

Automated sputum cytometry for detection of intraepithelial neoplasias in the lung

Gerald Li^{a,*}, Martial Guillaud^a, Jean leRiche^{a,1}, Annette McWilliams^a, Adi Gazdar^b, Stephen Lam^a and Calum MacAulay^a

^a*Integrative Oncology Department, BC Cancer Research Centre, Vancouver, BC, Canada*

^b*Department of Pathology, UT Southwestern Medical Center, Dallas, TX, USA*

Received: July 12, 2011

Accepted: January 6, 2012

Abstract. *Background:* Despite the benefits of early lung cancer detection, no effective strategy for early screening and treatment exists, partly due to a lack of effective surrogate biomarkers. Our novel sputum biomarker, the Combined Score (CS), uses automated image cytometric analysis of ploidy and nuclear morphology to detect subtle intraepithelial changes that often precede lung tumours.

Methods: 2249 sputum samples from 1795 high-risk patients enrolled in ongoing chemoprevention trials were subjected to automated quantitative image cytometry after Feulgen-thionin staining. Samples from normal histopathology patients were compared against samples from carcinoma *in situ* (CIS) and cancer patients to train the CS.

Results: CS correlates with several lung cancer risk factors, including histopathological grade, age, smoking status, and p53 and Ki67 immunostaining. At 50% specificity, CS detected 78% of all highest-risk subjects—those with CIS or worse plus those with moderate or severe dysplasia and abnormal nuclear morphology.

Conclusion: CS is a powerful yet minimally invasive tool for rapid and inexpensive risk assessment for the presence of precancerous lung lesions, enabling enrichment of chemoprevention trials with highest-risk dysplasias. CS correlates with other biomarkers, so CS may find use as a surrogate biomarker for patient assessment and as an endpoint in chemoprevention clinical trials.

Keywords: Intraepithelial neoplasia (IEN), lung cancer, risk assessment, intermediate or pre-neoplastic markers and risk factors, biomarkers and intervention studies, chemoprevention, biomarkers and intervention, cancer surveillance and screening, chemoprevention clinical trials, quantitative image cytometry, ploidy analysis, malignancy associated changes

1. Introduction

Lung cancer is the leading cause of cancer death worldwide [17]. Despite decades of work by clinicians and research scientists, lung cancer still has a bleak, 5-year survival rate of just 16% [18]. While the 5-year

survival rate is about 70% for resected early-stage lung cancer [20], the majority of lung cancers are detected at an advanced stage with metastasis already present [18, 20]. To decrease lung cancer mortality, a strategy is needed to identify both patients with early disease, for treatment, and those with precancerous disease at risk of cancer development, for chemoprevention [45].

Historical efforts to find an effective lung cancer screening strategy using chest X-rays and standard sputum cytology [20] unfortunately found no decrease in mortality [10]. Many promising techniques are currently being studied, including low-dose spiral CT

¹Posthumous author.

*Corresponding author: Gerald Li, Integrative Oncology Department, BC Cancer Research Centre, 675 West 10th Ave, Vancouver, BC, Canada, V5Z 1L3. Tel.: +1 604 675 8000 Ext. 7089; Fax: +1 604 675 8099; E-mail: gli@bccrc.ca.

[15, 49], fluorescence bronchoscopy [9, 33], and sputum and blood biomarkers [31, 57]. More recently, the National Lung Screening Trial reported a 20% reduction in lung cancer mortality in high-risk individuals after receiving low-dose CT screening [34]. While both spiral CT and fluorescence bronchoscopy are sensitive technologies to detect early lung cancer, both methods suffer from limitations [7] and are costly if applied to high-risk subjects defined by age and smoking history alone.

Consequently, a critical part of any lung cancer screening strategy is a means to assess a participant's risk of having cancer or precancerous lesions at high risk of progressing to cancer. Risk assessment is not only useful in directing costly or invasive screening methods, but also allows at-risk populations to be identified for chemoprevention [45]. Furthermore, if the risk methodology is based on molecular alterations in the lung's genetic material, then changes in the assessment can be used to monitor the effectiveness of treatment or chemoprevention. The progression of normal lung tissue to invasive cancer is a complex process, involving multiple stages of increasing genetic and molecular insult. In the central airway, Saccomanno et al. showed that squamous cell carcinomas arise from a series of distinct pathological stages [41]. Moreover, lower grade dysplasias have a considerably lower rate of progression than severe dysplasias and carcinomas *in situ* [20]. Ascertaining the degree to which areas of the lung have progressed down this path—and the corresponding increased risk of cancer—is clinically important, as it should guide screening and chemopreventative therapy decisions.

The internationally accepted standard prognostic factor for lung cancer risk is the histopathological grade of a bronchial biopsy based on the World Health Organization classifications [16]. Biopsy, taken either endoscopically or surgically, is an invasive procedure and so other attempts to quantify risk have focused on patient factors, or biomarkers in sputum or blood samples [31, 47, 57]. Blood screen methodologies have included circulating DNA and RNA markers and proteomic profiling [2, 54, 57]. A study employing the detection of nanoarchitectural changes in buccal cells to detect lung cancer gave promising results, but their analysis was based on a small number of manually selected cells from a small number of patients [40]. In sputum, markers such as Ras and hnRNP B1 and the aberrant methylation of tumour suppressor genes have all been investigated [31]. Recent studies

have also studied the presence of specific chromosomal abnormalities in sputum using fluorescence *in situ* hybridization [19, 53] and pulmonary function [51] as possible risk factors for lung cancer.

Sputum biomarkers are promising because they are relatively quick and inexpensive while being adaptable to large-scale population screening [51], making them practical tools to guide both subsequent screening and chemoprevention trials. Studies into the diagnostic utility of conventional sputum cytology (summarized in [44]) have reported widely varying results, likely due to differences in methodologies between studies and significant intra- and inter-observer variations in identifying abnormal cells [30]. However, most studies have been directed at the detection of tumours and a lack of research remains into the utility of sputum cytology as a risk assessment tool for precancerous lesions.

As the percentage of bronchial epithelial cells in sputum can be quite low, reported sensitivities of sputum cytology for lung cancer detection tend to be quite low [30]. Malignancy associated changes (MAC) are subtle morphological and physiological changes that have been observed in non-malignant cells when cancer is present in a patient [35]. These changes may be due to soluble factors secreted by the malignant cells. Due to the larger number of non-malignant cells expected to exhibit MACs, we expect techniques based on MACs to be more sensitive than conventional cytology. Previous work in our group and others has shown that automated image cytometry based on MACs can be used to detect lung cancer [21, 39].

In this study, we correlate a number of published lung cancer risk factors—histopathological grade of biopsies from the bronchial tree, age, smoking status, quantitative morphometry, p53 and Ki67 biopsy status—to a novel sputum biomarker assay based on cell population ploidy status and malignancy associated changes [37, 39].

2. Materials and methods

Cell samples and data were drawn from several NCI-sponsored lung cancer chemoprevention trials in high-risk smokers, as defined by age and smoking history, [24–26] and from patients undergoing investigation for suspected lung cancer. A total of 2249 sputum samples were obtained between 2000 and 2006 from 1795 participants.

2.1. Chemoprevention trial subject recruitment and eligibility

For this study, a former smoker is defined as someone who has not smoked in the previous 12 months. A current smoker has smoked in the previous 12 months and a non-smoker is neither a former nor a current smoker. Former and current smokers between 40 to 74 years of age with a smoking history of ≥ 30 pack-years were recruited for the chemoprevention studies through the community outreach network of the public relations department of the British Columbia Cancer Agency (BCCA) using television programs, radio broadcasts, and through local newspapers. Following an initial interview during which study subjects completed a questionnaire to document their smoking history, we obtained a sputum sample from each subject using simultaneous high-frequency chest wall oscillation with an ABI Vest (Advanced Respiratory Inc., St. Paul, MN) and inhalation of 3% hypertonic saline from an ultrasonic nebulizer for 12 minutes [24, 25]. The subjects were instructed to cough intermittently during the induction procedure and for at least 2 hours afterwards to produce sputum samples. This procedure was found to be well tolerated by patients. The sputum samples were fixed in 50% ethanol and each sample was cytospun onto a glass slide and DNA was stained with Feulgen-thionin.

Some patients who volunteered for the chemoprevention studies did not meet the eligibility requirements for continuing on to participate in the bronchoscopy examination phase of those studies after a sputum sample was collected. A total of 1312 sputum samples in the present study were from such patients. Approval was granted by the Clinical Investigations Committees of the BCCA and The University of British Columbia. Written informed consent was obtained from all participants.

