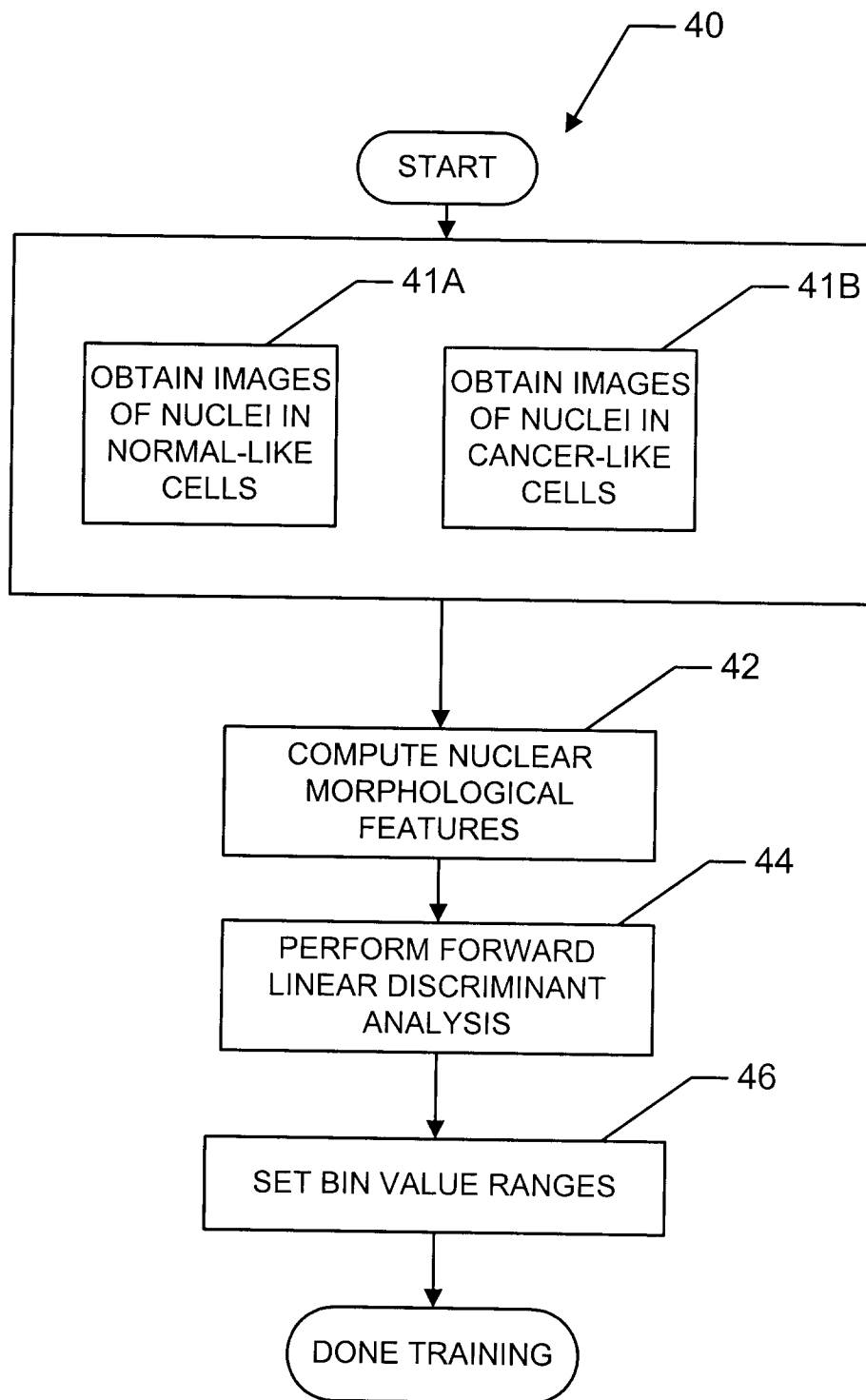


FIGURE 1B

