Automated Flow Cytometry Test Distinguishes Cancer from Non-Cancer in Sputum with High Sensitivity and Specificity

Lydia H. Bederka[1], Patricia Araujo[1], Jamila Sanchez[1], Marcia Grayson[1], Madeleine E. Lemieux[2], Jennifer Rebeles[1], Shao-Chiang Lai[1], Xavier T. Reveles[1] and **Vivienne I. Rebel** (**presenter**)[1]

[1]bioAffinity Technologies, San Antonio, TX, USA;
[2]Bioinfo, Plantagenet, ON, Canada

IASLC 2020 World Conference on Lung Cancer Singapore
JANUARY 28-31, 2021 | WORLDWIDE VIRTUAL EVENT

DISCLOSURES PRESENTER

Commercial Interest	Relationship(s)
bioAffinity Technologies	Employment, Stock Interest, Patent holder

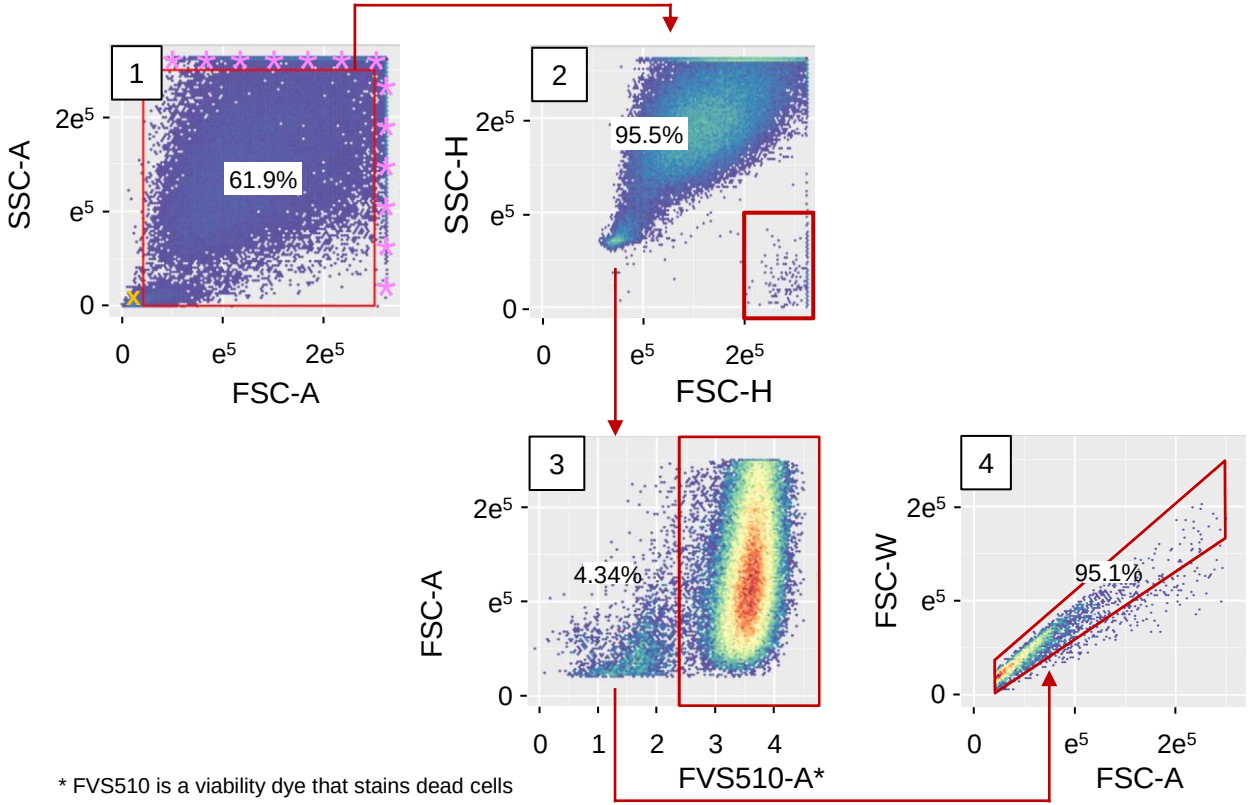
Introduction

Low-dose spiral computed tomography (LDCT) screening for lung cancer reduces mortality. It is therefore recommended in the U.S.A. for high-risk individuals between 55 and 80 years of age, who have smoked 30 pack years or more and have not quit smoking for more than 15 years. LDCT results may not always lead to a clear follow-up procedure when the nodules are small, particularly when they are between 4 mm to 20 mm in size. We therefore sought to develop a non-invasive test (CyPath® Lung) to detect lung cancer in individuals at high risk with emphasis on the test's ability to distinguish cancer in individuals with smaller nodules.