2.2. Semi-automated quantitative sputum analysis

An automated, high-resolution image cytometer (Cyto-SavantTM system from Oncometrics Inc., Vancouver, Canada) was programmed to measure the DNA content of at least 3000 objects per sample [8, 11, 39]. The image cytometer was subjected to the daily, weekly, monthly, and yearly quality assurance SOPs described in Chiu et al. [6] and Guillaud et al. [13] to ensure that the system's components (i.e.,

device and sample staining) were operating within their expected performance parameters. Each object detected on the slide was individually focused and scanned. Each object was then subjected to a discriminating function, in the form of a classification tree, which separated bronchial epithelial cells from other materials such as food particles, macrophages, lymphocytes, and other inflammatory cells [28, 38]. All cells were then reviewed by a trained cytotechnologist (certified by the Canadian Society of Laboratory Technologists). About 90% of all collected objects were identified to be epithelial cells after this procedure, which we have previously demonstrated yields comparable results to manually selecting nuclei [39].

For each epithelial cell, 110 nuclear features were calculated. These features can be divided into 6 categories: morphology (size and shape); densitometric properties (absorption amount and distribution); discrete texture features (euchromatin/heterochromatin); Markovian texture features (co-occurrence based); fractal texture features; and run-length texture features [8, 28]. For each slide, the average, standard deviation, skewness, and kurtosis were calculated for each feature from all epithelial cells found on the slide.

2.3. Ploidy measures

We have previously shown that Feulgen-thionin staining with our system is quantitative for DNA [11, 39]. Each cell's ploidy status was assessed by measuring the integrated optical density (IOD) of the nucleus and normalizing this against the mean IOD of the sample's diploid cell population, as determined from a frequency histogram of the nuclear IODs [13]. Diploid cells were assigned a DNA index of 1.0. A ploidy score for each slide was calculated by examining the frequency of cells falling within a series of DNA index ranges and then finding the range which had the most discriminating performance between normal and abnormal cases, where abnormal cases were defined as carcinomas *in situ* (CIS) and cancers. The ranges used were: <0.95, 0.95–1.00, 1.00–1.60, 1.60–1.85, 1.85–1.95, 1.95–2.09, 2.09–2.15, >2.15.

2.4. MAC measures

All MAC feature calculations considered only cells with DNA indices between 0.7 and 1.3. A training set was constructed by randomly sampling 100 cells

from each of the 36 normal, 6 CIS, and 36 cancer samples, as defined by the histopathological grading of their matching bronchial biopsies (see Section 2.5). The sampled cells from the CIS and cancer samples were then pooled together and compared against the sampled normal cells. A forward-stepping linear discriminant function analysis [14, 37] was performed on these two sets of about 4000 cells each, resulting in seven features selected to be indicative of malignancy associated changes.

The combined cytometric score for each slide was calculated from a linear combination of the 7 selected MAC features and the ploidy score. This combined cytometric score created a sputum-based biomarker that was used in the subsequent comparative analysis.

2.5. Biopsy collection and analysis

Atypia in a sample was defined as the presence of at least five cells which had DNA indices greater than 1.2 [24, 25]. All volunteers with atypical sputum were recalled and invited to be examined using autofluorescence bronchoscopy; they had an average of 7-8 bronchial biopsies taken per visit. Each of the 7934 biopsies collected was fixed in buffered formalin, embedded in paraffin, and serially sectioned. H&E-stained sections from each biopsy were systematically reviewed by two experienced lung pathologists (J LeRiche, A Gazdar), as previously described [24, 25]. Each biopsy was classified into one of the categories (normal, basal cell hyperplasia, metaplasia, mild/moderate/severe dysplasia, carcinoma *in situ*, cancer) in the histopathological system established by the World Health Organization [52]. Minor (i.e., one grade) differences in sample classification were resolved by telephone consultation between the two pathologists. If the diagnosis differed by two or more grades, both pathologists reviewed the slides again and reached a consensus diagnosis after communication by phone, email, or in person. The biopsies were matched by patient and date of collection to 1233 distinct sputum samples. For each of these sputum samples, the most severe consensus biopsy diagnosis associated with that sputum sample was recorded. In addition, all samples taken from patients who were subsequently diagnosed with CIS or cancer by non-bronchoscopic means (e.g., CT scans) within 8 months after sputum collection were also classified as CIS or cancer, as appropriate. All subjects who received, for any biopsy,

a biopsy grade of dysplasia, or worse, then had a Morphometry Index (MI) calculated for all their biopsies, according to the procedure set out in [14]. A total of 5060 biopsies had MIs calculated for this study.

178 biopsy samples taken from sites with a biopsy grading of at least dysplasia at baseline or follow-up were subjected to immunohistochemical analysis using 4 markers: p53, Ki67, bcl2, and cleaved caspase 3, as previously described [24]. They were graded visually on a 0-4 scale (with 0 indicating no stain and 4 indicating more than 75% of the nuclei staining positive) by experienced cytotechnologists. Of these biopsies, 159 corresponded to one of the sputum samples within the data set for the present analysis, matching both patient and time.

Some of the volunteers in this study either developed resectable lung cancer during the trial process or were discovered to have cancer upon enrolment. From 40 of these subjects who developed lung cancer, 73 sputum samples were collected either before or after surgical treatment. These 40 cases included patients with squamous cell carcinomas, adenocarcinomas, large cell lung cancers, and small cell lung cancers.

Except where otherwise noted, statistical significance in the present analysis was assessed using unpaired *t*-tests and ANOVA performed using STATISTICA software (StatSoft Inc., Tulsa, OK). $P < 0.05$ was considered statistically significant.

3. Results

The average age of the 1795 volunteers when samples were taken was 59.7 years (ranging from 39 to 83), and the average pack-years smoked was 48 (ranging from 8 to 221 amongst all current and former smokers). 57% of the samples were taken from male participants, 43% from females. 60.2% of the samples came from former smokers, 38.5% from current smokers, and 0.3% from non-smokers. The age distributions were similar between sexes: male average was 60 (range 39-83) and female average was 59 (range 39-81). However, there was some difference in their smoking history, with the male average pack-year exposure being 50, and the female average pack-year exposure being 44.

Upon comparing the ploidy characteristics of the normal and CIS/cancer training sets, the most

discriminating ploidy feature was found to be the frequency of epithelial cells with a ploidy amount between 1.6 and 1.85. Hence, this was used as the ploidy score. Seven features were found to be most indicative of malignancy associated changes: 1) the standard deviation (SD) of a nuclear morphology feature, *harmon05_fft* (a measure of nuclear roundness) [28], across all the epithelial cells measured for the sample, 2) the SD of 3 nuclear discrete texture features, *high_DNA_area*, *medium_DNA_amount* and *medium_average_distance* [8], across all the epithelial cells measured for the sample, 3) the SD of a Markovian texture feature, *correlation* [8], across all the epithelial cells measured for the sample, 4) the mean of a fractal texture feature, *fractal_dimension* [8], across all the epithelial cells measured for the sample and 5) the SD of a run-length texture feature, *maximum_run_length* (adapted from [8]), across all the epithelial cells measured for the sample. These 7 MAC-based features were combined with the ploidy score as a weighted sum to create a Raw Combined Score (Raw CS).

A plot of the Raw Combined Scores showed a dependency on the number of identifiable cells on each slide (Supplementary Data 1A). To correct for this, we subtracted from each Raw CS the value predicted by the distance-weighted least squares fit as a function of the number of cells measured on the slide. There were insufficient samples with more than 6000 cells to reliably estimate the trend, so the adjustment for samples with more than 6000 cells was set to zero. Except where otherwise specified, the adjusted Combined Score will be denoted simply as the Combined Score or CS for the remainder of the present analysis.

As with any cytological test, we must set a sample adequacy threshold that minimizes the scatter from measuring too few cells without excluding so many samples that it causes undue stress on patients and reduces the test's overall utility in a clinical setting. We chose 500 cells per slide as a threshold because below this level, the somewhat consistent patterns that the various histological categories exhibit break down (Supplementary Data 1B). Meanwhile, approximately 10% of the sample slides are excluded at this level, which was felt to be an acceptable rate. Hence, only sputum samples with at least 500 identifiable cells were used in the subsequent analysis.