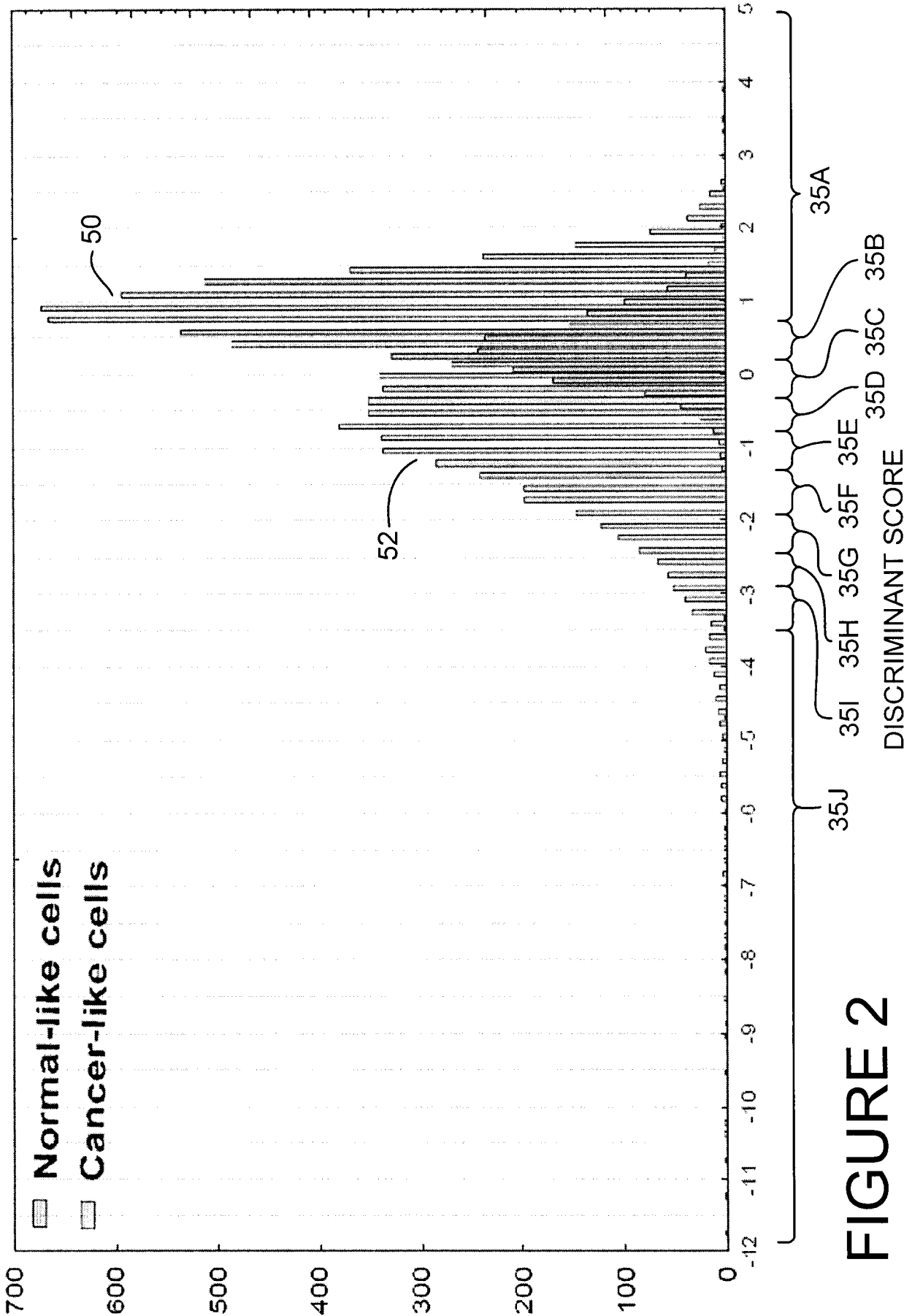


FIGURE 2

5/10

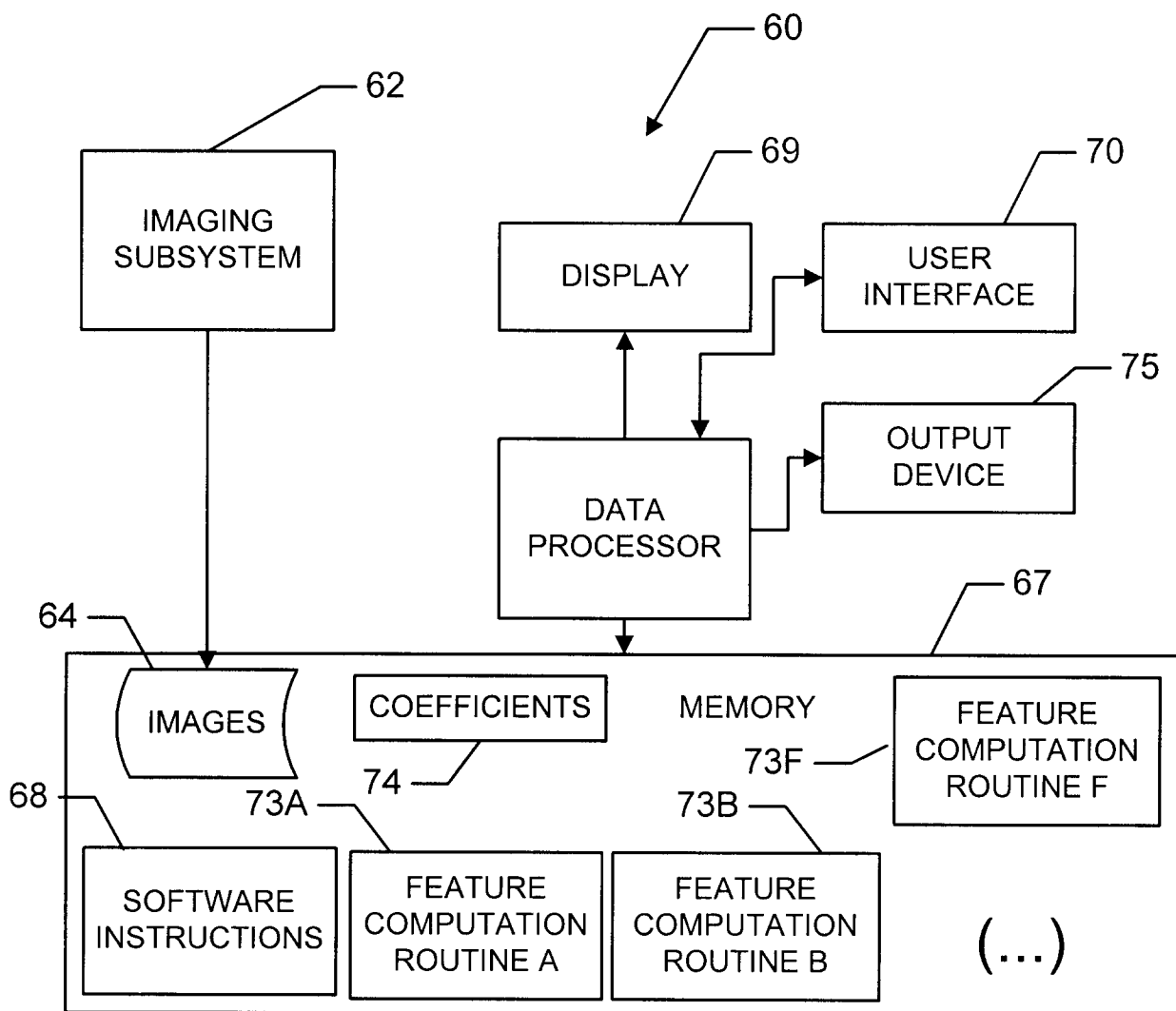


