

An Exhaled microRNA Panel Interrogated for Lung Cancer Case-Control Discrimination

Miao Shi

AECOM: Yeshiva University Albert Einstein College of Medicine <https://orcid.org/0000-0002-9355-025X>

Weiguo Han

The University of Arizona College of Pharmacy

Olivier Loudig

Center for Discovery and Innovation

Chirag D. Shah

AECOM: Albert Einstein College of Medicine

Jay B. Dobkin

AECOM: Albert Einstein College of Medicine

Steven Keller

Merck: Merck & Co Inc

Ali Sadoughi

AECOM: Albert Einstein College of Medicine

Changcheng Zhu

AECOM: Albert Einstein College of Medicine

Robert E. Siegel

James J Peters VA Medical Center

Maria Katherine Fernandez

Montefiore Health System

Lizzet DeLaRosa

AECOM: Albert Einstein College of Medicine

Dhruv Patel

Montefiore Health System

Aditi Desai

Montefiore Health System

Taha Siddiqui

AECOM: Albert Einstein College of Medicine

Saurabh Gombar

AECOM: Albert Einstein College of Medicine

Yousin Suh

Columbia University

Tao Wang

AECOM: Albert Einstein College of Medicine

H. Dean Hosgood

AECOM: Albert Einstein College of Medicine

Kith Pradhan

AECOM: Albert Einstein College of Medicine

Kenny Ye

AECOM: Albert Einstein College of Medicine

Simon D. Spivack (✉ Simon.spivack@einsteinmed.org)

AECOM: Albert Einstein College of Medicine

Research

Keywords: microRNA, exhaled breath condensate, lung cancer, RT-PCR

Posted Date: May 20th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-528874/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Background: An exhaled microRNA-based lung cancer case-control discriminant biomarker strategy is reported.

Methods: A microRNA-seq discovery effort compared paired tumor to non-tumor tissue, was reconciled with analogous TCGA and published literature-based tissue-discriminant microRNA data, yielding a candidate panel of 24 microRNAs that are upregulated in either adenocarcinomas and/or squamous cell carcinomas. The technical feasibility of microRNA-PCR assays in exhaled breath condensate (EBC) was tested. The airway origin of exhaled microRNAs was then topographically “fingerprinted”, using paired EBC and bronchoscopic samples. For initial EBC testing, a clinic-based case-control set of 351 individuals (166 NSCLC cases, 185 non-cancer controls) was interrogated with the 24-candidate microRNA panel by qualitative RT-PCR, and curated by melt curve analysis. Data were analyzed by both logistic regression (LR), and by random-forest (RF) models, validated by iterative resampling.

Results: Both feasibility of exhaled microRNA detection, and its origins in part from lower airway sources, were confirmed. LR models adjusted for age, sex, smoking status, pack years, quit-years, and underlying lung disease identified exhaled miR-21, 33b, 212 ($p_{adj}=0.019, 0.018, 0.033$, resp.) as case-control discriminant. For the RF analysis, the combined clinical + microRNA models showed modest added discrimination capacity (1.1–2.5%) beyond the clinical models alone: by subgroup, all subjects 1.1% ($p = 8.7e-04$); former smokers 2.5% ($p = 3.6e-05$); early stage 1.2% ($p = 9.0e-03$). Sensitivity, specificity, positive- and negative-predictive values of the clinical + microRNA models for the entire cohort were 71%-76%.

Conclusion: This work suggests that exhaled microRNAs are measurable qualitatively; reflect in part lower airway signatures; and if improved/refined, can potentially help distinguish lung cancer cases from controls.

Background

There is a consensus that the positive predictive value, efficiency, and mortality benefit of CT or other screening modalities for early lung cancer detection could be better leveraged by defining up-front an even higher risk subpopulation to screen [1–5]. Clinical risk factors of age, smoking status, and tobacco dose, when combined into sophisticated risk models [6, 7], still do not adequately capture overall risk nor define the highest risk subgroups, as most of the risk for lung cancer remains unexplained by standard clinical factor-based risk profiling [2, 8]. Therefore, the pursuit of molecular markers of risk is pivotal to improving current lung cancer screening efforts [1, 2, 8, 9], to focus on those individuals most likely to benefit. Blood-based markers have been suggested [10, 11]. Assessments of the broad epithelial field for messenger RNAs in bronchial brushings have been convincing [12, 13], and that for microRNAs suggestive [14]. Non-invasive molecular risk-assessment tools are not in clinical use at present.

This initial report describes an exhaled microRNA approach to non-invasive interrogation of the lung for the purpose of developing a lung cancer risk biomarker that is both airway compartment-derived, and population-applicable. As a first step, we derived a candidate microRNA pool, including a total of 20 upregulated microRNAs that differentiated frank human lung NSCLC tumors versus paired non-tumor tissue from the same individual surgical resections, using RNAseq efforts from our own sample sets (Supplemental data, Fig. 1), and verified them with analogous data from the TCGA [15, 16]. We also added 5 microRNAs of additional interest from the published lung cancer literature [17–20]. Next we tested the technical feasibility of detecting microRNAs in exhaled breath condensate (EBC), and then assessed the initial exhaled microRNA performance in discriminating those with non-small cell lung cancer (cases), and those without lung cancer (controls) drawn from the same clinical population of individuals destined for bronchoscopy or lung resection surgery. Starting with a robust base clinical model in the discovery set, for the three primary analyses (all categories, formers smokers, early stage), the exhaled microRNA biomarker data yield a modest 1.1–2.5% increment in case-control discrimination attributable to the addition of qualitative exhaled microRNA data, over and above clinical factor models alone.

Methods

EBC donor recruitment and sample collection

Subject recruitment: A series of 351 consenting individuals destined for lung sampling for clinical purposes (bronchoscopy or thoracic surgery) were enrolled under a protocol approved by the Einstein-Montefiore institutional review board (IRB). This observational series work was STROBE compliant [21]. This study included 166 cases of lung cancer and 185 controls without lung cancer (Table 1). It also included 4 lab volunteers and EBC was collected from every volunteer at three different timepoints. EBC (and other non-invasive airway specimen) collection occurred immediately prior to the planned bronchoscopy/thoracic surgery, to preclude procedure-induced spillage of lung materials into the EBC (and mouthwash) samples. Clinical data was obtained by direct interview in advance of any clinically-indicated bronchoscopic/surgical procedure (and therefore in advance of tissue diagnosis), and verified manually in the clinical electronic medical record. Inclusions were: age >21; fitness for the clinically-indicated (bronchoscopy/surgical) procedure; capacity and willingness to consent. Exclusions were: acute respiratory illness, contraindications to additional brushings/bronchoalveolar lavage (coagulopathy/known poorly controlled uremia); lack of capacity for consent. As such, subjects entailed a diversity of ages, ethnicities, smoking histories, clinical diagnoses, and underlying chronic lung diseases, which were accounted for in the models.