A comparison of the sputum-derived Combined Score (CS) with the maximum histopathological grade of all the bronchial biopsies of the test subject at the

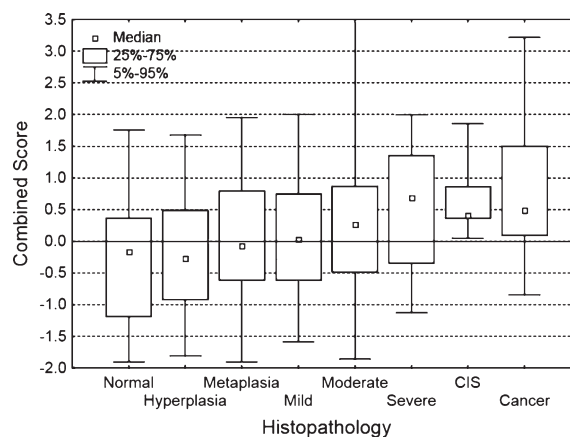


Fig. 1. Box plots of Combined Score for sputum samples containing more than 500 identifiable cells, grouped according to histopathological grade. There is a general increase in the median Combined Score when progressing to more pathologically severe cases.

corresponding time point is shown in Fig. 1. There is a clear trend that as pathological severity increases, so does the CS (F -test, $P < 10^{-5}$). Post hoc analysis using the Tukey Unequal N Honestly Significant Difference (HSD) test showed that the CS of the normal and hyperplasia groups were statistically significant from those of the cancers ($P = 0.003$ and 0.009 , respectively). As similar histopathological groups can often be difficult to distinguish, we created four new groups: normal/hyperplasia, metaplasia/mild dysplasia, moderate dysplasia to CIS, invasive cancer. When these groups are used, the trend between CS and pathological severity becomes even more evident, as shown in Fig. 2A. Post hoc analysis shows that the normal/hyperplasia group is significantly different from all other groups and the cancers are significantly different from the metaplasia/mild group (summarized in Fig. 2B).

To determine the degree to which the Combined Score can be used to ascertain lung cancer risk, we compared the CS to other known risk factors and biomarkers of lung cancer. Using the Morphometry Index in conjunction with histopathological grading, we created high- and low-risk subject groups, which we will denote m-risk. A given patient was considered low-m-risk if he or she had a histopathological grading of hyperplasia, or less, and a maximum MI < 1.36 , as described in [25]. High-m-risk subjects had a maximum MI > 1.36 and a histopathological grading of moderate dysplasia, or worse. Additionally, all CIS and cancer patients were denoted high-m-risk,

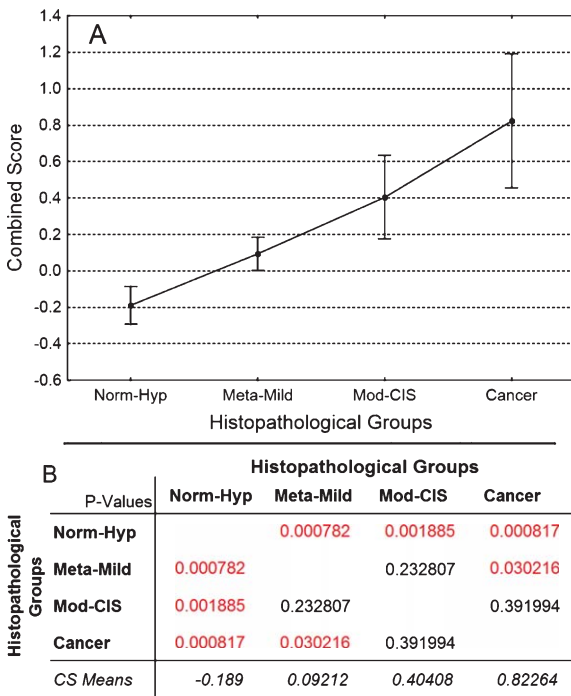


Fig. 2. Analysis of Combined Scores for samples from patients in each of the four histopathological groups created by combining similar grades: normal/hyperplasia, metaplasia/mild dysplasia, moderate dysplasia to CIS, invasive cancer. A: Plot of mean CS. Error bars denote 95% confidence intervals. B: Summary *P*-value matrix of Tukey Unequal N HSD Post Hoc analysis. Significant *P*-values ($p < 0.05$) are highlighted in red and the italicized row shows mean Combined Scores in each group.

regardless of MI. Given the strong correlation between the Morphological Index and cancer risk [14], we feel that this combined assessment represents a more accurate approximation of lung cancer risk than does a system that relies on histopathology alone. If all the subjects that fit into the high- or low-m-risk categories are grouped together, there is a significant correlation between the Combined Score and the m-risk groups ($P = 0.00004$) (plot shown as Supplementary Data 2). Removing the sample used for training from this analysis, this correlation between CS and m-risk groups still holds ($P = 0.008$).

Using the m-risk groups, we can construct a receiver operating characteristic (ROC) curve representing the ability of the Combined Score to distinguish high-m-risk patients from low-m-risk patients (Fig. 3). Patients with metaplasia and mild dysplasia are considered neutral m-risk. As we are unsure whether to consider them

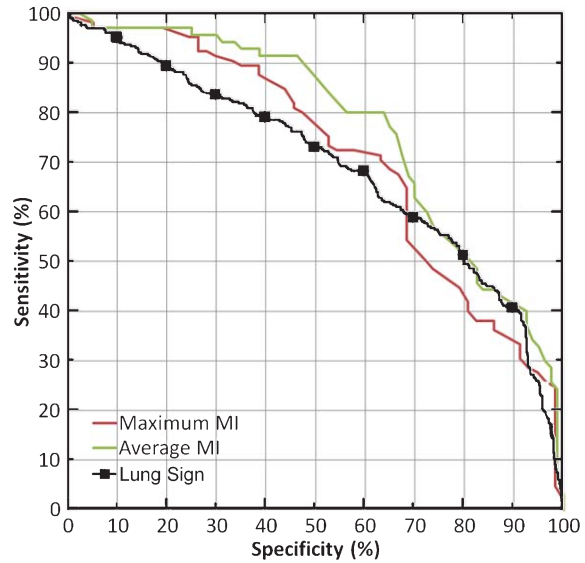


Fig. 3. ROC curves showing the ability of the Combined Score to distinguish between high- and low-m-risk patients. Patient m-risk groups were defined as described in the text. Patient m-risk can be defined using the maximum MI of the biopsies from the subject or the average MI from the biopsies taken from the subject, both cases for which are shown. For comparison, the ROC curve for LungSign is shown as well (adapted from Fig. 2 in [21]). Areas under the curve are 0.711, 0.766, and 0.692 for maximum MI, average MI, and LungSign, respectively.

high- or low-m-risk, we excluded them from this analysis. Since many patients received multiple biopsies over the course of the study, MI can be used to determine m-risk either by using the maximum MI at a given time point or the average MI at that time point. In either case, the worst histopathological diagnosis was used to determine the m-risk group. CS performed well by both definitions of m-risk, although using the average MI resulted in a noticeably better area under the curve than using the maximum MI (AUC by trapezoidal rule, 0.766 and 0.711, respectively). If all samples used in training the CS are removed from this analysis, the areas under the curve become 0.752 and 0.677, respectively.

Age and pack-years smoked are the most widely studied epidemiological lung cancer risk factors. Both have been shown to be key predictors of lung cancer risk [5]. Plotting each sample's Combined Score against age and pack-years smoked shows a positive correlation in each case (data not shown). A linear regression with a statistically significant positive slope can be calculated in each case, which

confirms that CS increases with age and smoking history. Furthermore, these trends remain even if the subject population is broken down into male/female and current smoker/former smoker groups. The subject population was divided into these groups because these subpopulations may exhibit differences in cancer risk or progression [23, 31].

There has been considerable work to find immunohistochemical markers—in blood, sputum, or biopsies—that correlate with lung cancer risk [12, 16]. Two of the more successful lung cancer immunohistochemical markers are p53 and Ki67. p53, which is the most studied marker for all cancers including lung cancer [16], has been shown to be overexpressed in many premalignant bronchial lesions [1]. Furthermore, overexpression of p53 in a lesion correlates with an increased risk of a lesion progressing to invasive cancer [1]. Immunostaining with the proliferation marker Ki67, which is expressed in the G₁, S, G₂, and M phases of the cell cycle [43], has been shown to be of prognostic value in a number of cancers, including lung cancer [4, 58]. Ki67 expression has further been demonstrated to increase as preneoplastic lung lesions progress from mild dysplasia to CIS [32].

