The LuCID study: Detection of lung cancer breath biomarkers via Breath Biopsy in a multi-centre trial

BREATH® BIOPSY

CANCER RESEARCH UK

Marc van der Schee[1], Max Allsworth[1], Yichen Chen[1], Rob Smith[1], Simon Kitchen[1], Mariana Ferreira Leal[1], Chris Hodkinson[1], Billy Boyle[1], Alexandra de Saedeleer[1], Duncan Apthorp[1], Jasper Boschman[1], Silvano Dragonieri[2], Hubert Wirtz[3], Mina Gaga[4], Anjani Prasad[5,6], Mohammed Haris[7], Liz Fuller[8], Jakki Faccenda[9], Lori Calvert[9], David Baldwin[10], Arnaud Scherpereel[11], Jonathan Bennett[12], Serena Chee[13], Andrew Barlow[14], Andrew Wight[15], Veronica Conteh[16], Martin Ledson[17], Eleanor Mishra[18], Amrithraj Bhatta[19], Mark Weatherhead[20,21], William Ricketts[22], John Wrightson[23], Seamus Grundy[24], Philip Crosbie[25], Mamta Ruparel[26], Samuel Janes[26], Robert Rintoul[27,28]

[1]Owlstone Medical Ltd., Cambridge, UK. [2]Università degli Studi di Bari Aldo Morom, Bari, Italy. [3]Universitätsklinikum Leipzig, Leipzig, Germany. [4]Athens Chest Hospital, Athens, Greece. [5]Wycombe Hospital, Buckinghamshire Healthcare NHS Trust, Wycombe, UK. [6]Stoke Mandeville Hospital, Buckinghamshire Healthcare NHS Trust, Aylesbury, UK. [7]Royal Stoke University Hospital, Stoke-on-Trent, UK. [8]South Tyneside District Hospital, South Shields, UK. [9]Peterborough City Hospital, Peterborough, UK. [10]Nottingham University Hospital NHS Trust, Nottingham, UK. [11]Service d'Explorations Fonctionnelles Respiratoires, Hôpital Calmette, Lille, France. [12]Glenfield Hospital, Leicester, UK. [13]University Hospital Southampton, Southampton, UK. [14]West Hertfordshire Teaching Hospitals NHS Trust, Watford, UK. [15]Barnet Hospital, Royal Free London NHS Foundation Trust, Barnet, UK. [16]Liverpool Heart and Chest Hospital NHS Foundation Trust, Liverpool, UK. [18]Norfolk & Norwich University Hospitals NHS Foundation Trust, Norwich, UK. [19]Blackpool Teaching Hospitals NHs Foundation Trust, Blackpool, UK. [20]Wansbeck General Hospital, Ashington, UK. [21]North Tyneside General Hospital, North Shields, UK. [22]St Bartholomew's Hospital, London, UK. [23]The Churchill Hospital, Oxford, UK. [24]University Hospital Aintree NHS Foundation Trust, Liverpool, UK. [25]Manchester University NHS Foundation Trust, Manchester, UK. [26]University College London, London, UK. [27]University of Cambridge, Cambridge, UK. [28]Royal Papworth Hospital, Cambridge, UK.

At A Glance:

- The LuCID (Lung Cancer Indicator Detection) study aims to provide a robust evaluation of the diagnostic potential of breath biomarkers in the intended use population by using state-of-the-art breath collection and analysis approaches.
- The breath of 1844 subjects presenting with a clinical suspicion of lung cancer was analyzed using two complementary analytical pipelines to capture the broadest range of exhaled biomarkers possible. Results for 814 controls and 574 cases are presented here.
- 11 breath biomarkers differed significantly between cases and controls. In a multivariate model the combined performance of breath biomarkers did not exceed or add to, the accuracy of an established epidemiological risk prediction model.

1. Background and Objectives

- The premise of breath analysis for the early detection of lung cancer is its ability to reflect metabolic changes inside the body by analyzing Volatile Organic Compounds (VOCs) that occur in trace amounts in the breath. As metabolic changes occur early in the development of cancer and breath can be collected in virtually limitless quantities, VOCs may have value as a (complementary) screening modality.
- There is a notable discrepancy between the promising research thus far and its translation into clinical practice. Key limitations relate to study design, the analytical methodology and statistical approaches taken.
- **The LuCID study aims to address these challenges by evaluating the potential of exhaled breath biomarkers in a large cohort of individuals with a clinical suspicion of lung cancer by:**
  - **Identifying breath biomarkers that differ significantly between subjects with and without primary lung cancer.**
  - **Evaluating the diagnostic performance relative to and in addition to established epidemiological risk prediction models.**

2. Methods

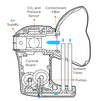
Trial design

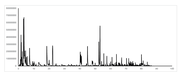
Prospective case-control study enrolling individuals suspected of having lung cancer from 26 sites across Europe and the United Kingdom.