Materials and Methods

Sputum was collected at home over three days and shipped overnight to the laboratory. Participants included people at high-risk for developing lung cancer, who at the time of providing the sample were either cancer-free based on LDCT results and, in some cases, received a negative lung biopsy or were diagnosed with lung cancer confirmed by biopsy. Sputum was processed upon receipt into a single-cell suspension before labeling with a viability dye to exclude dead cells, antibodies directed against CD45, CD206, CD66b, CD3, CD19, EpCAM and panCytokeratin to distinguish blood and non-blood cell populations, and a porphyrin (TCPP) to identify cancer and/or cancer-associated cells. Samples were run on a flow cytometer and the data of 170 participants were used to develop an automated flow cytometry analysis platform. Of the 171 samples, 154 samples were used to build a predictive model to distinguish between cancer and non-cancer samples. Three samples were discarded because of technical difficulties experienced at the time of data acquisition. Quality control measures that became part of the automated analysis discarded an additional thirteen samples because of a lack of sufficient cells and one sample was discarded because of too few lung macrophages (See also **Fig 2**).

Automated analysis: eliminating debris, dead cells and cell doublets



* FVS510 is a viability dye that stains dead cells

Fig 1. Cleaning up the sample before analysis. Depicted are the first four steps of the automation that leads to the identification of the sputum cells of interest. Events outside the red box in profile (1) were eliminated, which included small debris (X) and cell aggregates (*). The events selected in the red box in profile (1) were further analyzed and additional debris (red box in profile (2)) and dead cells (red box in profile (3)) were eliminated. Lastly, single cells (red box in profile (4)) were separated from cell doublets and small aggregates (cells outside the red box in profile (4)).

The events in the red box in profile (4) represent single, live sputum cells ("cells of interest"), which were then subjected to a sample quality control check (**Fig 2**).

Automated analysis: quality control

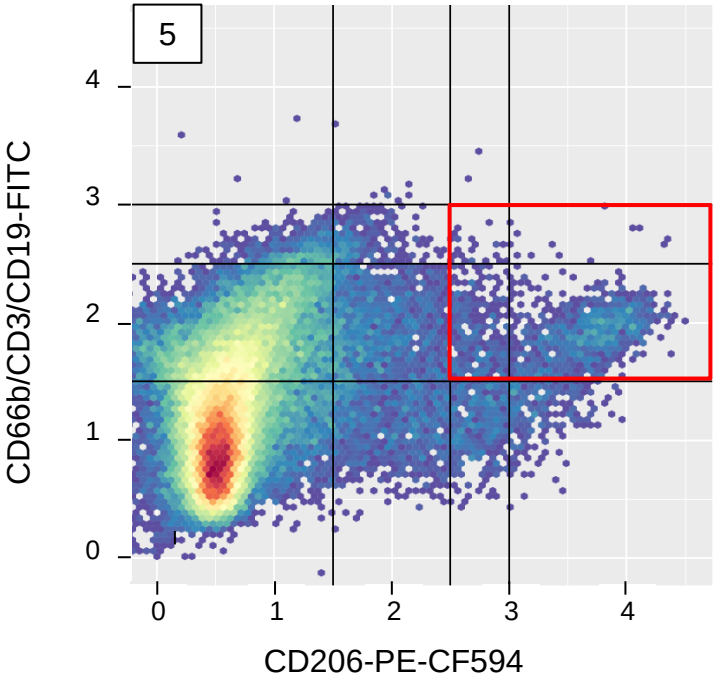


Fig 2. Cleaning up the sample before analysis. Depicted are the sputum cells of interest, selected as described in **Fig 1**. Cells were stained with a cocktail of FITC-labeled antibodies (CD66b, CD3 and CD19) and a CD206 antibody labeled with PE-CF594. CD206 specifically identifies alveolar and interstitial lung macrophages (cells indicated by the red box). CD206-positive cells therefore can serve as an indication of a sample being from the lung as opposed to merely being saliva. Samples were considered inadequate if less than 0.05% of the cells of interest fell into the red box.