Figure 4A and B plot the immunohistochemical staining score of p53 and Ki67 respectively, for each of the 159 biopsies selected (see Methods) against its histopathological grading. Overall, there was a positive correlation between p53 staining and histological grade. There was also a statistically significant difference between hyperplasia and all of the more severe grades (Fig. 4A). The proliferation marker Ki67 shows an even more pronounced progression of increasing staining with increasing severity of pathological grading (Fig. 4B).

To compare p53 and Ki67 immunohistochemical staining to the sputum-based Combined Score, the sample populations were separated into current and former smoker subgroups. To more clearly illustrate the trends in the data, the five immunohistochemical staining scores were combined into two groups: no/weak staining (scores 0–1) and stronger staining (scores 2–4). The CS correlates with the p53 immunohistochemical staining for both former smokers and current smokers (Fig. 5A and B, respectively). Ki67 staining correlated with increasing Combined Score values for the former smoker subgroup (Fig. 5C). However, there was no discernable pattern in the plot comparing Ki67 immunostaining to the Combined Score values for current smokers (Fig. 5D). This is

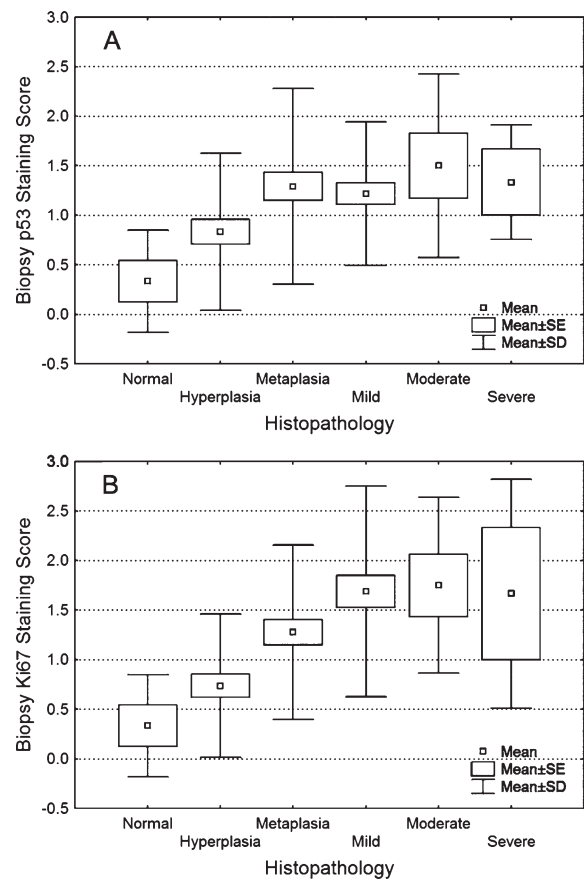


Fig. 4. Box plots comparing p53 (A) and Ki67 (B) staining score to histopathological grade for the 159 biopsy samples from sites that had a biopsy grading of dysplasia or worse at baseline or follow-up.

mainly because the CS for the current smoker cases with weaker Ki67 staining is as high as the scores for all cases with stronger Ki67 staining regardless of smoking status. For the comparisons shown in each of the four panels (Fig. 5A–D), $P=0.08$, 0.008 , 0.1 , 0.6 , respectively, although we can likely attribute the findings of insignificance in the former smoker comparisons to an insufficient number of cases.

Given the impact of smoking status on the interaction between the Ki67 measurements and Combined Score, we turned our attention to other risk factors that might have a confounding effect on our analysis of CS. We found that there were small but not significant differences between the high- and low-m-risk groups in terms of age, smoking history, and sex (t -tests, sex by Pearson χ^2 , $P=0.1$, 0.2 , 0.3 , respectively), but high-m-risk patients were significantly more likely to be current smokers (Pearson χ^2 , $P=0.0009$). Analyzing

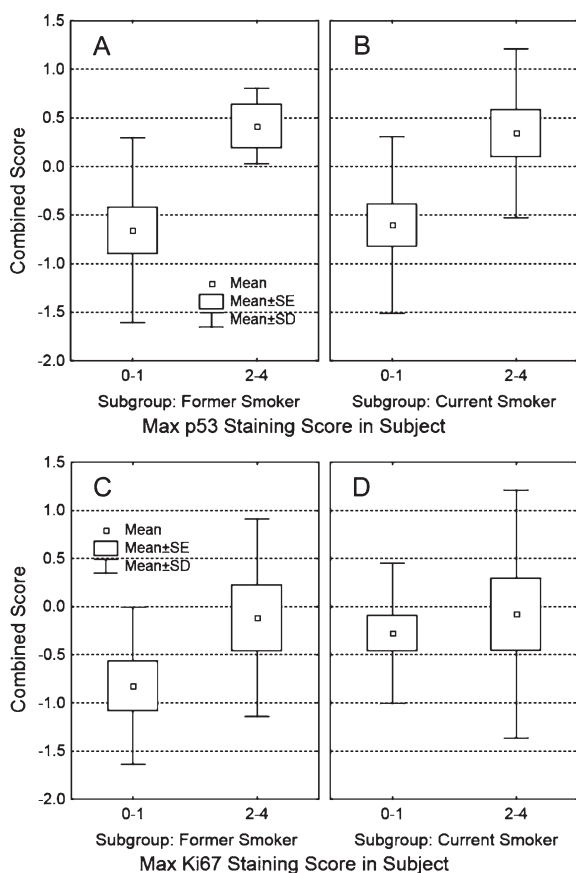


Fig. 5. Box plots comparing Combined Score to the maximum p53 (A, B) and Ki67 (C, D) staining score in that patient at that point in time. Immunostaining scores were grouped into two categories: no/weak staining (scores 0–1) and stronger staining (scores 2–4). Cases were further subdivided into former smoker (A, C) and current smoker (B, D) groups. For the comparisons shown in each of the four panels (A–D), $P = 0.08, 0.008, 0.1, 0.6$, respectively.

current and former smokers separately, we found that in both cases, high-m-risk patients had significantly higher CS than low-m-risk patients ($P = 0.01, 0.002$, respectively). Furthermore, Fig. 1 replotted with current and former smokers separately shows the same general trends in each subgroup as the original figure, demonstrating that smoking status does not have a confounding effect on our analysis of CS overall.

The most important feature of any surrogate biomarker is its correlation with cancer risk or progression. While our participant criteria were not designed to find lung cancer patients, a number of study participants developed lung cancer over the course of the study. Additionally, some patients recruited on account of receiving a bronchoscopy for other

clinical indications were found to have lung cancer upon enrolment. We compared sputum samples taken within eight months before surgery with sputum samples collected at least six months after the surgical resection treatment protocol. There was a statistically significant difference ($P = 0.003$) between the CS for the samples taken before and after surgery. The t -test was unpaired because not enough patients had data both before and after surgery for a pair-wise test to be statistically meaningful.

Among the sputum samples linked to a positive cancer diagnosis, there was no significant difference in CS between distal and proximal tumours ($P = 0.9$). When broken down by cancer subtype, there was no significant difference between adenocarcinomas and squamous carcinomas ($P = 0.1$). There were insufficient samples of small cell and other non-small cell lung cancers to make any other statistically meaningful comparisons.

4. Discussion

It has been suggested that the traditional view that cancer begins when invasive disease is first detected should be replaced by one in which carcinogenesis itself is the disease, with invasive or symptomatic cancer being merely the final outcome [48]. Consequently, treatments should aim to “reverse, suppress, or prevent the process of carcinogenesis” [12]. This is the goal of chemoprevention, with past work in our group and others showing promise [24–26, 42].

However, many early chemoprevention studies for lung cancer have actually shown neutral or even negative effect from chemopreventative agents. Disappointing results from these early studies may be due to the fact that many of these studies used smoking status as the primary selection criterion, resulting in a study population with an insufficiently high risk to benefit from chemoprevention [55]. Many precancerous lesions never progress even without treatment [36] and so chemoprevention will offer these patients no additional benefit. If we can remove these patients from a study population and only study those who are likely to progress to invasive disease without treatment, any effect from chemoprevention should become more evident.