FIGURE 3

6/10

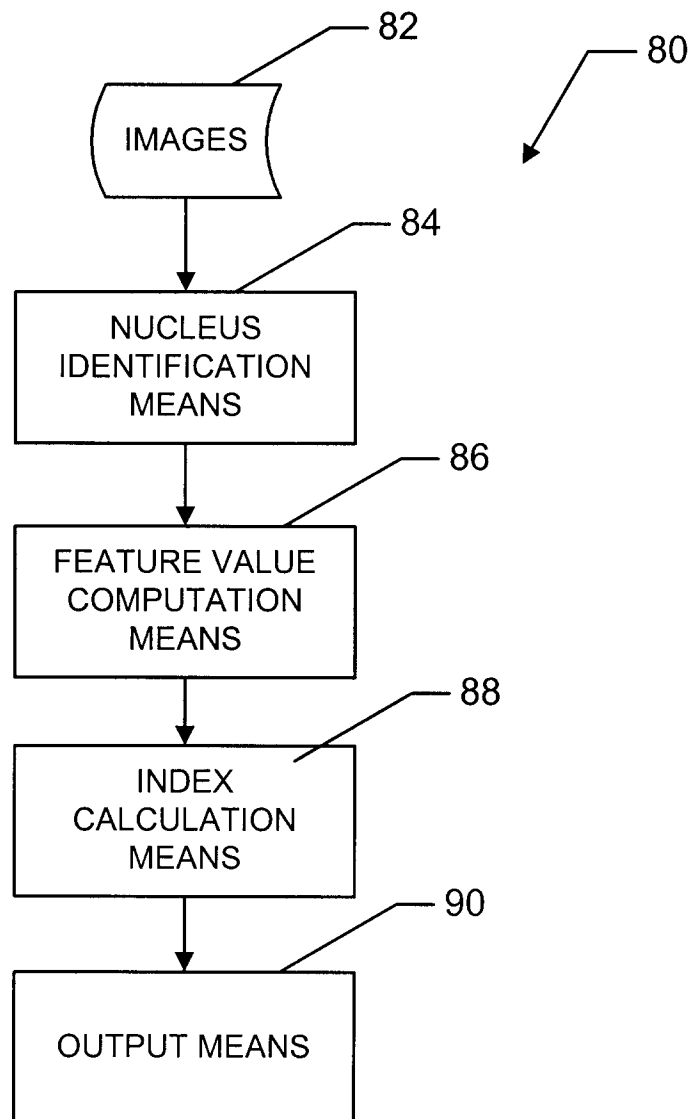