Additionally, a sample that yielded a profile with less than 10,000 cells of interest was also considered inadequate because subsequent analysis of less than 10,000 cells was statistically inferior.

Of the 171 samples initially provided for development of the automated analysis and building the model 13 (7.6%) contained too few cells and 1 (0.6%) had too few macrophages. An additional 3 samples were excluded because of technical difficulties with the acquired data, leading to 154 samples available for model building.

Identifying flow cytometric parameters for building the model

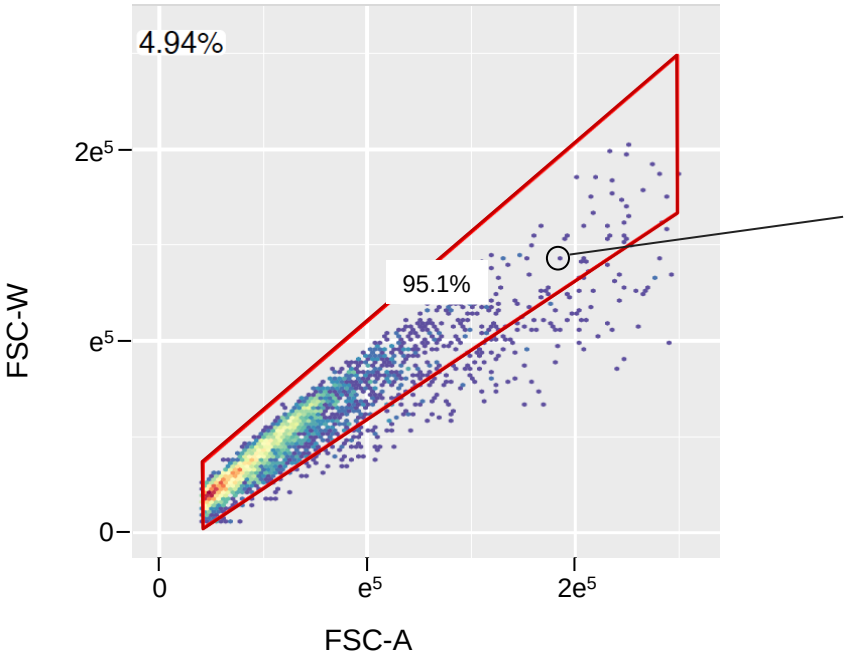


Fig 3. Single live sputum cells (in red box) identified from samples obtained from high-risk patients with or without lung cancer were used to identify predictive markers belonging to the cancer group rather than the non-cancer group.

Each cell is defined by information about:

- SSC pulse (height, width and area under the curve)
- FSC pulse (height, width and area under the curve)

- APC (TCPP) (fluorescence intensity)
- BV510 (viability dye) (fluorescence intensity)

dyes

- PE (fluorescence intensity)
- FITC (fluorescence intensity)
- PE-CF594 (fluorescence intensity)

antibodies

Development of the model to distinguish cancer from non-cancer

The distribution pattern of flow cytometric parameters (SSC, FSC, dyes or antibodies; see **Fig 3**) as well as patient features (such as age, gender, smoking history; see **Table 1**) were evaluated individually and in combination as predictors of belonging to the cancer cohort rather than the high-risk group.

Samples were divided into training (2/3) and test (1/3) sets that maintained the relative proportion of cancer to high-risk samples of the whole group. The training set was used to fit the model while the test set was used to validate the model.

This process of random selection of training and test sets was repeated to evaluate the robustness of the model and avoid over-fitting the data.

Four parameters were found to be predictive and used in the model algorithm:

- APC (TCPP) density of signal (APC/SSC-A)
- BV510 density (viability dye) of signal (BV510/FSC-A)
- Relative size of the population with FITC between 2.5-3 and PE-CF594 < 1.5, suggesting a non-macrophage population may be of importance
- Age