End points for chemoprevention studies are typically the incidence of invasive cancer or mortality

[55]. Since many pre-cancers never develop into invasive disease, regardless of whether chemopreventative agents are used, this makes such trials long and costly. Governments and pharmaceutical companies may be reluctant to invest in the development of cancer chemoprevention drugs and strategies due to the immense research cost, especially while lung cancer chemoprevention is still not a universally accepted approach to the management of the disease. In order to accelerate the development and verification of new chemopreventative agents, intermediate end points need to be identified and validated.

Computer analysis of sputum samples has previously been used to detect lung cancer, using criteria based on ploidy [56], MACs [37], or both [21]. The LungSign test, for example, combined ploidy and MAC analysis, and was effective in detecting 40% of all lung cancers with 91% specificity consistently across all subtypes and stages, far better than the results from conventional cytology [21]. MAC analysis is appealing because it can be measured on non-malignant cells, which typically greatly outnumber malignant ones in sputum samples. The features we used describe various aspects of the nuclear architecture. Changes in the chromatin distribution and organization may be indicative of changes in activation and expression of genes. Genetic and epigenetic alterations, which may be related to cell cycle, metabolic, or differentiation status of the cell, are reflected in these MAC features [8]. Using a similar approach to LungSign, we have devised a novel biomarker combining ploidy and MAC analysis. Unlike LungSign, which was optimized for the detection of invasive cancers with high specificity, the Combined Score presented in this paper is designed to detect dysplasias. By detecting pre-cancerous lesions before they become invasive cancers, the CS could allow the highest-risk patients to be enrolled in chemopreventative therapy trials in an effort to reduce their risk of progression to cancer. Since the Combined Score correlates with dysplasia grade, the effectiveness of any such intervention can also be safely and easily monitored over time.

As a biomarker for lung cancer risk, our analysis shows that the Combined Score correlates with a number of other known lung cancer risk factors. When compared to histopathological grade as shown in Fig. 1, there is a clear trend towards higher Combined Scores with increasing disease severity. This trend is apparent even if we were to remove all the normal, CIS, and cancer cases, the sample sets from which the

training set was derived. Since Morphometry Indices for biopsies from patients who progressed to cancer were significantly higher than non-progressing lesions of the same histopathological grade [14], MI can supplement histopathology. By combining histopathology and MI, we can get a better assessment of cancer risk (which we denoted m-risk). In the context of chemoprevention trials, then, we would enrol patients with high m-risk as we feel that they would benefit most from treatment. Adding the MI to our analysis, we found that CS correlates even better with m-risk (supplementary Data 2) than with histopathology alone.

The fact that the correlation between CS and histopathological grading is not as strong as that between MI and pathological grades in previous studies [14] reflects the difference between the subtle malignancy associated changes that occur lung-wide and the more pronounced changes found in the diagnostic cells of biopsies. However, collecting biopsies (upon which both histopathological grading and MI are based) is still an invasive technique. The correlation between CS and the combination of histopathology and MI suggest that CS could be used as a rapid, non-invasive, and relatively inexpensive alternative to these techniques for both risk assessment and the conduction of chemoprevention studies.

We used MI and histopathology as a gold standard to assess the performance of CS in identifying those patients at highest risk to progress to invasive cancer. Figure 3 shows ROC curves using either maximum or average MI to define m-risk. Both these methods of determining m-risk generate noticeably different ROC curves. Clinically, a physician may be interested in determining the risk of progression of the most severe lesion and so a risk assessment using the maximum MI is most appropriate. However, a sputum biomarker is based on a sampling of cells from throughout the lungs. As expected, then, when the average MI is used as the criterion for determining m-risk, the ROC curve for CS looks improved over the maximum MI case. Nonetheless, even when using the maximum MI as the m-risk criterion, the ROC curve for CS compares very well with that for LungSign. This is despite the fact that LungSign seeks to distinguish between cancerous (CIS or worse) and non-cancerous samples, whereas the Combined Score is able to separate high-grade dysplasias from normals, an arguably much more challenging task. While our samples were not routinely screened with conventional cytology, a subset of our samples overlaps with those used in the LungSign

study, where they reported a sensitivity of 16% and a specificity of 99.1% for cytology. The CS showed a similar level of sensitivity at that level of specificity.

The ideal analysis for any novel biomarker would be to see which patients ultimately develop cancer. This requires extensive follow-up and even then, only a small number ever progress. In the absence of data on actual cancer progression, the next best alternative is to ensure that the novel biomarker correlates with known biomarkers. We found that Combined Score correlated with age, smoking history, and immunohistochemical staining of p53 and Ki67, all of which have been previously found to correlate with lung cancer risk [1, 5, 12, 32]. Except for Ki67 staining in current smokers, these trends were further found to be applicable to both current and former smokers.

In former smokers, the Combined Score shows a particularly strong correlation with p53 staining. This may be due in part to “field cancerization”, a concept first proposed by Slaughter et al. [46] to explain the propensity of individuals with one malignancy to develop second primary tumours. Mutations and changes in expression level of the p53 gene across wide areas of the lung have previously been reported in subjects with dysplasia or preinvasive lesions [1]. Since CS is based on a sampling of cells from throughout the lungs, we might expect a better correlation with an immunohistochemical marker whose expression has likewise been altered over a large region of the pulmonary mucosa.

Ongoing exposure to cigarette smoke causes inflammation in the lungs and has been shown to be associated with an increased expression of not only Ki67 [27] but also PCNA [22], another important proliferation marker. This confounds our analysis and may help explain why we do not observe a trend between CS and Ki67 staining in current smokers, as any correlation between CS and Ki67 may be dwarfed by the impact of smoking on proliferation across the lung. Further, smoking is known to alter the expression not only of a large number of genes [48] but the chromatin structure as well [50] and these changes are different in current and former smokers. Our Ki67 staining results, when compared to the Combined Score, illustrate one more example of the difference between the lungs of current and former smokers, underscoring the necessity of taking smoking status into consideration for any proliferation-based diagnosis or treatment.

To address the issue of potential confounding effects in our analyses of CS, we compared the age, smoking history, sex, and smoking status of patients in our two

m-risk groups. As these are all documented to affect lung cancer risk, we expected to see some differences between the groups. Except in the case of smoking status, the differences we observed were too small to be considered potential confounders. A follow-up analysis showed that among both current and former smokers, the general trends we observed in Fig. 1 still hold and high-m-risk patients have higher CS, so smoking status does not impart any additional confounding effect on our analyses. However, as current smokers generally had higher CS than former smokers ($P = 0.00001$), different thresholds may need to be set if CS were to be used in a clinical setting.

Since the Combined Score is presented on a continuous numeric scale, it allows smaller changes in lung health to be detected. The use of automated image analysis also means that it should be more objective than standard histopathology. We’ve shown that the Combined Score, like the Morphometry Index, reflects a patient’s cancer risk, with high-m-risk patients having a significantly higher CS than low-m-risk patients. We believe that CS can be used to monitor chemoprevention trials as an alternative end point. Unlike the MI, however, the CS is a sputum-based biomarker, which is less invasive and more likely to be tolerated by patients. This further allows CS to be measured repeatedly over the course of the trial.

In chemoprevention trials, these advantages mean that trials can be designed to use a reduction in CS by a certain threshold amount as an alternative end point, instead of waiting for invasive disease to develop. We can consider our analysis of surgically resected lung cancer cases to be an example of this, as we can think of surgery and chemoprevention as two different interventions and CS as a common scale by which to assess their effectiveness. In cases where lung cancer has been treated by surgery, the Combined Scores before and after surgery are significantly different. We are further encouraged by the observation that our sample contains a mix of squamous cell carcinomas and adenocarcinomas. With the small sample size, however, it is difficult to properly assess the ability of CS to detect successful surgery. There is also insufficient data to assess whether CS performs better with squamous cell carcinomas or adenocarcinomas. The present study was not designed to test CS in this setting, but the initial results suggest this is another potential application of CS that merits further study. Similar to how CS can be tuned for optimal detection of precancerous lesions, the continuous scale of the Combined Score allows us in the future to

select a good threshold for detecting successful surgery once data from a larger and more comprehensive study is available.