FIGURE 3A

7/10

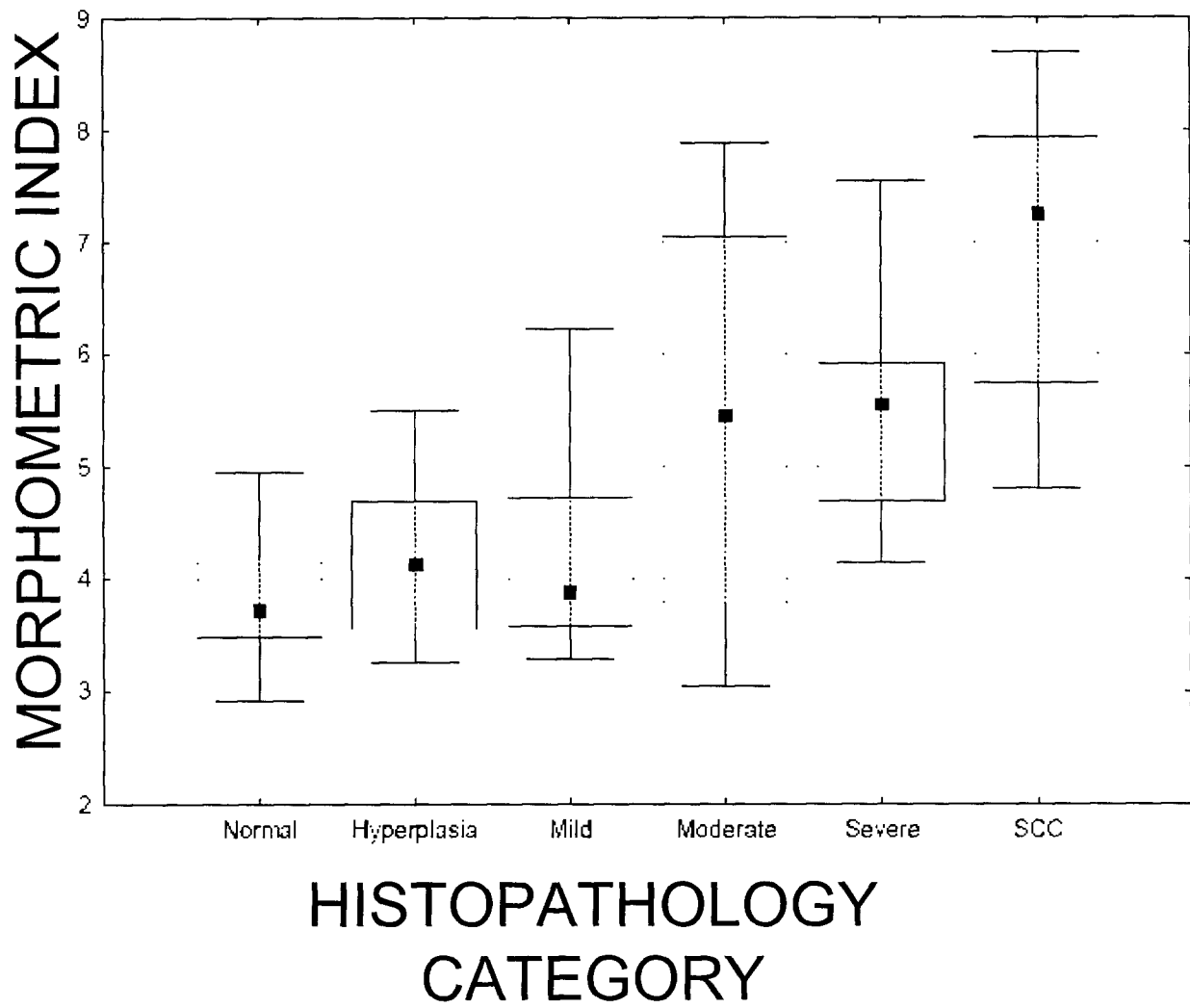


FIGURE 4

8/10

