CLAIMS

What is claimed is:

1. A flow cytometer method for automatically analyzing, with a computer, a sputum sample from a subject suspected of having lung cancer comprising:

obtaining *in vitro* a plurality of cells from the sputum sample from the subject suspected of having lung cancer;

marking the plurality of cells with *i*) a plurality of cell lineage specific markers, *ii*) a cell viability marker and *iii*) a tetra (4-carboxyphenyl) porphyrin (TCPP) marker;

analyzing with the flow cytometer the plurality of cells marked with *i-iii* to obtain a subpopulation selected for cell size from the plurality of cells based upon an automatically selected bead size exclusion gate;

from the cell size selected subpopulation, automatically selecting, by the computer, a viable singlet population of cells using an automated non-debris gate and an automated singlets gate;

from the viable singlet population of cells, automatically obtaining, by the computer, flow cytometer values based upon the plurality of cell lineage specific markers, the cell viability marker and the TCPP marker;

automatically applying, by the computer, a trained classifier to a meta data from the subject and the flow cytometric values obtained wherein the meta data of the subject includes age; and

automatically generating, by the computer, based upon the application of the trained classifier, a classification for the sputum sample wherein the classification is selected from a plurality of classification options comprising cancer and non-cancer, wherein the cell viability marker is FVS510, the plurality of cell lineage specific markers are selected from fluorescent anti-CD206, fluorescent anti-CD3, fluorescent anti-CD19, and fluorescent anti-CD66b and the trained classifier is

$$Cancer \sim b_0 + b_1 * age + b_2 * CD206^{low}CD66b/CD3/CD19^{mid} + b_3 * \frac{FVS510}{FSC - A_{log_{10}}}[R2] + b_4 * \frac{TCPP}{SSC - A_{log_{10}}}[R3] - b_5 * age * \frac{FVS510}{FSC - A_{log_{10}}}[R2]$$

wherein:

b_0 is an intercept value when all variables in the trained classifier are set to zero;

b_1-b_5 are weights for each variable determined by fitting the trained classifier to the training set;

$CD206^{low}CD66b/CD3/CD19^{mid}$ is a population of cells that are labeled with fluorescent anti-CD206 and identified as “low” on the signal intensity logicle scale dotplot and simultaneously labeled with fluorescent anti-CD66b/anti-CD3/anti-CD19 and identified as “mid” on the signal intensity logicle scale dotplot;

*FVS*510/FSC-A_{log10} [R2] is a proportion of cells with a *FVS*510 signal density greater than or equal to 0.3 but less than 0.6 (“[R2]”) where the *FVS*510 signal density is a per cell *FVS*510 fluorescent signal on the logicle scale divided by the FSC-A_{log10}; and

TCPP/SSC-A_{log10} [R3] is a proportion of cells with a TCPP signal density greater than or equal to 0.6 (“[R3]”) where the signal density is a per cell TCPP fluorescent signal on the logicle scale divided by the SSC-A_{log10}.

2. The method of claim 1 wherein the sputum sample is a single cell suspension.

3. The method of claim 1 wherein the plurality of cell lineage specific markers are selected from CD206, CD3, CD19, CD66b, CD45, EpCAM, PanCK, and any combination thereof.

4. The method of claim 1 wherein the cell viability marker labels dead cells preferentially to live cells.

5. The method of claim 1 wherein the automatically analyzing step comprises obtaining, from the plurality of cells, the flow cytometry values for side scatter, forward scatter, fluorescence from TCPP, fluorescence from the cell viability marker, and fluorescence from the plurality of cell lineage specific markers.

6. The method of claim 1 wherein the bead size exclusion gate is set between 5μm and 30μm wherein events less than 5μm and greater than 30μm are not further analyzed.

7. The method of claim 1 wherein the automated non-debris gate excludes a majority of dead cells from a non-debris population.

8. The method of claim 1 wherein the automated singlets gate is applied to a population of cells selected in the automated non-debris gate.

9. The method of claim 1 wherein the sputum sample includes a minimum number of CD206 expressing cells for the sputum sample to be acceptable for determination of a lung health.

10. A system for automated analysis of flow cytometry data, the system comprising:

a computer processor in communication with a memory having stored therein flow cytometry data from a plurality of markers in a plurality of cells from a sputum sample of a subject wherein the plurality of markers include *i*) a plurality of cell lineage specific markers, *ii*) a cell viability marker and *iii*) a tetra (4-carboxyphenyl) porphyrin (TCPP) marker;

a computer-program product embodied in a non-transitory computer readable medium, the computer-program product comprising instructions for causing the computer processor to:

receive the flow cytometry data acquired from the plurality of cells from a sputum sample;

select from the plurality of cells in the sputum sample a subpopulation of cells automatically selected based upon application of automatic gates selected from a bead size exclusion gate, a viability gate and a singlets gate;

determine, from the subpopulation, flow cytometric values of interest for the plurality of cell lineage specific markers, the cell viability marker and the TCPP marker;

automatically apply a classifier to the flow cytometric values of interest and a meta data of the subject;