The observation that post-surgery Combined Scores are lower than scores from before surgery suggests that CS is sensitive to MACs, which was the intention in training the CS. While the correlative evidence is weak, an advantage of using MACs as a pre-screening test is the extra sensitivity inherent in being able to detect malignancy even when a sputum sample consists primarily of non-malignant cells, as is often the case. In addition to being able to detect MACs, the Combined Score appears to be able to detect the effects of field cancerization. The correlation of CS with p53 staining is suggestive of this, as is the observation that CS is better able to assess m-risk when m-risk is calculated on the basis of average MI as opposed to maximum MI (Fig. 3). The magnitude of a MAC effect might be expected to correlate with the most severe lesion present releasing soluble factors to which the surrounding cells respond, but the CS appears to correlate more with the severity of the overall “cancer field” as reflected in the average MI. This is of benefit to the design of future chemoprevention studies as it would be informative to be able to monitor the overall level of field cancerization in response to a candidate chemopreventive therapy. Our data weakly suggests that the CS can act as a surrogate biomarker in this regard.

We view MACs and field cancerization as separate but possibly related phenomena. While the prevailing field cancerization hypothesis suggests that cancers arise from a field of altered cells, previous work with pre- and post-surgery patients suggests that cancer cells themselves influence histologically normal cells [29]. As these effects can be reversed by removing the tumour, it has been hypothesized that such effects may be a response in histologically normal cells to autocrine signals released by malignant cells [29]. Although our data cannot provide insight into the mechanisms underlying the morphological changes detected by the Combined Score, our results weakly suggest that CS may correlate with both MACs and field cancerization.

Despite decades of work, there remains no widely accepted screening test for early lung cancer detection. Studies using spiral CT, for example, showed high sensitivity for detecting non-calcified pulmonary nodules, but had a low specificity, which, coupled with a low overall prevalence of lung cancer even amongst heavy smokers, led to a low positive predictive value [31] and consequently increased costs due to follow-up test-

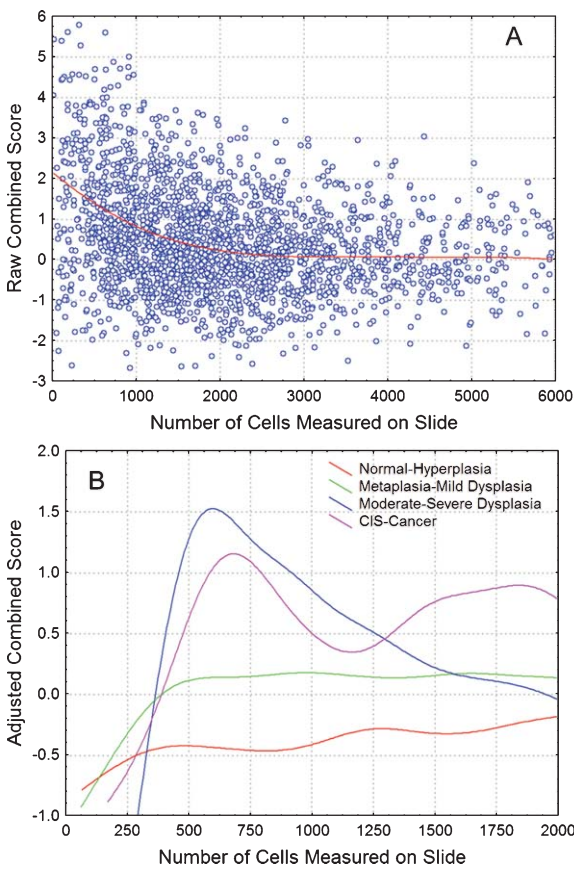
ing and unnecessary surgical interventions. To address these shortcomings, it has been suggested that automated sputum cytometry could be used as an initial screening test, thereby increasing the disease prevalence amongst those subsequently screened by CT and autofluorescence bronchoscopy [41]. While our intent was not to design a novel pre-screening tool, our analysis of the Combined Score as a pre-screen for patients most likely to benefit from chemoprevention suggests that CS could be potentially be used to pre-screen for patients most likely to benefit from secondary lung cancer screening with CT and autofluorescence bronchoscopy. Our study population was at high risk of developing lung cancer on the basis of demographic risk factors (i.e., age and smoking history). We envision any potential use of CS in a pre-screening setting would also focus on such a subset of patients, as these are patients most likely to benefit from additional screening. Moreover, patients at high risk due to age and smoking history are readily identified by the use of a patient questionnaire.

Like the LungSign test, the Combined Score is a sputum biomarker that has an adjustable classification threshold. This allows the performance to be optimized to best complement other early lung cancer detection methods [21]. In such a pre-screening scenario, we would like a test with a high sensitivity, while tolerating a lower specificity. The performance of the Combined Score matches that of the LungSign test very well. At a specificity of 50%, for example, we can achieve 78% sensitivity, which is slightly better than LungSign at that level of specificity. This means we could reduce the number of CTs by half and still catch roughly three-quarters of all high-m-risk pre-cancers. This would have significant cost savings and mean less risk of increased cancer incidence caused by radiation exposure due to unnecessary CT scans [3]. However, one must keep in mind that our analysis of the ability of CS in identifying high-m-risk lesions excludes metaplasias and mild dysplasias, which may result in better perceived performance. More study will be needed to validate the use of CS as a pre-screener in conjunction with more invasive screening tools in a clinical setting.

Attempts to develop effective screening tools for lung cancer have faced many challenges. Just as importantly, where patients have been found to harbour precancerous lesions, there remain no widely accepted interventions as research into chemoprevention is currently hampered by a lack of effective

surrogate biomarkers to serve as end points for trials. We have presented evidence that the Combined Score, a novel automated sputum image cytometry biomarker based on ploidy and MAC analysis, correlates with other known lung cancer risk factors like histopathology, age, smoking status, and immunohistochemistry of p53 and Ki67. Compared to LungSign, a similar sputum biomarker, Combined Score achieves a similar

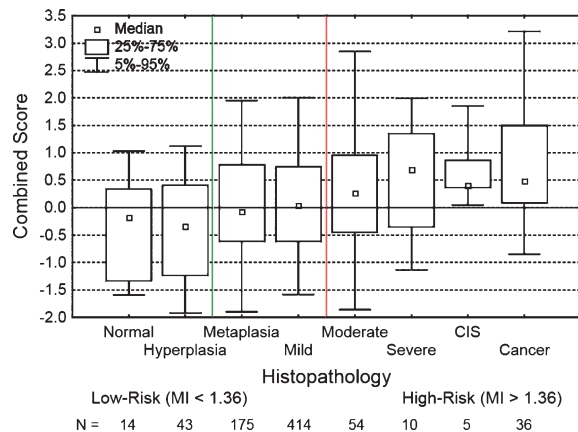
performance separating high- and low-m-risk precancers to that of LungSign separating cancers and non-cancers. Patients with high Combined Scores are prime candidates for enrolment in chemoprevention studies, where the Combined Score may be most useful as a method of monitoring response and screening for a higher risk study population more likely to benefit from treatment. This will hopefully spur more interest in investigating chemopreventative therapies that will treat the carcinogenic process before invasive disease appears, saving money and patient lives in the long run.



Supplementary Data 1. Dependence of the Combined Score on the number of identifiable cells. Trend lines are distance-weighted least squares fits. A: As the number of identifiable cells increases, the scatter decreases and the Raw Combined Score becomes a more consistent measurement. The distribution of histopathological grades, meanwhile, is quite consistent across the range of identifiable cell counts. However, a distinct trend towards higher Raw Combined Scores at lower cell counts necessitated a cell count normalization procedure. B: Adjusted Combined Scores, with data categorized according to the highest grade of abnormality found in that patient's biopsies. Data points have been removed to highlight the trends. Below a count of 500 cells per slide, the consistent patterns that the various histological categories exhibit break down, as seen in the rapid changes and convergence of the four running average curve lines.

Acknowledgments

The authors would like to thank Jagoda Korbelik, Anita Carraro, Priscilla Fung, Deanna Ceron, Anne Dy Buncio, and Sukhinder Khattrra for technical assistance and database management. Funding for this project was provided by the National Institutes of Health (project numbers 1P01-CA96964-05 and 1U01-CA96109-05).



Supplementary Data 2. Box plots of Combined Score for sputum samples, sorted according to m-risk. Normal and hyperplasia groups are low-m-risk and only data for which the low-m-risk MI criterion is also met (max MI < 1.36) is shown here. Moderate and severe dysplasia are considered high-m-risk and the data shown here only includes cases where the high-m-risk MI criterion is also met (max MI > 1.36). Additionally, all CIS and cancer cases were counted as high-m-risk, regardless of MI. Metaplasia and mild dysplasia are neutral m-risk and all data in these groups is shown. The numbers at the bottom indicate the number of sputum samples in each group. The Combined Scores for the low- and high-m-risk groups are significantly different ($P=0.00004$).