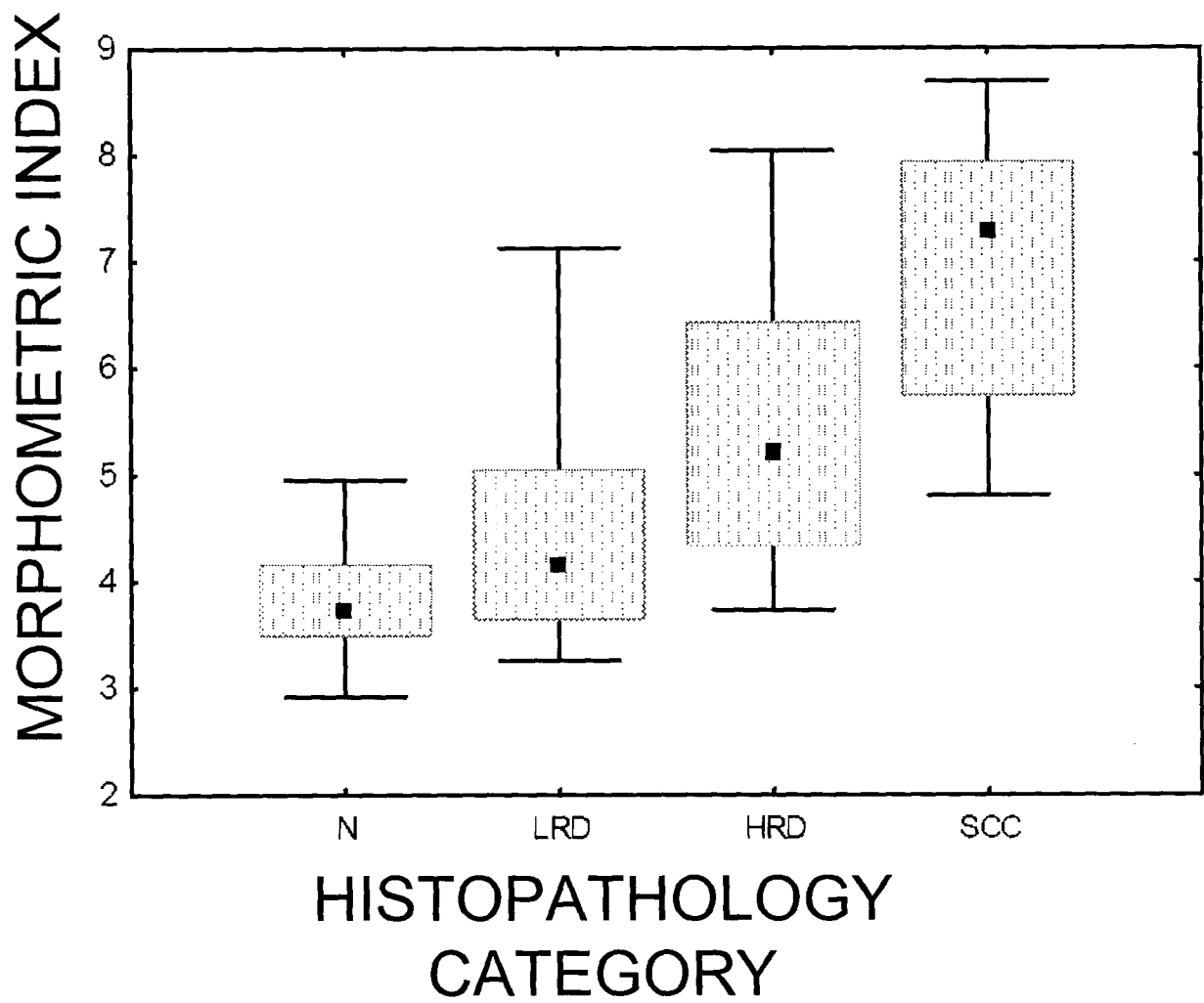


FIGURE 4A

9/10