automatically generate an output at a display device with an identification of one or more classifications for the sputum sample comprising cancer or non-cancer, wherein the cell viability marker is FVS510, the plurality of cell lineage specific markers are selected from fluorescent anti-CD206, fluorescent anti-CD3, fluorescent anti-CD19, and fluorescent anti-CD66b and the trained classifier is

$$Cancer \sim b_0 + b_1 * age + b_2 * CD206^{low}CD66b/CD3/CD19^{mid} + b_3 * \frac{FVS510}{FSC - A_{log_{10}}}[R2] + b_4 * \frac{TCPP}{SSC - A_{log_{10}}}[R3] - b_5 * age * \frac{FVS510}{FSC - A_{log_{10}}}[R2]$$

wherein:

b_0 is an intercept value when all variables in the trained classifier are set to zero;

b_1-b_5 are weights for each variable determined by fitting the trained classifier to the training set;

$CD206^{low}CD66b/CD3/CD19^{mid}$ is a population of cells that are labeled with fluorescent anti-CD206 and identified as “low” on the signal intensity logicle scale dotplot and simultaneously labeled with fluorescent anti-CD66b/anti-CD3/anti-CD19 and identified as “mid” on the signal intensity logicle scale dotplot;

*FVS*510/FSC-A_{log10} [R2] is a proportion of cells with a *FVS*510 signal density greater than or equal to 0.3 but less than 0.6 (“[R2]”) where the *FVS*510 signal density is a per cell *FVS*510 fluorescent signal on the logicle scale divided by the FSC-A_{log10}; and

TCPP/SSC-A_{log10} [R3] is a proportion of cells with a TCPP signal density greater than or equal to 0.6 (“[R3]”) where the signal density is a per cell TCPP fluorescent signal on the logicle scale divided by the SSC-A_{log10}.

11. The system of claim 10 wherein the cell viability marker labels dead cells preferentially to live cells.

12. The system of claim 10 wherein the flow cytometry values for the plurality of cells are obtained for side scatter, forward scatter, fluorescence from TCPP, fluorescence from the cell viability marker, and fluorescence from the plurality of cell lineage specific markers.

13. The system of claim 10 wherein the bead size exclusion gate is set to exclude events having a size of less than 5μm and greater than 30μm.

14. The system of claim 10 wherein the automated non-debris gate excludes a majority of dead cells from a non-debris population.

15. A non-transitory computer-readable medium comprising program code that, when executed, causes processing circuitry to:

obtain flow cytometer values from a viable singlet population of a subject's sputum sample based upon side scatter, forward scatter, fluorescence from TCPP, fluorescence from a cell viability marker, and fluorescence from a plurality of cell lineage specific markers;

apply a trained classifier to a meta data from the subject and the flow cytometric values obtained, wherein the meta data of the subject is age; and

generate, based upon the application of the trained classifier, a classification for the sputum sample wherein the classification is selected from a plurality of classification options comprising cancer and non-cancer, wherein the cell viability marker is FVS510, the plurality of cell lineage specific markers are selected from fluorescent anti-CD206, fluorescent anti-CD3, fluorescent anti-CD19, and fluorescent anti-CD66b and the trained classifier is

$$\begin{aligned} \textit{Cancer} \sim b_0 + b_1 * \textit{age} + b_2 * CD206^{\textit{low}}CD66b/CD3/CD19^{\textit{mid}} \\ + b_3 * \frac{FVS510}{FSC - A_{log_{10}}}[R2] + b_4 * \frac{TCPP}{SSC - A_{log_{10}}}[R3] \\ - b_5 * \textit{age} * \frac{FVS510}{FSC - A_{log_{10}}}[R2] \end{aligned}$$

wherein:

b_0 is an intercept value when all variables in the trained classifier are set to zero;

b_1-b_5 are weights for each variable determined by fitting the trained classifier to the training set;

$CD206^{low}CD66b/CD3/CD19^{mid}$ is a population of cells that are labeled with fluorescent anti-CD206 and identified as "low" on the signal intensity logicle scale dotplot and simultaneously labeled with fluorescent anti-CD66b/anti-CD3/anti-CD19 and identified as "mid" on the signal intensity logicle scale dotplot;

*FVS*510/FSC-A_{log10} [R2] is a proportion of cells with a *FVS*510 signal density greater than or equal to 0.3 but less than 0.6 ("[R2]") where the *FVS*510 signal density is a per cell *FVS*510 fluorescent signal on the logicle scale divided by the FSC-A_{log10}; and

TCPP/SSC-A_{log10} [R3] is a proportion of cells with a TCPP signal density greater than or equal to 0.6 ("[R3]") where the signal density is a per cell TCPP fluorescent signal on the logicle scale divided by the SSC-A_{log10}.

16. The non-transitory computer-readable medium of claim 15 wherein the cell viability marker labels dead cells preferentially to live cells.

17. The non-transitory computer-readable medium of claim 15 wherein a minimum number of CD206 positive cells are present in the sputum sample to be analyzed.

18. The non-transitory computer-readable medium of claim 15 wherein the bead size exclusion gate is set between 5μm and 30μm.

19. The non-transitory computer-readable medium of claim 15 wherein the automated non-debris gate excludes a majority of dead cells from a non-debris population.