References

- [1] E. Brambilla, S. Gazzeri, S. Lantuejoul, J.L. Coll, D. Moro, A. Negoescu and C. Brambilla, p53 mutant immunophenotype and deregulation of p53 transcription pathway (Bcl2, Bax, and Waf1) in precursor bronchial lesions of lung cancer, *Clin Cancer Res* **4** (1998), 1609–1618.
- [2] R.M. Bremnes, R. Sirera and C. Camps, Circulating tumour-derived DNA and RNA markers in blood: A tool for early detection, diagnostics, and follow-up? *Lung Cancer* **49** (2005), 1–12.
- [3] D.J. Brenner, Radiation risks potentially associated with low-dose CT screening of adult smokers for lung cancer, *Radiology* **231** (2004), 440–445.
- [4] D.C. Brown and K.C. Gatter, Ki67 protein: The immaculate deception? *Histopathology* **40** (2002), 2–11.
- [5] A. Cassidy, S.W. Duffy, J.P. Myles, T. Liloglou and J.K. Field, Lung cancer risk prediction: A tool for early detection, *Int J Cancer* **120** (2007), 1–6.
- [6] D. Chiu, M. Guillaud, D. Cox, M. Follen and C. MacAulay, Quality assurance system using statistical process control: An implementation for image cytometry, *Cell Oncol* **26** (2004), 101–117.
- [7] J.M. Crowell, S.G. Baker, P.M. Marcus, J.D. Clapp and B.S. Kramer, Cumulative incidence of false-positive test results in lung cancer screening: A randomized trial, *Ann Intern Med* **152** (2010), 505–512.
- [8] A. Doudkine, C. MacAulay, N. Poulin and B. Palcic, Nuclear texture measurements in image cytometry, *Pathologica* **87** (1995), 286–299.
- [9] E. Edell, S. Lam, H. Pass, Y.E. Miller, T. Sutedja, T. Kennedy, G. Loewen, R.L. Keith and A. Gazdar, Detection and localization of intraepithelial neoplasia and invasive carcinoma using fluorescence-reflectance bronchoscopy: An international, multicenter clinical trial, *J Thorac Oncol* **4** (2009), 49–54.
- [10] R.S. Fontana, D.R. Sanderson, L.B. Woolner, W.F. Taylor, W.E. Miller and J.R. Muhm, Lung cancer screening: The Mayo program, *J Occup Med* **28** (1986), 746–750.
- [11] D.M. Garner, A. Harrison, C. MacAulay and B. Palcic, Cyto-Savant™ and its use in automated screening of cervical smears, In: G.L. Wied, P.H. Bartels, P.H. Rosenthal and U. Schenck, eds., *Compendium on the Computerized Cytology and Histology Laboratory*, Tutorials of Cytology, Chicago, 1994, pp. 346–352.
- [12] J. Gray, J.T. Mao, E. Szabo, M. Kelley, J. Kurie and G. Bepler, Lung cancer chemoprevention: ACCP evidence-based clinical practice guidelines (2nd Edition), *Chest* **132** (2007), 56S–68S.
- [13] M. Guillaud, D. Cox, A. Malpica, G. Staerkel, J. Maticic, D. Van Niekirk, K. Adler-Storthz, N. Poulin, M. Follen and C. MacAulay, Quantitative histopathological analysis of cervical intra-epithelial neoplasia sections: Methodological issues, *Cell Oncol* **26** (2004), 31–43.
- [14] M. Guillaud, J.C. le Riche, C. Dawe, J. Korbelik, A. Coldman, I.I. Wistuba, I.-W. Park, A. Gazdar, S. Lam and C.E. MacAulay, Nuclear morphometry as a biomarker for bronchial intraepithelial neoplasia: Correlation with genetic damage and cancer development, *Cytometry A* **63** (2005), 34–40.
- [15] C.I. Henschke, D.F. Yankelevitz, D.M. Libby, M.W. Pasmantier, J.P. Smith and O.S. Miettinen, Survival of patients with Stage I lung cancer detected on CT screening, *N Engl J Med* **355** (2006), 1763–1771.
- [16] F.R. Hirsch, D.T. Merrick and W.A. Franklin, Role of biomarkers for early detection of lung cancer and chemoprevention, *Eur Respir J* **19** (2002), 1151–1158.
- [17] A. Jemal, F. Bray, M.M. Center, J. Ferlay, E. Ward and D. Forman, Global cancer statistics, *CA Cancer J Clin* **61** (2011), 69–90.
- [18] A. Jemal, R. Siegel, J. Xu and E. Ward, Cancer statistics, *CA Cancer J Clin* **60** (2010), 277–300.
- [19] R.L. Katz, T.M. Zaidi, R.L. Fernandez, J. Zhang, W. He, C. Acosta, M. Daniely, L. Madi, M.A. Vargas, Q. Dong, X. Gao, F. Jiang, N.P. Caraway, A.A. Vaporciyan, J.A. Roth and M.R. Spitz, Automated detection of genetic abnormalities combined with cytology in sputum is a sensitive predictor of lung cancer, *Mod Pathol* **21** (2008), 950–960.
- [20] G.J. Kelloff, S.M. Lippman, A.J. Dannenberg, C.C. Sigman, H.L. Pearce, B.J. Reid, E. Szabo, V.C. Jordan, M.R. Spitz, G.B. Mills, V.A. Papadimitrakopoulou, R. Lotan, B.B. Aggarwal, R.S. Bresalier, J. Kim, B. Arun, K.H. Lu, M.E. Thomas, H.E. Rhodes, M.A. Brewer, M. Follen, D.M. Shin, H.L. Parnes, J.M. Siegfried, A.A. Evans, W.J. Blot, W.H. Chow, P.L. Blount, C.C. Maley, K.K. Wang, S. Lam, J.J. Lee, S.M. Dubinett, P.F. Engstrom, F.L. Meyskens Jr, J. O'Shaughnessy, E.T. Hawk, B. Levin, W.G. Nelson, W.K. Hong, The AACR Task Force on Cancer Prevention and Progress in chemoprevention drug development: The promise of molecular biomarkers for prevention of intraepithelial neoplasia and cancer – A plan to move forward, *Clin Cancer Res* **12** (2006), 3661–3697.
- [21] R.A. Kemp, D.M. Reinders and B. Turic, Detection of lung cancer by automated sputum cytometry, *J Thorac Oncol* **2** (2007), 993–1000.
- [22] F.R. Khuri, J.S. Lee, S.M. Lippman, J.J. Lee, S. Kalapurakal, R. Yu, J.Y. Ro, R.C. Morice, W.K. Hong and W.N. Hittelman, Modulation of proliferating cell nuclear antigen in the bronchial epithelium of smokers, *Cancer Epidemiol Biomarkers Prev* **10** (2001), 311–318.
- [23] C. Kiyohara and Y. Ohno, Sex differences in lung cancer susceptibility: A review, *Gen Med* **7** (2010), 381–401.
- [24] S. Lam, J.C. leRiche, A. McWilliams, C. MacAulay, Y. Dyachkova, E. Szabo, J. Mayo, R. Schellenberg, A. Coldman, E. Hawk and A. Gazdar, A randomized Phase IIb trial of pulmicort turbuhaler (budesonide) in people with dysplasia of the bronchial epithelium, *Clin Cancer Res* **10** (2004), 6502–6511.
- [25] S. Lam, C. MacAulay, J.C. leRiche, Y. Dyachkova, A. Coldman, M. Guillaud, E. Hawk, M.-O. Christen and A.F. Gazdar, A randomized Phase IIb trial of anethole dithiolethione in smokers with bronchial dysplasia, *J Natl Cancer Inst* **94** (2002), 1001–1009.
- [26] S. Lam, A. McWilliams, J. leRiche, C. MacAulay, L. Wattenberg and E. Szabo, A Phase I study of myo-inositol for lung cancer chemoprevention, *Cancer Epidemiol Biomarkers Prev* **15** (2006), 1526–1531.
- [27] J.J. Lee, D. Liu, J.S. Lee, J.M. Kurie, F.R. Khuri, H. Ibarguen, R.C. Morice, G. Walsh, J.Y. Ro, A. Broxson, W.K. Hong and