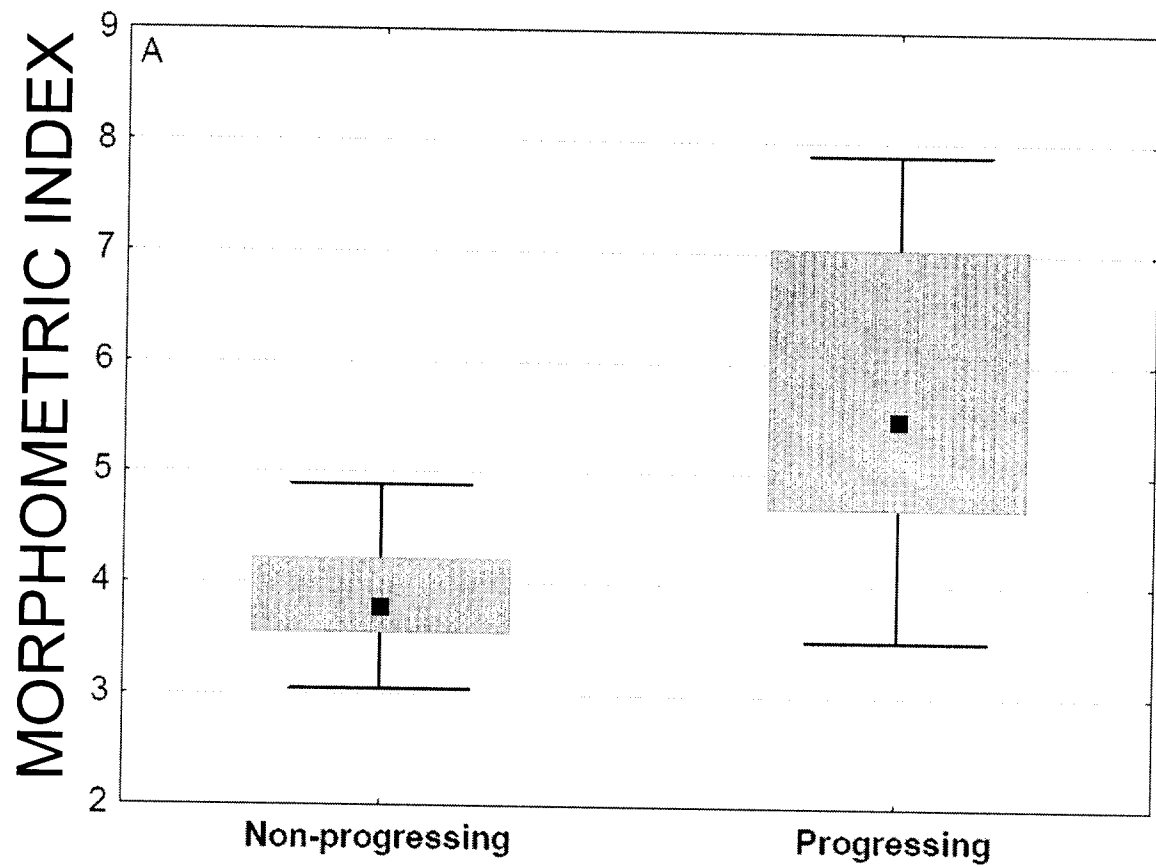
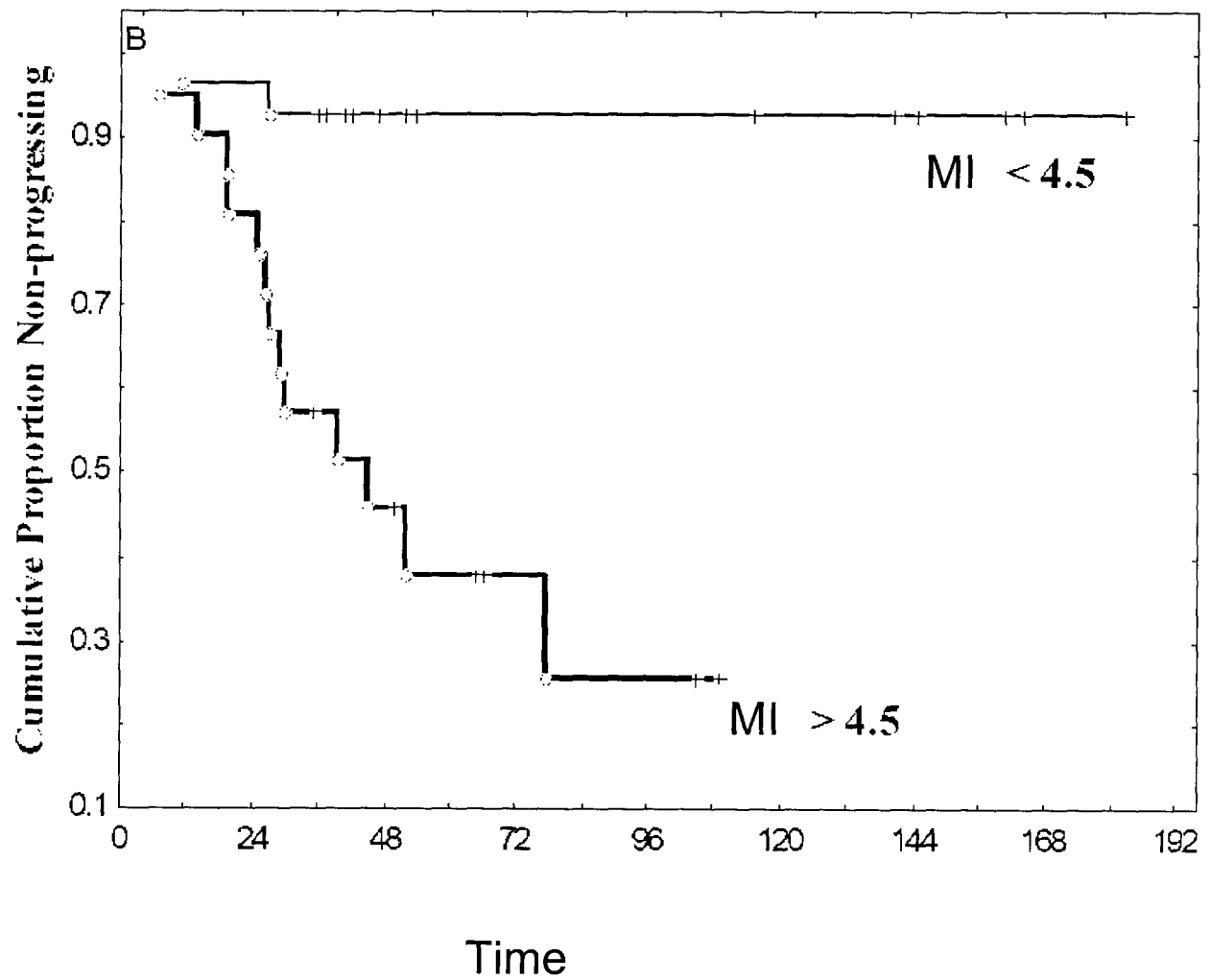