Table 1
Clinical characteristics among 351 cases and controls

	Cases (n = 166)	Controls (n = 185)	Statistics	p-value
Age (years)	66.93	56.40	t-test	1.23E-15
Gender (% male)	48.80	49.73	chi-square	0.86
Smoking Status (%)	overall		chi-square	1.51E-10
Current	43.37	20.54		
Former	47.59	41.62		
Never	9.04	37.84		
Pack Years	43.43	19.21	t-test	3.12E-09
Quit Years (former smokers)	7.33	9.40	t-test	6.00E-03
Pack years-Quit Years	31.14	5.82	t-test	7.32E-06
Tumor Histology (%)			N/A	
Adeno	50.0	N/A		
Squam	21.1	N/A		
Undiff NSCLC	15.7	N/A		
Small Cell	9.0	N/A		
Mets/Other	4.2	N/A		
Stage (%)			N/A	
I	33.13	N/A		
II	12.05	N/A		
III	31.33	N/A		
IV	11.45	N/A		
ULD (%)				
COPD	56.02	19.46	Fisher	1.29E-12
Fibrosis	1.20	2.16	Fisher	0.69
Inflammation, NOS	1.20	10.27	Fisher	2.16E-04
Asthma	12.05	16.76	Fisher	0.23
Sarcoid	1.20	8.65	Fisher	1.30E-03

	Cases (n = 166)	Controls (n = 185)	Statistics	p-value
Bronchiectasis	2.41	2.70	Fisher	1.00
None	31.33	49.19	Fisher	7.00E-04

Clinical characteristics of the case versus control subjects. Former smoker, defined as quit > 1 year; COPD, defined clinically (MD report, medications), radiographically, and/or pathologically in medical records; Pack-yrs – Quit-yrs, in former smokers, a constructed variable combining cumulative dose (pack years) minus proximity of smoking (quit-years); NOS, not otherwise specified. Adeno = adenocarcinoma; Squamous = squamous cell carcinoma; NSCLC-Undifferentiated non-small cell lung cancer; Small cell = small cell carcinoma; Mets/Other = metastases from other organs to lung or other tumor histologies. ULD = Underlying (chronic) lung disease.

EBC Sample collection

The EBC collection followed the recommendations of the American Thoracic Society/European Respiratory Society Task Force on EBC [22]. RTube™ (Respiratory Research, Inc) was used to collect patient's Exhaled Breath condensate (EBC), per standard protocol. The essentials of the simple RTube® device are (i) One way inhalation/exhalation valve; (ii) Small port for exhaled breath mixing and turbulence; (iii) Exhalation cooling chamber, polypropylene; (iv) Manually operated piston for condensate capture. Briefly, before any clinically-indicated lung procedure, subjects were equipped with RTube /mouthpiece/noseclips and performed quiet, tidal volume breathing plus one deeper breath (sigh) per minute, collected over a 10–15 minute span, while seated; saliva was to be swallowed, and excess saliva was trapped by RTube® device by design. Any coughing was instructed to be done off of the mouthpiece, to minimize oral contamination. A bare minimum of 100ul of EBC was the goal, and achieved in > 75% of subjects. Over 50% of individuals collected > 500ul EBC.

RNA extraction

For total RNA extraction, EBC was concentrated by ethanol precipitation and then was purified by Trizol (Invitrogen) per manufacturer protocol and lab optimized protocol. The following components were added into a capped polypropylene tube and thoroughly mixed, including 100–400 ul of EBC sample, 40 ul of 3M sodium acetate (pH 5.5), 5 ul of 5 ug/ul glycogen carrier, and 1100 ul of 100% cold ethanol. The mixture was chilled at -80 °C for 30 min and then centrifuged at 14,000 rpm for 20 min at 4 °C. Then, the supernatant was discarded and the pellet was rinsed with cold 70% ethanol twice, and air-dried. The pellet was then dissolved in 0.5 ml of Trizol®. Total RNA was purified per the Trizol® manufacturer protocol. The RNA pellet was dissolved in 15 ul of RNase-free water.

microRNA PCR analysis

The overall strategy was to amplify mature microRNAs by a previously published lab protocol involving poly-A tailing using a one-base anchored and tagged oligo-dT-RT strategy, and a microRNA-specific

forward primer coupled to a universal, unique tag-specific reverse primer, in aggregate precluding false gDNA amplification [23–25]. Individual steps and details follow.

Cell culture samples used for microRNA PCR development

For positive controls in microRNA-PCR assay development, a set of cell lines including NHBE, HBEC, A549, Hela, HTB-119 and CRL-1995 was RNA extracted in conventional column (RNEasy, Qiagen), and provided a stock solution of total RNA for initial testing of microRNA-specific primers.

Poly(A) Tailing

The Poly(A) Tailing Kit (Ambion) was used to polyadenylate the 3' termini of microRNA. First, ATP was diluted to 1% of the original concentration. Then, the following components were added into a PCR tube and thoroughly mixed, including 2 ul of 5x buffer, 0.8 ul of $MnCl_2$ (25 mM), 0.4 ul of diluted ATP, 0.25 ul of enzyme and 6.55 ul of total RNA from EBC. The mixture was incubated at 37°C for 30 min.

Reverse transcription. Reverse transcription was performed with 10 μ l of the *E. coli* Poly(A) Polymerase (E-PAP) treated total RNA using Superscript III reverse transcriptase (Invitrogen) as follows. RNA template was added to a master mix containing 1 μ l of 100 μ M universal oligo-dT-adapted universal RT primer [25], 1 μ l of dNTP mix (each base 10 μ M) and 1 μ l of DNase/RNase-free water. Total volume was adjusted to 13 μ l with DNase/RNase-free water. The solution was incubated at 65°C for 5 min and then cooled on ice. A master mix containing 4 μ l of 5X first-strand buffer, 1 μ l of 0.1 mM DTT, 1 μ l RNaseOUT (Invitrogen) and 1 μ l SuperScript III per RT sample was prepared and added to each sample. The samples were incubated at 42°C for 30 min, 50°C for 30 min, followed by 70°C for 15 min.

Realtime PCR. Typically, the RT reaction was diluted 1:20 and 2 μ l used in the realtime PCR of microRNAs with the transcript specific forward PCR primers (Supplementary Table 3, n = 25 primersets) and a matched (tag-directed) reverse primer. cDNA template was added to a master mix containing 10 ul of 2x PowerSYBR green master mix (Applied Biosystem), 1 ul of 10 uM primers mix and 7 ul of DNase/RNase-free water. The reaction was incubated in an Applied Biosystems 7500 realtime PCR system at 95 °C for 10 min, followed by 45 cycles of 95 °C for 15 s, 60 °C for 15 s and 72 °C for 32 s. After that, dissociation stage/melting curve analysis was performed. In developing each primerset, primers were designed to produce a single unique melting curve on known microRNA extracts from lung cell lines. Multiple separate positive and negative controls in both lung cell lines and EBC sample standards were run, including (a) gDNA spike (to exclude false gDNA amplification) (b) no-RTase (to exclude false gDNA amplification); (c) no poly Adenylation (to exclude false messenger RNA amplification); (d) no template (to exclude reagent contamination by PCR product).