Inclusion criteria: Clinical suspicion of lung cancer warranting further diagnostic work-up based on symptoms and/or suspicious imaging.

Exclusion criteria: Subjects undergoing treatment for any type of malignancy.

Cases were individuals with histopathologically confirmed diagnosis of lung cancer. Controls were participants without any type of cancer during 24 month follow-up. Individuals with an inconclusive diagnosis such as an indeterminate pulmonary nodule or inconclusive biopsy were excluded from the primary analysis.





Exhaled Breath Sampling

- Breath sampling prior to completion of diagnostic work-up.
- ReCIVA® Breath Sampler (Owlstone Medical) collecting 4 sorbent tubes per individual.
- Real-time quality control of breath sample collection.
- Liverpool Lung Project Risk model data obtained at baseline.

Breath Analysis

- Breath samples were analyzed by Gas-Chromatography Mass-Spectrometry.
- Samples analyzed by Method 1 (M1): non-polar, general-purpose method, OR Method 2 (M2): polar, improved signal-to-noise levels and increased sensitivity method.
- Targeted analysis was performed on 65 literature-reported compounds.
- Untargeted analysis was performed to identify novel candidate compounds.
- Stability of analytical performance was evaluated by running frequent quality controls and chemical standards.

Statistical Analysis

- All data was normalized and adjusted for potential biases; age, gender, BMI, smoking status.
- Univariate analysis was performed to identify potential biomarkers. Both uncorrected and corrected p-values were computed.
- Multivariate analysis was performed to generate Receiver Operator Characteristic Curves with associated AUC, Sensitivity and Specificity.
- Performance of breath biomarkers was evaluated relative to and in addition to the Liverpool Lung Project Risk Model.

3. Results

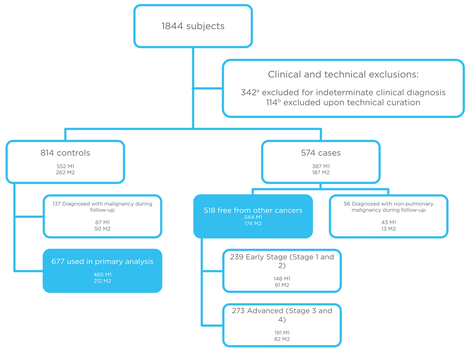


Figure 1: LuCID population. [a] 6 ineligible subjects, 336 in whom no unequivocal per protocol tissue-based diagnosis or exclusion of lung cancer could be obtained, typically unsuccessful biopsies or indeterminate pulmonary nodules. [b] 67 technically failed breath collections and 47 analytical failures.

Table 1: Clinical and pathological characteristics of LuCID subjects.

	M1				M2			
	Lung Cancer			Controls (N=465)	Lung Cancer			Controls (N=212)
	Advanced (N=191)	Early (N=148)	All Cases* (N=344)		Advanced (N=82)	Early (N=91)	All Cases* (N=174)	
Age [years; median (Q1-Q3)]	71 (65-76)*	72 (67-79)*	72 (66-77.2)*	67 (58-74)	71 (67-76.8)	74 (69-77.5)*	72 (68-77)*	69.5 (59-77)
BMI [median (Q1-Q3)]	25.5 (22.9-29.4)*	26.3 (23.8-29.8)	25.8 (23.2-29.4)	26.6 (23-30.9)	25.5 (23.4-28.2)*	26.8 (23.7-31.3)	25.9 (23.4-29)	27 (24.1-30.8)
Gender [Female; N(%)]	72 (37.7)	58 (39.2)	132 (38.4)	194 (41.7)	26 (31.7)	47 (51.6)	73 (42)	87 (41)
Smoker Status [N(%)]								
Current Smoker	63 (33)	41 (27.7)	106 (30.8)	102 (21.9)	32 (39)	20 (22)	52 (29.9)	44 (20.8)
Ex-smoker	111 (58.1)	91 (61.5)	204 (59.3)	221 (47.5)	43 (52.4)	57 (62.6)	101 (58)	103 (48.6)
Never Smoked	17 (8.9)	16 (10.8)	34 (9.9)	142 (30.5)	7 (8.54)	14 (15.4)	21 (12.1)	65 (30.7)
Years Smoked [median (Q1-Q3)]	45 (33-53)*	43 (25-53)*	44 (30-53)*	26 (0-44)	46 (31-54)*	44 (26-52)*	45 (28-53)*	27 (0-44)
History of COPD [N(%)]	86 (45)	84 (56.8)*	171 (49.7)	199 (42.8)	25 (30.5)	26 (28.6)	51 (29.3)	49 (23.1)
History of Pneumonia [N(%)]	2 (1.05)	5 (3.38)	7 (2.03)	18 (3.87)	1 (1.22)*	0 (0)*	1 (0.575)*	15 (7.08)
LLPv2 Score [median (Q1-Q3)]	0.0717 (0.0288-0.134)	0.0752 (0.0275-0.169)*	0.0716 (0.0283-0.153)*	0.0198 (0.00503-0.0675)	0.0675 (0.0305-0.133)*	0.0606 (0.0235-0.111)*	0.0636 (0.0266-0.119)*	0.0194 (0.00490-0.0722)
Histological subtype [N(%)]								
NSCLC	169 (88.5)	128 (86.5)	299 (86.9)	-	73 (89.0)	88 (96.7)	162 (93.1)	-
Adenocarcinoma	100 (52.4)	69 (46.6)	170 (49.4)	-	35 (42.7)	52 (57.1)	88 (50.6)	-
Squamous cell carcinoma	50 (26.2)	48 (32.4)	98 (28.5)	-	26 (31.7)	34 (37.4)	60 (34.5)	-
Large cell neuroendocrine carcinoma	3 (1.6)	3 (2.0)	7 (2.0)	-	0 (0)	1 (1.1)	1 (0.6)	-
Adenosquamous carcinoma	3 (1.6)	0 (0)	3 (0.9)	-	2 (2.44)	1 (1.1)	3 (1.7)	-
Sarcomatoid carcinoma	2 (1.0)	0 (0)	2 (0.6)	-	1 (1.22)	0 (0)	1 (0.6)	-
Other (NSCLC not otherwise specified)	11 (5.8)	8 (5.4)	19 (5.5)	-	9 (11)	0 (0)	9 (5.2)	-
SCLC	22 (11.5)	5 (3.4)	29 (8.4)	-	8 (9.76)	2 (2.2)	10 (5.75)	-
Typical Carcinoid	0 (0)	12 (8.1)	13 (3.8)	-	1 (1.22)	1 (1.1)	2 (1.15)	-
Atypical Carcinoid	0 (0)	3 (2.0)	3 (0.9)	-	0 (0)	0 (0)	0 (0)	-