FIGURE 5

10/10

**FIGURE 6**

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CA2006/001933

A. CLASSIFICATION OF SUBJECT MATTER IPC: G01N 15/00 (2006.01) , C12M 1/34 (2006.01) , C12Q 1/00 (2006.01) , G01N 15/02 (2006.01) , G01N 21/84 (2006.01) , G01N 33/483 (2006.01) (more IPCs on the last page) According to International Patent Classification (IPC) or to both national classification and IPC																	
B. FIELDS SEARCHED Minimum documentation searched (classification system followed by classification symbols) G01N-15/00, C21M-1/34, C12Q-1/00, G01N-15/02, G01N-21/84, G01N-33/483, G06F-19/00, G06T-7/40 Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched Electronic database(s) consulted during the international search (name of database(s) and, where practicable, search terms used) USPTO WEST, PUBMED, IEEE: Keywords: morphological, weighted, cancer, index, indices, computer aided, nuclear texture, optical density, displasia, hyperplasia, heterozygosity																	
C. DOCUMENTS CONSIDERED TO BE RELEVANT <table border="1"> <thead> <tr> <th>Category*</th> <th>Citation of document, with indication, where appropriate, of the relevant passages</th> <th>Relevant to claim No.</th> </tr> </thead> <tbody> <tr> <td>X</td> <td>Mathie M. E. et al: "PATHMASTER: Content Based Cell Image Retrieval Using Automated Feature Extraction": Journal of The American Medical Association, Vol. 7, No. 4, Jul/Aug 2000, pages 404 - 415. See the abstract, page 405, col. 1, second paragraph to col. 2, first paragraph under background, page 406, including the caption of figure 2, page 409, col. 1, and table 1, page 410, col. 1, under Digital Image Library, page 411, page 412, first and second paragraphs under Discussion and the caption of figure 9, page 413 - 414.</td> <td>1, 41</td> </tr> <tr> <td>Y</td> <td></td> <td>2 - 15, 19 - 23, 25, 30 - 39, 43 - 57, 60 - 62, 70, 72 - 81, 83 - 92</td> </tr> <tr> <td>X</td> <td>Poulin N et al: "Histometric Texture Analysis of DNA in Thin Sections from Breast Biopsies": Analytical and Quantitative Cytology and Histology, vol. 17, no. 5, October, 1995, pages 291 - 299. See page 291, under study design, page 292, col. 1, last paragraph - page 293, col. 2, second paragraph, page 294 - 295 and tables 1, 2 and 3 and figure 1, page 297, second paragraph and figure 3.</td> <td>1, 41</td> </tr> <tr> <td>Y</td> <td></td> <td>2 - 15, 19 - 23, 25, 35 - 39, 43 - 57, 60 - 62, 70, 78 - 81, 83 - 92</td> </tr> </tbody> </table>			Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.	X	Mathie M. E. et al: "PATHMASTER: Content Based Cell Image Retrieval Using Automated Feature Extraction": Journal of The American Medical Association, Vol. 7, No. 4, Jul/Aug 2000, pages 404 - 415. See the abstract, page 405, col. 1, second paragraph to col. 2, first paragraph under background, page 406, including the caption of figure 2, page 409, col. 1, and table 1, page 410, col. 1, under Digital Image Library, page 411, page 412, first and second paragraphs under Discussion and the caption of figure 9, page 413 - 414.	1, 41	Y		2 - 15, 19 - 23, 25, 30 - 39, 43 - 57, 60 - 62, 70, 72 - 81, 83 - 92	X	Poulin N et al: "Histometric Texture Analysis of DNA in Thin Sections from Breast Biopsies": Analytical and Quantitative Cytology and Histology, vol. 17, no. 5, October, 1995, pages 291 - 299. See page 291, under study design, page 292, col. 1, last paragraph - page 293, col. 2, second paragraph, page 294 - 295 and tables 1, 2 and 3 and figure 1, page 297, second paragraph and figure 3.	1, 41	Y		2 - 15, 19 - 23, 25, 35 - 39, 43 - 57, 60 - 62, 70, 78 - 81, 83 - 92
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.															
X	Mathie M. E. et al: "PATHMASTER: Content Based Cell Image Retrieval Using Automated Feature Extraction": Journal of The American Medical Association, Vol. 7, No. 4, Jul/Aug 2000, pages 404 - 415. See the abstract, page 405, col. 1, second paragraph to col. 2, first paragraph under background, page 406, including the caption of figure 2, page 409, col. 1, and table 1, page 410, col. 1, under Digital Image Library, page 411, page 412, first and second paragraphs under Discussion and the caption of figure 9, page 413 - 414.	1, 41															
Y		2 - 15, 19 - 23, 25, 30 - 39, 43 - 57, 60 - 62, 70, 72 - 81, 83 - 92															
X	Poulin N et al: "Histometric Texture Analysis of DNA in Thin Sections from Breast Biopsies": Analytical and Quantitative Cytology and Histology, vol. 17, no. 5, October, 1995, pages 291 - 299. See page 291, under study design, page 292, col. 1, last paragraph - page 293, col. 2, second paragraph, page 294 - 295 and tables 1, 2 and 3 and figure 1, page 297, second paragraph and figure 3.	1, 41															
Y		2 - 15, 19 - 23, 25, 35 - 39, 43 - 57, 60 - 62, 70, 78 - 81, 83 - 92															
☒ Further documents are listed in the continuation of Box C. ☐ See patent family annex. <table border="1"> <tr> <td> * Special categories of cited documents : "A" document defining the general state of the art which is not considered to be of particular relevance "E" earlier application or patent but published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed </td> <td> "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art "&" document member of the same patent family </td> </tr> </table>			* Special categories of cited documents : "A" document defining the general state of the art which is not considered to be of particular relevance "E" earlier application or patent but published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed	"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art "&" document member of the same patent family													
* Special categories of cited documents : "A" document defining the general state of the art which is not considered to be of particular relevance "E" earlier application or patent but published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed	"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art "&" document member of the same patent family																
Date of the actual completion of the international search 19 March 2007 (19 - 03-2007)		Date of mailing of the international search report 10 April 2007 (10-04-2007)															
Name and mailing address of the ISA/CA Canadian Intellectual Property Office Place du Portage I, C114 - 1st Floor, Box PCT 50 Victoria Street Gatineau, Quebec K1A 0C9 Facsimile No.: 001-819-953-2476		Authorized officer Patrick Norman 819- 997-2156															