Data cleaning/scoring

Since microRNAs are all of near-identical size, base composition/melting temperature was a major distinguishing feature. The criteria for including or excluding a micro-RNA-derived PCR product as present

were extracted from the melting curves. If a sample had the same melting curve maximum temperature (T_m) as the positive control from cell lines for that microRNA primerset, it was called “positive”. If a reaction sample had no visible melting curve, or the visible melting curve displayed greater than $\pm 1.5^\circ\text{C}$ different T_m from the melting curve from the positive, individual miR-specific control, it was called “negative”. We used one convention for overall scoring of samples – at least one of two replicates must be positive. The housekeeper control chosen, based on literature, and ubiquitous presence in our EBC samples, was miR-423-3p [26, 27]. From previously described studies [28–30], hsa-miR-16, hsa-miR-26b, hsa-miR-92, hsa-miR-423, hsa-miR-374, are often used as housekeeper controls.

Statistical Analysis

Logistic Regression (LR)

Logistic Regression was performed for each miRNA with cancer case-control status as the response, with and without the clinical variables included as the covariates. The clinical covariates are age, gender, smoking status (never smokers, former smokers, current smokers), pack years, quit years, and underlying lung diseases (categorized in three groups, (1) any of COPD, fibrosis, generic inflammation and/or asthma; (2) sarcoid and bronchiectasis; (3) none and others.

Random Forests (RF): Two types of Random Forest [31, 32] classifiers were built for comparison, using R package random forest [32]. First, an RF classifier was built on the clinical variables alone: age, gender, smoking status, pack-years, quit-years, underlying lung disease (type), tumor histology, stage. Two-fold cross-validation was repeated 20 times to gauge the accuracy of this classifier, and its sensitivity, specificity, positive and negative predictive value. Second, an RF classifier was built on the clinical variable plus the microRNA variables together. To compare the performance of the two types of RF classifiers, we further generated 100 resampled ROC curves for each one and compared the average area under the curve (AUC) between the two models using a two-independent sample t-test. A resampled ROC curve was generated by repeatedly splitting the dataset into 50% training, 50% testing (100 times), building the two random forest models (clinical and clinical + microRNA), and predicting the outcomes of the testing split.

Airway topography similarity statistic

A subset of 12 EBC donors provided bronchoscopic samples of deep alveolar (BAL) and major airway (bronchial, BB) levels, as well as sputum, mouthrinse and other specimens. The pilot sub-study (Supplemental Table 4) was designed to evaluate if an individual microRNA profile from EBC retains the distinct features of the microRNA profile from deep lung (bronchial brushings or bronchoalveolar lavage), or alternately resembles contaminating upper airway/mouthwash tissues. This was done by applying an arbitrary panel of 13-microRNAs interrogated by qualitative RT-PCR against samples from 12 individuals, each donating five airway level samples for comparison [bronchoalveolar lavage (BAL), bronchial brush (BB), sputum (SP), mouthrinse (MW), EBC]. To statistically test the surrogacy of EBC-microRNA for deeper lung specimens (bronchial brushings and bronchoalveolar lavage), we developed a similarity statistic of

two tissue types based on Hamming distance. That is, where and are (binary) miRNA profiles from two tissue types of the same individual i . The Hamming distance H gives the total number of miRNAs for which the two profiles d and d' are discordant. The smaller the statistic SH is, the more similar are two tissue types in miRNA profiles within each subject. If the two tissue types from the same individual are not closer than two tissue types from two random individuals, then there is no information in one of the tissues to infer the miRNA profile of the other tissue. To test that the two tissues from the same individuals are closer than two random individuals, we performed a permutation test that permutes the miRNA profiles within each tissue type among individuals.

Temporal stability of EBC miRNA for an individual across time

The EBC samples from 4 lab volunteers at three different timepoints were used to evaluate temporal stability of EBC miRNA for an individual across time. The three targets were miR-141, miR-142-3p and miR-205 and the housekeeping gene is miR-423-3p. Heat matrix was built by delta Ct (normalize Ct of target miRNA to Ct of housekeeping miR-423-3p).

Results

The clinical characteristics of the 351 subjects are described in Table 1. Baseline clinical characteristics that differed between cases and controls included age, smoking status, pack-years, quit years, underlying lung disease. Former smokers were defined as quit greater than one year from enrollment. Cases significantly differed from controls for: age (66.9 vs. 56.4, resp.); smoking status (current 43.4 vs. 20.5%, former 47.6% vs. 41.6%, never 9.0 vs. 37.8% never smokers); pack-years among current/former smokers (43.4 vs. 19.2); quit years among former smokers (7.3 vs. 9.4), pack years-quit years index (31.1 vs. 5.8); Underlying lung disease including COPD % (56.0 vs. 19.5); inflammation NOS% (1.2 vs. 10.3); sarcoidosis% (1.2 vs 8.6); none% (31.8 vs. 49.2). Both logistic regression (LR) and random forest (RF) discriminant models took these clinical inter-group differences into account. For RF, this included measuring the incremental impact on case-control discrimination of microRNAs over and above these clinical factors alone.

EBC surrogacy for the lung

The similarity statistic of two tissue types that were based on Hamming distance,, and 1000 permutations of miRNA profiles within each tissue type among individuals, gave an estimated P-value of 0.007, suggesting that the miRNA profiles of EBC are closer to miRNA profiles of BAL of the same individual than to miRNA profiles of BAL of random individuals. The same analysis was applied between EBC and BB ($p = 0.23$), EBC and SP ($p = 0.18$), EBC and MW ($p = 0.04$).

Logistic regression: LR models were created (Table 2), using individual exhaled microRNA presence or absence as univariate predictors of case-control status, with adjustment for clinical factors: age, gender, smoking status (current, former, never), smoking pack years and quit years, and presence of underlying

lung disease. For the entire data set, miR-21, 33b and 212 appeared to be somewhat informative for case-control status ($p < 0.05$), after adjustment for the above-listed clinical factors.

Table 2
 Logistic regression, univariate miR, All subjects, n = 351

##	miRNA	p	p.adj
1	miR.324.5p	0.967	0.817
2	miR.9	0.383	0.779
3	miR.21	0.090	0.020
4	miR.31	0.343	0.231
5	miR.33b	0.011	0.017
6	miR.96	0.631	0.342
7	miR.105	0.677	0.142
8	miR.146a.5p	0.358	0.340
9	miR.182.5p	0.840	0.687
10	miR.196b	0.385	0.587
11	miR.199b.5p	0.640	0.562
12	miR.200a	0.799	0.396
13	miR.200b	1.000	0.959
14	miR.205	0.664	0.793
15	miR.212	0.153	0.033
16	miR.221	0.932	0.386
17	miR.345	0.113	0.081
18	miR.429	1.000	0.601
19	miR.767	0.236	0.476
20	miR.944	0.053	0.404
21	miR.1269a	0.059	0.496
22	miR.1293	0.102	0.230
23	miR.1910	0.261	0.154
24	miR.3662	0.862	0.539

The models included adjustments for clinical factors of age, gender, smoking status (current, former, never), smoking pack-years and quit-years, and underlying lung disease. Underlying lung disease: For all models, underlying lung disease was treated as trichotomous (COPD /fibrosis//inflammation NOS, asthma) *versus* sarcoidosis/bronchiectasis) *versus* none/other. Housekeeper miR423-5p, not listed.