- W.N. Hittelman, Long-term impact of smoking on lung epithelial proliferation in current and former smokers, *J Natl Cancer Inst* **93** (2001), 1081–1088.
- [28] C. MacAulay, *Development, implementation and evaluation of segmentation algorithms for the automatic classification of cervical cells*, Ph.D. Dissertation, The University of British Columbia, 1989.
- [29] C. MacAulay, S. Lam, P.W. Payne, J.C. LeRiche and B. Palcic, Malignancy-associated changes in bronchial epithelial cells in biopsy specimens, *Anal Quant Cytol Histol* **17** (1995), 55–61.
- [30] A. McWilliams, B. Lam and T. Sutedja, Early proximal lung cancer diagnosis and treatment, *Eur Respir J* **33** (2009), 656–665.
- [31] A. McWilliams, C. MacAulay, A.F. Gazdar and S. Lam, Innovative molecular and imaging approaches for the detection of lung cancer and its precursor lesions, *Oncogene* **21** (2002), 6949–6959.
- [32] A.P. Meert, F. Feoli, B. Martin, J.M. Verdebout, C. Mascaux, A. Verhest, V. Ninane and J.P. Sculier, Ki67 expression in bronchial preneoplastic lesions and carcinoma *in situ* defined according to the new 1999 WHO/IASLC criteria: A preliminary study, *Histopathology* **44** (2004), 47–53.
- [33] K. Moghissi, K. Dixon and M.R. Stringer, Current indications and future perspective of fluorescence bronchoscopy: A review study, *Photodiagnosis Photodyn Ther* **5** (2008), 238–246.
- [34] National Lung Screening Trial Research Team, Reduced lung-cancer mortality with low-dose computed tomographic screening, *N Engl J Med* **365** (2011), 395–409.
- [35] H.E. Nieburgs, Recent progress in the interpretation of malignancy associated changes (MAC), *Acta Cytol* **12** (1968), 445–453.
- [36] J.A. O’Shaughnessy, G.J. Kelloff, G.B. Gordon, A.J. Dannenberg, W.K. Hong, C.J. Fabian, C.C. Sigman, M.M. Bertagnolli, S.P. Stratton, S. Lam, W.G. Nelson, F.L. Meyskens, D.S. Alberts, M. Follen, A.K. Rustgi, V. Papadimitrakopoulou, P.T. Scardino, A.F. Gazdar, L.W. Wattenberg, M.B. Sporn, W.A. Sakr, S.M. Lippman and D.D. Von Hoff, Treatment and prevention of intraepithelial neoplasia: An important target for accelerated new agent development, *Clin Cancer Res* **8** (2002), 314–346.
- [37] B. Palcic, D.M. Garner, J. Beveridge, X.R. Sun, A. Doudkine, C. MacAulay, S. Lam and P.W. Payne, Increase of sensitivity of sputum cytology using high-resolution image cytometry: Field study results, *Cytometry* **50** (2002), 168–176.
- [38] B. Palcic, C. MacAulay, S. Shlien, W. Treurniet, H. Tezcan and G. Anderson, Comparison of three different methods for automated classification of cervical cells, *Anal Cell Pathol* **4** (1992), 429–441.
- [39] P.W. Payne, T.J. Sebo, A. Doudkine, D. Garner, C. MacAulay, S. Lam, J.C. LeRiche and B. Palcic, Sputum screening by quantitative microscopy: A reexamination of a portion of the National Cancer Institute Cooperative Early Lung Cancer Study, *Mayo Clin Proc* **72** (1997), 697–704.
- [40] H.K. Roy, H. Subramanian, D. Damania, T.A. Hensing, W.N. Rom, H.I. Pass, D. Ray, J.D. Rogers, A. Bogojevic, M. Shah, T. Kuzniar, P. Pradhan and V. Backman, Optical detection of buccal epithelial nanoarchitectural alterations in patients harboring lung cancer: Implications for screening, *Cancer Res* **70** (2010), 7748–7754.
- [41] G. Saccomanno, V.E. Archer, O. Auerbach, R.P. Saunders and L.M. Brennan, Development of carcinoma of the lung as reflected in exfoliated cells, *Cancer* **33** (1974), 256–270.
- [42] M. Saito, H. Kato, T. Tsuchida and C. Konaka, Chemoprevention effects on bronchial squamous metaplasia by folate and vitamin B12 in heavy smokers, *Chest* **106** (1994), 496–499.
- [43] T. Scholzen and J. Gerdes, The Ki-67 protein: From the known and the unknown, *J Cell Physiol* **182** (2000), 311–322.
- [44] G. Schreiber and D.C. McCrory, Performance characteristics of different modalities for diagnosis of suspected lung cancer: Summary of published evidence, *Chest* **123** (2003), 115S–128S.
- [45] T. Shaipanich, A. McWilliams and S. Lam, Early detection and chemoprevention of lung cancer, *Respirology* **11** (2006), 366–372.
- [46] D.P. Slaughter, H.W. Southwick and W. Smejkal, “Field cancerization” in oral stratified squamous epithelium; clinical implications of multicentric origin, *Cancer* **6** (1953), 963–968.
- [47] M.R. Spitz, C.J. Etzel, Q. Dong, C.I. Amos, Q. Wei, X. Wu and W.K. Hong, An expanded risk prediction model for lung cancer, *Cancer Prev Res (Phila)* **1** (2008), 250–254.
- [48] M.B. Sporn, Chemoprevention of cancer, *Lancet* **342** (1993), 1211–1213.
- [49] S.J. Swensen, J.R. Jett, T.E. Hartman, D.E. Midthun, S.J. Mandrekar, S.L. Hillman, A.M. Sykes, G.L. Aughenbaugh, A.O. Bungum and K.L. Allen, CT screening for lung cancer: Five-year prospective experience, *Radiology* **235** (2005), 259–265.
- [50] P. Szulakowski, A.J. Crowther, L.A. Jimenez, K. Donaldson, R. Mayer, T.B. Leonard, W. MacNee and E.M. Drost, The effect of smoking on the transcriptional regulation of lung inflammation in patients with chronic obstructive pulmonary disease, *Am J Respir Crit Care Med* **174** (2006), 41–50.
- [51] M.C. Tammemagi, S.C. Lam, A.M. McWilliams and D.D. Sin, Incremental value of pulmonary function and sputum DNA image cytometry in lung cancer risk prediction, *Cancer Prev Res (Phila)* **4** (2011), 552–561.
- [52] W.D. Travis, T.V. Colby, B. Corrin, Y. Shimosato and E. Brambilla, Histologic and graphical text slides for the histological typing of lung and pleural tumors, In: *International histological classification of tumors*, 3rd ed., World Health Organization Pathology Panel, Springer Verlag, Berlin, 1999, p. 5.
- [53] M. Varella-Garcia, A.P. Schulte, H.J. Wolf, W.J. Feser, C. Zeng, S. Braudrick, X. Yin, F.R. Hirsch, T.C. Kennedy, R.L. Keith, A.E. Baron, S.A. Belinsky, Y.E. Miller, T. Byers and W.A. Franklin, The detection of chromosomal aneusomy by fluorescence *in situ* hybridization in sputum predicts lung cancer incidence, *Cancer Prev Res (Phila)* **3** (2010), 447–453.
- [54] Y.C. Wang, H.S. Hsu, T.P. Chen and J.T. Chen, Molecular diagnostic markers for lung cancer in sputum and plasma, *Ann NY Acad Sci* **1075** (2006), 179–184.
- [55] R.C. Winterhalder, F.R. Hirsch, G.K. Kotantoulas, W.A. Franklin and P.A. Bunn Jr, Chemoprevention of lung cancer—from biology to clinical reality, *Ann Oncol* **15** (2004), 185–196.
- [56] S. Xing, B. Khanavkar, J.A. Nakhosteen, Z. Atay, K.-H. Jöckel, W. Marek and The Research Institute for Diagnosis Treatment of Early Lung Cancer (RIDTELC) Lung Study Group, Predictive value of image cytometry for diagnosis of lung cancer in heavy smokers, *Eur Respir J* **25** (2005), 956–963.

- [57] J. Yee, M.D. Sadar, D.D. Sin, M. Kuzyk, L. Xing, J. Kondra, A. McWilliams, S.F. Man and S. Lam, Connective tissue-activating peptide III: A novel blood biomarker for early lung cancer detection, *J Clin Oncol* **27** (2009), 2787–2792.
- [58] C.-Q. Zhu, W. Shih, C.-H. Ling and M.-S. Tsao, Immunohistochemical markers of prognosis in non-small cell lung cancer: A review and proposal for a multiphase approach to marker evaluation, *J Clin Pathol* **59** (2006), 790–800.