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CA2006/001933

G06F 19/00 (2006.01) , ***G06T 7/40*** (2006.01) , ***G01N 1/30*** (2006.01)

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CA2006/001933

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	Henderson et al: "Development of Ganglion Cell Topography in Ferret Retina", The Journal of Neuroscience, April, 1988, pages 1194 - 1205. See page 1197, col. 1, second paragraph and the caption of figure 3.	2 - 7, 43 - 45, 51 - 57, 60 - 62, 70, 72 - 76, 78 - 81, 83 - 92, 99 - 102
Y	Raimondo F. et al: "Automated Evaluation of Her - 2/neu Status in Breast Tissue from Fluorescent In Situ Hybridization Images", IEEE transactions on Image Processing, vol. 14, No. 9, September, 2005, pages 1288 - 1299. See page 1291, col. 2, second paragraph.	2 - 7, 43 - 45, 51 - 57, 60 - 62, 70, 72 - 76, 78 - 81, 83 - 92, 99 - 102
Y	Boustany N. et al: "Calcium Induced alterations in Mitochondrial Morphology Quantified in Situ with Optical Scatter Imaging", Biophysical Journal, vol. 83, September, 2002, pages 1691 - 1700. See the caption of figure 3 and page 1694, col. 1, last paragraph	2 - 7, 43 - 45, 51 - 57, 60 - 62, 70, 72 - 76, 78 - 81, 83 - 92, 99 - 102

INTERNATIONAL SEARCH REPORTInternational application No.
PCT/CA2006/001933**Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of the first sheet)**

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons :

1. ☐ Claim Nos. :
because they relate to subject matter not required to be searched by this Authority, namely :

2. ☒ Claim Nos. :
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically :

Claims 16 - 18, 40, 42, 58 - 59, 63 - 68, 77, 82, 93 - 98 have not been searched because they add features which are unrelated to technical features of the inventions. PCT Rule 6.3a. Claims 103 and 104 have not been searched because they are omnibus claims which are so broad that a meaningful search could not be done.
3. ☐ Claim Nos. :
because they are dependant claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows :

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying additional fees, this Authority did not invite payment of additional fees.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claim Nos. :
4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claim Nos. :

Remark on Protest ☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.
☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.
☐ No protest accompanied the payment of additional search fees.