Random Forests: Clinical, exhaled microRNA, and combined clinical + exhaled microRNA RF models discriminating cases from controls were constructed (Table 3). For lung cancer overall, including all subjects (n = 351) and all case primary lung malignant tumor histologies, the clinical RF model included age, gender smoking status, pack-years, quit-years, underlying lung disease. For the *clinical only* RF model alone, case-control discriminant accuracy, sensitivity, specificity, positive predictive value, negative predictive value, AUC-ROC, were: 0.74, 0.74, 0.74, 0.76, 0.72, 0.814, respectively. For the *microRNAs only* model, the respective values were: 0.57, 0.63, 0.50, 0.58, 0.55, 0.611. For the *combined clinical + microRNA* model, the respective performance values were: 0.74, 0.74, 0.74, 0.76, 0.73, 0.826. The added AUC discrimination conferred by exhaled microRNAs for the overall group of subjects (n = 351) was 1.2% (0.814 = > 0.826; p = 0.07, Welch t-test).

Table 3
Exhaled microRNA RF models

Table 3. RF models, lower stringency	Individual component factors	Accuracy (p-value)	Sensi, Speci	PPV, NPV	ROC-AUC	AUC difference, Clinical vs Clinical + micro-RNA, % (p-value)
All subjects, all smoking categories, all tumor histologies, n = 166 cases, 185 controls						
Clinical Variables Alone (unselected)	All clinical variables [age, gender, smoking status, pack-years, quit-years, underlying lung disease, tumor histology for cases]	0.74 ($< 2.2e-16$)	0.74, 0.74	0.76, 0.72	0.814	
microRNAs alone	All 24 microRNAs.	0.57	0.63,	0.58,	0.611	
	Important miRs: 21, 33b, 944, 1269a, 1910.	($< 2.2e-16$)	0.50	0.55		
Clinical + microRNA	All Clinical factors and All 24 miRs	0.74 ($2e-16$)	0.74, 0.74	0.76, 0.73	0.826	1.2% (0.07)
Former smokers only n = 79cases, 77 controls						
Clinical Variables Alone	All clinical variables [age, gender, smoking status, pack-years, quit-years, underlying lung disease]	0.69 ($< 2.2e-16$)	0.69 0.69	0.69 0.70	0.777	
microRNAs alone	All 24 microRNAs	0.59	0.57	0.59	0.656	
	Important miRs: 33b, 146a.5p, 200a, 212, 1293.	($< 2.0e-16$)	0.61	0.59		
Clinical + microRNA	All Clinical and All miRs	0.70 ($< 2.0e-16$)	0.67 0.72	0.70 0.69	0.807	3.0% ($6.0e-03$)

Underlying lung disease (trichotomous: (3). COPD or fibrosis or inflammation or asthma; versus (2). sarcoid bronchiectasis; versus (1) none/other)

Table 3. RF models, lower stringency	Individual component factors	Accuracy (p-value)	Sensi, Speci	PPV, NPV	ROC-AUC	AUC difference, Clinical vs Clinical + micro-RNA, % (p-value)
Early Stage only (stages I and II) n = 78 cases, 184 controls						
Clinical Variables Alone	All clinical variables [age, gender, smoking status, pack-years, quit-years, underlying lung disease]	0.75 (2.2e-16)	0.86 0.50	0.80, 0.60	0.806	
microRNAs alone	All 24 microRNAs. Important miRs: 96, 146a.5p, 944, 1269a, 1910.	0.700 (NS)	0.90, 0.23	0.73 0.49		
Clinical + microRNA	All Clinical variables and All miRs	0.76 ($< 2.2e-16$)	0.90, 0.45	0.79 0.65	0.828	(2.2% (5.1e-03))
Former Smoker x Early Stage Sub-subgroup n = 34 cases, 77 controls						
Clinical Variables Alone	All clinical variables [age, gender, smoking status, pack-years, quit-years, underlying lung disease, tumor histology for cases]	0.71 (1.9e-03)	0.87 0.35	0.75 0.54	0.714	
microRNAs alone	All 24 microRNAs. Important miRs: 96, 146a.5p, 200b, 345, 1910	0.67 (NS)	0.86 0.25	0.72 0.44	0.641	
Clinical + microRNA	All Clinical and All miRs	0.71 (1.4e-02)	0.90 0.28	0.74 0.54	0.738	2.4% (NS)

Underlying lung disease (trichotomous: (3). COPD or fibrosis or inflammation or asthma; versus (2). sarcoid bronchiectasis; versus (1) none/other)

Table 3. RF models, lower stringency	Individual component factors	Accuracy (p-value)	Sensi, Speci	PPV, NPV	ROC-AUC	AUC difference, Clinical vs Clinical + micro-RNA, % (p-value)
Current Smokers only, n = 38cases, 72 controls						
Clinical Variables Alone	All clinical variables [age, gender, smoking status, pack-years, quit-years, underlying lung disease, tumor histology for cases]	0.78 (2.2e-16)	0.59 0.87	0.71 0.80	0.731	
microRNAs alone	All 24 microRNAs. Important miRs: 105, 146a.5p, 182.5p, 200a, 205.	0.59 (NS)	0.15 0.82	0.30 0.65	0.464	
Clinical + microRNA	All Clinical and All miRs	0.76 (2.2e-16)	0.48 0.90	0.73 0.77	0.764	3.3% (3.5e-02)
Adenocarcinoma only, n = 87cases, 184 controls						
Clinical Variables Alone	All clinical variables [age, gender, smoking status, pack-years, quit-years, underlying lung disease]	0.74 ($< 2.2e-16$)	0.83 0.53	0.79 0.60	0.817	
microRNAs alone	All 24 microRNAs. Important miRs: 96, 146a.5p, 221, 944, 1269a	0.65 (NS)	0.87 0.18	0.69 0.40	0.579	
Clinical + microRNA	All Clinical and All miRs	0.73 (2.2e-16)	0.87 0.43	0.76 0.61	0.796	-2.1% (1.1e-02), neg direction
Late Stage (III, IV) only, n = 77cases, 184 controls						

Underlying lung disease (trichotomous: (3). COPD or fibrosis or inflammation or asthma; versus (2). sarcoid bronchiectasis; versus (1) none/other)