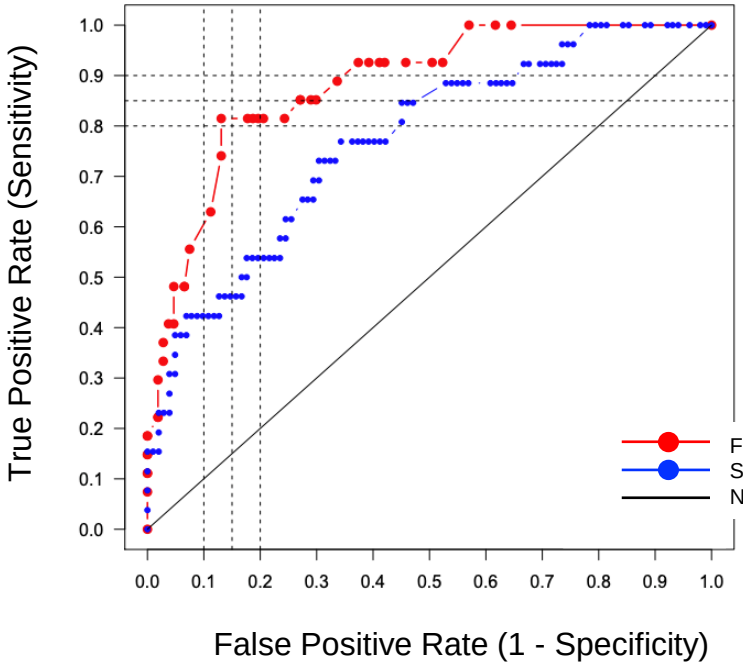
The status (cancer versus non-cancer) of four of the samples used to build the model could not be confirmed and therefore, the test performance (**Fig 4**) was determined on 150 samples.

Patient features of the high-risk and cancer group

Variable		High Risk; n=122	Cancer, n=28	p-value
Patient demographics				
Age (years)	mean (SD)	65.0 (5.5)	71.4 (5.1)	< 0.0001
Male	No. (%)	57 (46.7)	21 (75)	< 0.0001
Female	No. (%)	65 (53.3)	7 (25)	n.s.
Smoking Status				
Never	No. (%)	0 (0)	1 (3.6)	
Former	No. (%)	69 (56.6)	15 (53.6)	
Pack years	mean (SD)	56.1 (24.3)	53.3 (36.3)	n.s.
Current	No. (%)	53 (43.4)	12 (42.9)	
Pack years	mean (SD)	55.2 (26.5)	51.8 (14.1)	n.s.
Comorbidities				
COPD	No. (%)	81 (66.4)	13 (46.4)	n.s.
LDCT Characteristics				
Nodule-negative	No. (%)	38 (31.1)	3 (11.7)*	n.s.
No. of Nodules**	mean (SD)	3.1 (2.1)	2.2 (1.5)	0.026
Nodule size (mm)**	mean (SD)	6.9 (7.3)	24.3 (16.7)	< 0.0001
	size range	2 - 56	4 - 63	

Table 1 * no nodules were reported in nodule-negative cancer patient. Instead, abnormalities were described as “masses” or “opacities” without a size
** scans in high risk group with no nodules were excluded n.s. = not significant

Results of the flow-cytometry based CyPath® Lung test



Sensitivity:	82%
Specificity:	88%
PPV:	62%
NPV:	95%
Accuracy:	87%

Fig 4

* a previous version of the CyPath test based on sputum smears stained with the porphyrin (TCPP) label and a microscope to identify TCPP-labeled cells

CyPath Lung ® performance according to nodule size

Largest nodule size (mm)	CA (No.)	HR (No.)	Sensitivity (%)	Specificity (%)
≤ 6	2	98	100	89.8
> 6 - < 12	5	14	100	78.6
> 6 - < 20	11	21	90.9	76.2
< 20	13	119	92.3	87.4
> 6 - < 30	17	22	88.2	100
> 30	6	2	66.7	100
Data not available*	3	0	N/A	N/A

Table 2

* no nodules were reported in nodule-negative cancer patients. Instead, abnormalities were described as “masses” or “opacities” and no size was provided

No. = number of samples; CA = cancer; HR = high-risk; N/A = not applicable;

Tumor information in cancer patient group (n = 28)

Tumor type	No. (%)
Non-small cell carcinoma	1 (3.6)
Adenocarcinoma	11 (39.3)
Squamous cell carcinoma	13 (46.4)
Large cell carcinoma	1 (3.6)
Small cell carcinoma	2 (7.1)

Stage	No. (%)*
I	9 (42.9)
II	2 (9.5)
III	5 (23.8)
IV	5 (23.8)
NA	7

Table 3

* Percentages are calculated on total of samples for which information was available
NA = information not available

Conclusions

The automated CyPath® Lung flow cytometric assay correctly classifies study participants into cancer or high-risk cohorts with high accuracy, including participants with nodules smaller than 20 mm.

The CyPath® Lung assay thus has the potential to complement LDCT screening and improve diagnosis of early stage lung cancer.