Table 3. RF models, lower stringency	Individual component factors	Accuracy (p-value)	Sensi, Speci	PPV, NPV	ROC-AUC	AUC difference, Clinical vs Clinical + micro-RNA, % (p-value)
Clinical Variables Alone	All clinical variables [age, gender, smoking status, pack-years, quit years, underlying lung disease, underlying lung disease]	0.74 (2.2e-16)	0.84 0.49	0.80 0.56	0.797	
microRNAs alone	All 24 microRNAs. Important miRs: 31, 33b, 105, 212, 944	0.67 (NS)	0.90 0.12	0.71 0.35	0.541	
Clinical + microRNA	All Clinical and All miRs	0.74 (2.2e-16)	0.89 0.38	0.77 0.58	0.809	1.2% (NS)
Late Stage (III, IV) x Former Smoker, n = 40 cases, 77 controls						
Clinical Variables Alone	All clinical variables [age, gender, smoking status, pack-years, quit years, underlying lung disease, underlying lung disease] <i>histologies?</i>	0.70 (1.6e-13)	0.82 0.46	0.75 0.57	0.781	
microRNAs alone	All 24 microRNAs. Important miRs: 33b, 200a, 212, 345, 1293	0.63 (NS)	0.84 0.22	0.67 0.41	0.654	
Clinical + microRNA	All Clinical and All miRs	0.69 (1.1e-07)	0.85 0.37	0.72 0.57	0.789	0.8% (NS)
<i>Underlying lung disease (trichotomous: (3). COPD or fibrosis or inflammation or asthma; versus (2). sarcoid bronchiectasis; versus (1) none/other)</i>						

For *a priori* selected subgroups, data analyses are also tabulated (Table 3). For example, *Former smoker* cases versus former smoker controls comparison of case-control discriminant performance were again described in terms of discriminant accuracy, sensitivity, specificity, positive predictive value, negative predictive value, AUC-ROC. For the *clinical only* model, performance parameters were: 0.69, 0.69, 0.69, 0.69, 0.70, 0.777, respectively. For the *microRNA-only* model, performance were: 0.59, 0.57, 0.61, 0.59, 0.59, 0.656, respectively. For the *combined clinical + microRNA* model, performance were: 0.70, 0.67, 0.72,

0.70, 0.69, 0.807, respectively. The added AUC discrimination conferred by exhaled microRNAs was 3.0% ($0.777 = > 0.807$; $p = 6.0e-03$, Welch t-test) for former smokers. Similarly, early stage (I + II combined) cases showed 2.2% added case-control AUC discrimination from the exhaled microRNA panel ($p = 5.1e-03$). For additional clinically important combined subgroups, the case versus controls models' performance characteristics are described in Table 3.

Temporal stability of EBC miRNA for an individual across time

Three target miRNAs (miR-141, miR-142-3p and miR-205) in EBC of three different timepoints of four individuals were detected by realtime PCR and normalized to housekeeping miR-423-3p. It shows the EBC samples from different timepoints of the same subject were stable to a large extent (Fig. 2).

TABLES:

Discussion

This report entails the most comprehensive interrogation of microRNAs in exhaled breath, here uniquely performed to distinguish subjects with and without primary lung cancers [33]. Starting with a lung tissue microRNA-seq discovery effort combined with published literature-suggested microRNAs, we interrogated a panel of 25 microRNAs in exhaled breath condensate using our RNA-specific qualitative RT-PCR. We found that: (i) microRNAs are detectable in exhaled breath condensate; (ii) there are individual exhaled microRNAs that offer case-control discrimination by logistic regression (microRNAs 21, 33b, 212), and (iii) additional RF models can be developed, using the entire microRNA panel, that also suggest some modest additional case-control discrimination, particularly in the subsets of former smoker, and early stage subjects, over and above that demonstrated in comprehensive clinical models.

Technical challenges abound in examining nucleic acids in EBC. While EBC is widely available non-invasively, this specimen entails only trace levels of microRNA template. This is perhaps because the templates are by definition, higher in molecular weight (22 nucleotides in length) than is typically true for exhaled airstream-suspended molecules such as H_2O_2 , 8-isoprostane, and others [22]. Nonetheless, the PCR confers capacity for detection of microRNAs at the low template copy level, as is suggested here. The trace concentrations inherent to EBC specimens for most analytes, including nucleic acids has, to date, precluded performing discovery efforts such as microRNA next gen sequencing, directly from this matrix.

The microRNA interrogation panel choice was therefore based on: (i) a previously unpublished microRNA seq effort (GEO#: GSE33858) inter-tissue comparison of 32 lung resected bronchogenic carcinoma versus remote lung tissue. (stratified for adenocarcinoma, squamous cell carcinoma histologies), with 10 representative overexpressed microRNAs included from each of those two histologies. The remainder came from: (ii) TCGA [15, 16]; and (iii) several literature-identified microRNA markers of lung cancer.

We used our previously published microRNA-PCR that is micro/mRNA – specific, as it excludes gDNA fragment false priming by employing a uniquely tagged RT-primer strategy [23, 25], and in primer design precluded false amplification of messenger RNA fragments. We chose to treat the data as qualitative (individual miR, present/absent) because we were insufficiently confident of robust quantitative RT-PCR data that could not be reliably scaled to a robustly quantifiable internal housekeeper at these trace levels. Performance of the fluorescent intercalating (SYBR®) dye detection strategy coupled to URT-PCR on the realtime PCR platform allowed quality assurance using quantitation curve, melt-curve, melt temperature with each PCR reaction. This was superimposed on a series of other analyses invoked during primer design, using multiple positive and negative controls, described in the Methods and Additional Studies sections. We additionally piloted a commercial qPCR platform (In Vitrogen/Taqman®) without readily apparent additional precision nor microRNA-PCR sensitivity.

This cross-sectional case-control design was chosen as representing a typical initial step in early development of potential risk biomarkers [34, 35]. Clinical-demographic differences were observed in cases versus controls for age, smoking, pack-years, quit years, a pack-years minus quit-years composite index, underlying lung disease (COPD, inflammation/fibrosis, asthma, sarcoidosis, bronchiectasis). However, these differences were equally modelled in both clinical-only models and in the clinical + microRNA combined models identically, so they should not have biased the incremental microRNA-attributable risk prediction. We emphasized current and former smokers predominantly, as they are at elevated risk for lung cancer, and therefore commonly come to clinical attention for surveillance, biopsy/resection, and thus were considered appropriately efficient for enrollment in this initial study. Our case and control ascertainment was crisp, minimizing misclassification as subjects were all confirmed histologically by virtue of their bronchoscopic/surgical procedures, underwent further verification of case and control status by an additional 3–6 month period of clinical follow-up, facilitated by electronically-retrieved clinical assessments from the engaged clinical pathologists, radiologists, surgeons, and pulmonologists on each subject. Recruited subjects with disputed case-control ascertainment (< 1% of enrolled) were excluded from the study.

In this moderate size case-control subject set, with an already selected candidate 24-microRNA panel, we initially performed logistic regression, using case-control status as the main outcome variable, and a clinical model tested with/without each individual miR on the panel. Separately, we then employed iterative cross validation by random forests to assure stability of our results, rather than separate discovery and validation sets. The RF approach iteratively and randomly splits the data, substantively cross validating in truly random fashion, and minimizing over-fit.

The clinical versus clinical-microRNA incremental differences are admittedly modest (~ 0.0–3.0%). We surmise that this is, in part, due to the strength of the clinical model alone displaying ROCs ~ 0.75-80. These were unusually robust clinical models, we believe for two reasons. First is the clinical model comprehensiveness, in part attributable to inclusion of all major known substantive risk factors for lung cancer (including quit years, underlying lung disease, others). Secondly, there is positive selection inherent to enrolling clinical bronchoscopy and surgical subjects such as these (above), wherein both (case and

control) sets of subjects are drawn from the same base (procedural-destined) population that is itself selected on clinical criteria to be at high risk for lung cancer. By definition, that high risk is perceived by the clinician as sufficient to warrant an invasive diagnostic/therapeutic procedure, the enrollment point for a majority of our subjects. Both of these factors (clinical model comprehensiveness, and clinical series enrollment bias) contribute to high risk in this clinical series, and imply that clinical risk model performance will be elevated. Thus, the difference between this comprehensive clinical model alone, and that for this clinical model plus microRNA could potentially be artificially narrowed (as compared to that using conventional sparse clinical models) by virtue of the comprehensiveness of the clinical model. We believe the negative impact of such bias on the estimate of the actual contribution of exhaled microRNAs to case-control discrimination, is counter-balanced by the strength inherent in using the same (robust) clinical model when comparing clinical-only models versus combined clinical + microRNA models. Additionally, the definitive diagnoses inherent in recruiting those destined for lung sampling/pathologic readout was another strength. Overall, then, the above considerations suggest ours is a conservative estimate of the exhaled biomarker contribution in real clinical conditions.

Among study limitations, we were forced to use a dichotomous (present/absent) signal for a given microRNA in a given EBC sample, despite being run on a realtime machine, for technical reasons. The realtime CT values, using the chosen platform, were not robust enough to generate reliable quantitative data, worth re-addressing in optimization studies, which are ongoing.

Additionally the discriminant microRNA signal may in fact be small in magnitude, as our data suggest. This small magnitude of microRNA change in the “field” of bronchial epithelium itself has been suggested in a comprehensive RNA-seq study of bronchial brushings in a similar case control setting [14]. Notably, of the discriminant four bronchial epithelial case-control discriminant microRNAs in that report, only one (146-5p) was interrogated in our study. While 146-5p was not individually case-control discriminant in LR models, it was contributory in the RF models for former smokers and early stage. There is a very recent pilot report that EBC miRNAs might allow the identification, stratification and monitoring of lung cancer [31].

We set out to survey the “state of the epithelium” rather than detection of a small peripheral tumor itself. This view of broad epithelial “field” interrogation is appropriate to risk assessment, rather than that of a suspect tumor diagnostic tool. That the signal was likely from the field of normal cell material, rather than spillage of a tumor is supported by the observation that early stage subset showed more case-control discrimination than the late stage cases, which would not be expected if the tumor itself was spilling microRNA material.

Conclusions

In conclusion, this is one of the first reports of exhaled microRNAs in lung cancer. Given the technical demands of this application, we plan to refine the exhaled microRNA interrogation technique, including miR-PCR quantitation, and microRNA panel adjustments, to better serve case-control discrimination.

Assuming improved performance with these refinements, risk assessment efforts can be pursued in future prospective cohorts. Such trials could evaluate whether the biomarker platform can predict future events, the “mother lode” of risk assessment [5]. If such utility was demonstrated, it would then allow for actionable clinical interventions, such as focusing early detection, or alternately perhaps directing chemoprevention, onto those individuals at highest risk.

Abbreviations

TCGA: The Cancer Genome Anatomy project. NCBI/NCI/NIH.

AUC: Area under curve, for receiver operating curve (ROC).

ROC: Test performance plots sensitivity versus specificity.

Declarations

Ethics approval and consent to participate

See the attachment.

Consent for publication

Not applicable.

Availability of data and materials

All data generated or analyzed during this study are included in this published article [and its supplementary information files].

Competing interests

The author(s) declare that they have no competing interests.

Funding

NIH R21 CA192168 (SDS); NIH [K24 CA139054](#) (SDS); Department of Defense, LCRP-Expansion Award (SDS)

Authors' contributions

MS and SDS conceived the study. MS, WH, HDH and SDS designed the study. CDS, JBD, SK, AS, MKF, LD, DP, AD and TS recruited subjects. SG, YS and TW performed and analyzed the microRNA-seq analyses. MS and WH performed the microRNA-PCR Assays. KP, KY and SDS performed the PCR based data analyses: MS and SDS wrote the manuscript.

All authors read and approved the manuscript.

Acknowledgements

We thank Einstein Epigenomics Center for sequencing our microRNA libraries.

Literature References

1. Gould MK: **Clinical practice. Lung-cancer screening with low-dose computed tomography.** *N Engl J Med* 2014, **371**(19):1813-1820.
2. Bach PB, Mirkin JN, Oliver TK, Azzoli CG, Berry DA, Brawley OW, Byers T, Colditz GA, Gould MK, Jett JR *et al.* **Benefits and harms of CT screening for lung cancer: a systematic review.** *JAMA* 2012, **307**(22):2418-2429.
3. Tan BB, Flaherty KR, Kazerooni EA, Iannettoni MD: **The solitary pulmonary nodule.** *Chest* 2003, **123**(1 Suppl):89S-96S.
4. Gould MK, Donington J, Lynch WR, Mazzone PJ, Midthun DE, Naidich DP, Wiener RS: **Evaluation of individuals with pulmonary nodules: when is it lung cancer? Diagnosis and management of lung cancer, 3rd ed: American College of Chest Physicians evidence-based clinical practice guidelines.** *Chest* 2013, **143**(5 Suppl):e93S-e120S.
5. Mazzone PJ, Sears CR, Arenberg DA, Gaga M, Gould MK, Massion PP, Nair VS, Powell CA, Silvestri GA, Vachani A *et al.* **Evaluating Molecular Biomarkers for the Early Detection of Lung Cancer. When Is a Biomarker Ready for Clinical Use? An Official American Thoracic Society Policy Statement.** *Am J Respir Crit Care Med* 2017, **196**(7):e15-e29.
6. Raji OY, Duffy SW, Agbaje OF, Baker SG, Christiani DC, Cassidy A, Field JK: **Predictive accuracy of the Liverpool Lung Project risk model for stratifying patients for computed tomography screening for lung cancer: a case-control and cohort validation study.** *Ann Intern Med* 2012, **157**(4):242-250.
7. D'Amelio AM, Jr., Cassidy A, Asomaning K, Raji OY, Duffy SW, Field JK, Spitz MR, Christiani D, Etzel CJ: **Comparison of discriminatory power and accuracy of three lung cancer risk models.** *Br J Cancer* 2010, **103**(3):423-429.
8. Tanoue LT, Tanner NT, Gould MK, Silvestri GA: **Lung cancer screening.** *Am J Respir Crit Care Med* 2015, **191**(1):19-33.
9. Kanodra NM, Silvestri GA, Tanner NT: **Screening and early detection efforts in lung cancer.** *Cancer* 2015, **121**(9):1347-1356.
10. Nadal E, Truini A, Nakata A, Lin J, Reddy RM, Chang AC, Ramnath N, Gotoh N, Beer DG, Chen G: **A Novel Serum 4-microRNA Signature for Lung Cancer Detection.** *Sci Rep* 2015, **5**:12464.
11. Wozniak MB, Scelo G, Muller DC, Mukeria A, Zaridze D, Brennan P: **Circulating MicroRNAs as Non-Invasive Biomarkers for Early Detection of Non-Small-Cell Lung Cancer.** *PLoS One* 2015, **10**(5):e0125026.

12. Silvestri GA, Vachani A, Whitney D, Elashoff M, Porta Smith K, Ferguson JS, Parsons E, Mitra N, Brody J, Lenburg ME et al: **A Bronchial Genomic Classifier for the Diagnostic Evaluation of Lung Cancer.** *N Engl J Med* 2015, **373**(3):243-251.
13. Spira A, Beane JE, Shah V, Steiling K, Liu G, Schembri F, Gilman S, Dumas YM, Calner P, Sebastiani P et al: **Airway epithelial gene expression in the diagnostic evaluation of smokers with suspect lung cancer.** *Nat Med* 2007, **13**(3):361-366.
14. Pavel AB, Campbell JD, Liu G, Elashoff D, Dubinett S, Smith K, Whitney D, Lenburg ME, Spira A: **Alterations in Bronchial Airway miRNA Expression for Lung Cancer Detection.** *Cancer Prev Res (Phila)* 2017, **10**(11):651-659.
15. **Comprehensive molecular profiling of lung adenocarcinoma.** *Nature* 2014, **511**(7511):543-550.
16. **Comprehensive genomic characterization of squamous cell lung cancers.** *Nature* 2012, **489**(7417):519-525.
17. Liu ZL, Wang H, Liu J, Wang ZX: **MicroRNA-21 (miR-21) expression promotes growth, metastasis, and chemo- or radioresistance in non-small cell lung cancer cells by targeting PTEN.** *Mol Cell Biochem* 2013, **372**(1-2):35-45.
18. Liu X, Sempere LF, Ouyang H, Memoli VA, Andrew AS, Luo Y, Demidenko E, Korc M, Shi W, Preis M et al: **MicroRNA-31 functions as an oncogenic microRNA in mouse and human lung cancer cells by repressing specific tumor suppressors.** *J Clin Invest* 2010, **120**(4):1298-1309.
19. Li Y, Zhang D, Chen C, Ruan Z, Huang Y: **MicroRNA-212 displays tumor-promoting properties in non-small cell lung cancer cells and targets the hedgehog pathway receptor PTCH1.** *Mol Biol Cell* 2012, **23**(8):1423-1434.
20. Garofalo M, Di Leva G, Romano G, Nuovo G, Suh SS, Ngankee A, Taccioli C, Pichiorri F, Alder H, Secchiero P et al: **miR-221&222 regulate TRAIL resistance and enhance tumorigenicity through PTEN and TIMP3 downregulation.** *Cancer Cell* 2009, **16**(6):498-509.
21. Vandenbroucke JP, von Elm E, Altman DG, Gotzsche PC, Mulrow CD, Pocock SJ, Poole C, Schlesselman JJ, Egger M: **Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): explanation and elaboration.** *Epidemiology* 2007, **18**(6):805-835.
22. Horvath I, Hunt J, Barnes PJ, Alving K, Antczak A, Baraldi E, Becher G, van Beurden WJ, Corradi M, Dekhuijzen R et al: **Exhaled breath condensate: methodological recommendations and unresolved questions.** *Eur Respir J* 2005, **26**(3):523-548.
23. Hurteau GJ, Spivack SD: **mRNA-specific reverse transcription-polymerase chain reaction from human tissue extracts.** *Anal Biochem* 2002, **307**(2):304-315.
24. Hurteau GJ, Carlson JA, Spivack SD, Brock GJ: **Overexpression of the microRNA hsa-miR-200c leads to reduced expression of transcription factor 8 and increased expression of E-cadherin.** *Cancer Res* 2007, **67**(17):7972-7976.
25. Hurteau GJ, Spivack SD, Brock GJ: **Potential mRNA degradation targets of hsa-miR-200c, identified using informatics and qRT-PCR.** *Cell Cycle* 2006, **5**(17):1951-1956.

26. Babion I, Snoek BC, van de Wiel MA, Wilting SM, Steenbergen RDM: **A Strategy to Find Suitable Reference Genes for miRNA Quantitative PCR Analysis and Its Application to Cervical Specimens.** *J Mol Diagn* 2017, **19**(5):625-637.
27. Link F, Krohn K, Schumann J: **Identification of stably expressed housekeeping miRNAs in endothelial cells and macrophages in an inflammatory setting.** *Sci Rep* 2019, **9**(1):12786.
28. Lange T, Stracke S, Rettig R, Lendeckel U, Kuhn J, Schluter R, Rippe V, Endlich K, Endlich N: **Identification of miR-16 as an endogenous reference gene for the normalization of urinary exosomal miRNA expression data from CKD patients.** *PLoS One* 2017, **12**(8):e0183435.
29. Torres A, Torres K, Wdowiak P, Paszkowski T, Maciejewski R: **Selection and validation of endogenous controls for microRNA expression studies in endometrioid endometrial cancer tissues.** *Gynecol Oncol* 2013, **130**(3):588-594.
30. Sauer E, Babion I, Madea B, Courts C: **An evidence based strategy for normalization of quantitative PCR data from miRNA expression analysis in forensic organ tissue identification.** *Forensic Sci Int Genet* 2014, **13**:217-223.
31. Perez-Sanchez C, Barbarroja N, Pantaleao LC, Lopez-Sanchez LM, Ozanne SE, Jurado-Gamez B, Aranda E, Lopez-Pedreira C, Rodriguez-Ariza A: **Clinical Utility of microRNAs in Exhaled Breath Condensate as Biomarkers for Lung Cancer.** *Journal of Personalized Medicine* 2021, **11**(2).
32. Breiman L: **Random forests.** *Machine Learning* 2001, **45**(1):5-32.
33. Mozzoni P, Banda I, Goldoni M, Corradi M, Tiseo M, Acampa O, Balestra V, Ampollini L, Casalini A, Carbognani P *et al*: **Plasma and EBC microRNAs as early biomarkers of non-small-cell lung cancer.** *Biomarkers* 2013, **18**(8):679-686.
34. Pepe MS, Etzioni R, Feng Z, Potter JD, Thompson ML, Thornquist M, Winget M, Yasui Y: **Phases of biomarker development for early detection of cancer.** *J Natl Cancer Inst* 2001, **93**(14):1054-1061.
35. Pepe MS, Feng Z, Janes H, Bossuyt PM, Potter JD: **Pivotal evaluation of the accuracy of a biomarker used for classification or prediction: standards for study design.** *J Natl Cancer Inst* 2008, **100**(20):1432-1438.

Figures

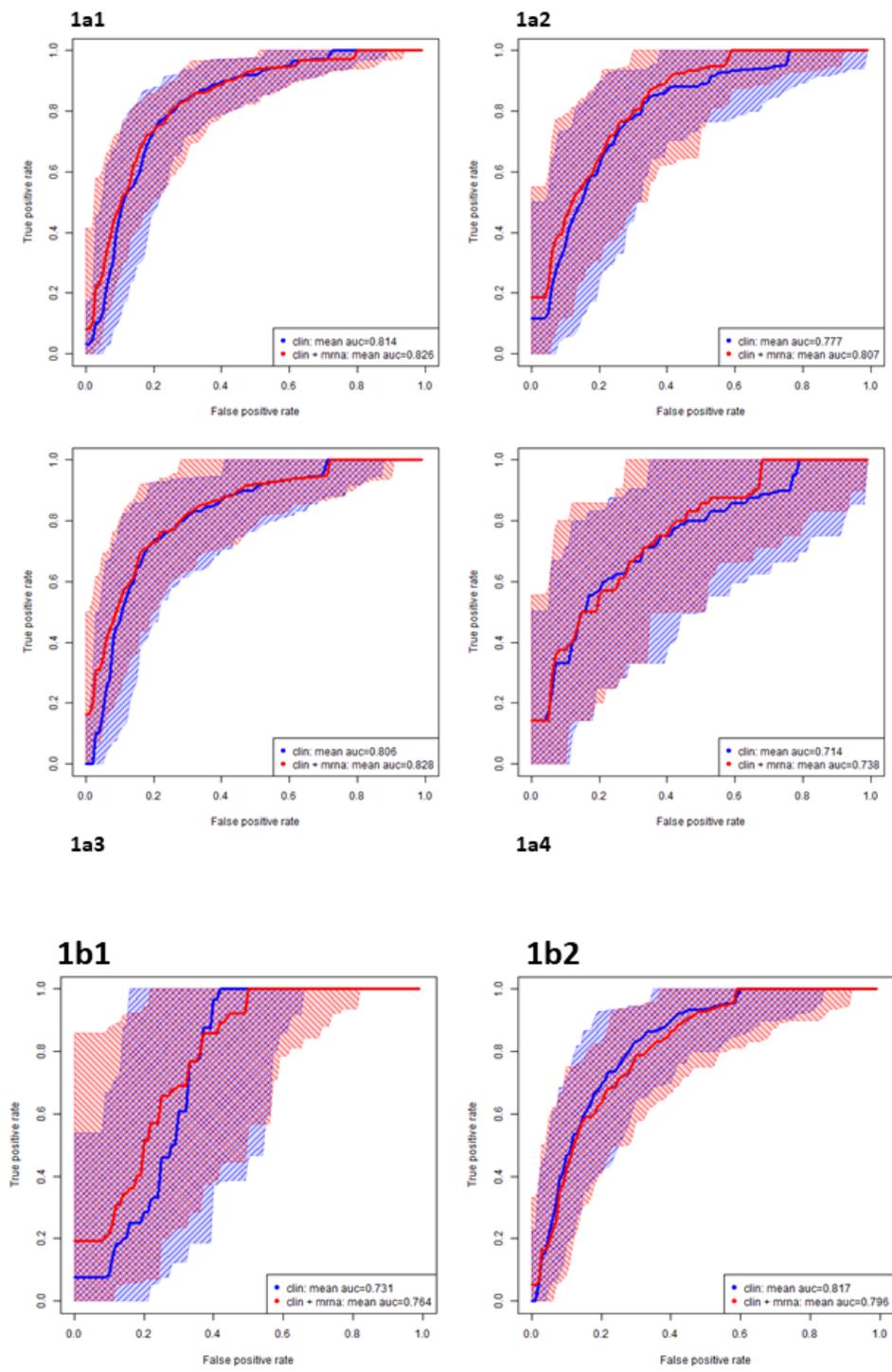


Figure 1

a. Receiver Operating Curves (ROC). Compare clinical factors alone model versus clinical factors + exhaled microRNAs combined model (red), to distinguish early stage (I+II) lung cancer all tumor histologies case donors combined, versus non-cancer controls. (1a1) All smoking categories combined; (1a2), former smokers only; (1a3) early stage only. Random forests, recursive partitioning and cross-validation were employed as described in the statistical analysis section. This ROC plot of true positives

versus false positives shows borderline incremental information (Fig. 1a1, +1.2% ($p=0.07$)) value of the exhaled microRNAs over and above the clinical model alone, in all subjects combining all smoking status, stages, and histologies. It particularly held true in main subgroup analyses, separating out former smokers (Fig. 1a2, +3.0% ($p=6.0e-03$)), and in early stage (I,II) models (Fig. 1a3, +2.2% ($5.1e-03$)). In combining these subgroups (1a4), early stage x former smokers combined did not show significant case-control discrimination (Fig. 1a4, NS). Model components and significance testing of area under curve (AUC) differences are described in Table 3. b. Current smokers only (left), cases versus non-cancer controls; Adenocarcinoma subjects only (right). Random forests, recursive partitioning and cross-validation were employed as described in the statistical analysis section. For Current smokers (left, 1b1), value of the exhaled microRNAs over and above the clinical model alone, was apparent +3.3% ($p=3.5e-02$). For adenocarcinoma case subsets (right, 1b2), the miR model detracted from the clinical+miR combined case-control discrimination, (-2.1% ($p=1.1e-02$)). Model components and significance testing of area under curve (AUC) differences are described in Table 3.

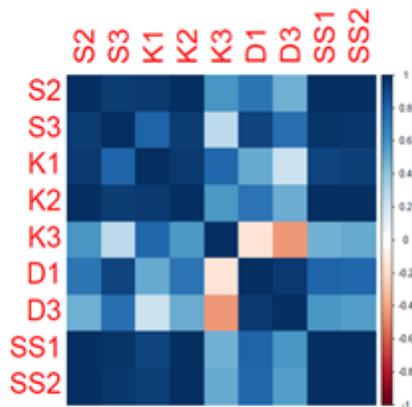


Figure 2

Temporal stability of EBC miRNA for an individual across time. S, K, D and SS represented different volunteers and S2 and S3 represented different timepoints of the same subject (S). Correlation coefficients represented on right.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Supplementarymaterial.docx](#)