*p-value < 0.05 when comparing cases to controls using the Mann-Whitney U test (continuous variable), Fisher's Exact test (binary variable), or Chi-squared test (multi-category variable). *Including patients with undefined tumour stage. Early cancer: any tumour up 2b TNM stage. LLPv2: Liverpool Lung Risk Project version 2.

Table 2: Breath biomarkers appearing at significantly different concentrations in the breath of lung cancer cases and controls after correcting for gender, age, BMI and smoker status.

Method	Volatile Organic Compound	Early			Advanced			All stages		
		Coefficient	p-value	Corrected p-value	Coefficient	p-value	Corrected p-value	Coefficient	p-value	Corrected p-value
M1	Allyl Methyl Sulfide	-0.124	0.265	0.963	-0.263	0.011	0.339	-0.218	0.011	0.427
	Acetoin	-0.075	0.512	0.963	-0.250	0.016	0.339	-0.209	0.015	0.427
	D-Limonene	0.126	0.262	0.963	-0.314	0.003	0.239	-0.131	0.143	0.950
	2 Undecanone	0.114	0.337	0.963	0.273	0.015	0.339	0.228	0.014	0.427
M2	Cyclohexane	-0.128	0.494	0.831	-0.357	0.034	0.384	-0.210	0.122	0.546
	Allyl Methyl Sulfide*	-0.048	0.762	0.999	-0.444	0.010	0.245	-0.319	0.032	0.410
	Acetoin	-0.203	0.332	0.815	-0.454	0.005	0.245	-0.327	0.012	0.301
	Acetophenone	-0.344	0.033	0.425	-0.215	0.219	0.662	-0.295	0.047	0.531
	Trans-2-Hexenyl isovalerate*	-0.346	0.033	0.425	-0.049	0.762	0.955	-0.201	0.122	0.546
	7-hexyl-Eicosane*	-0.462	0.004	0.203	-0.245	0.134	0.651	-0.329	0.011	0.301
	Hexadecane	-0.515	0.013	0.344	-0.140	0.421	0.839	-0.235	0.116	0.546
	1-ethylnonyl-Benzene*	-0.368	0.023	0.408	-0.258	0.126	0.645	-0.305	0.018	0.301
	1-pentylheptyl-Benzene*	-0.517	0.001	0.119	-0.321	0.057	0.470	-0.389	0.002	0.245

* Tentative ID, chemical elucidation ongoing.

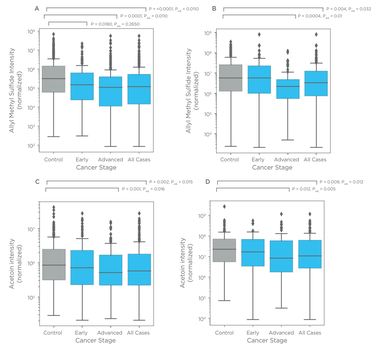


Figure 2: Normalized levels of volatile organic compounds on breath. A) Allyl Methyl Sulfide on breath with M1 cohort; B) Allyl Methyl Sulfide (tentative ID) on breath with M2 cohort; C) Acetoin on breath with M1; D) Acetoin on breath with M2. Uncorrected p-values before and after covariate adjustment are also displayed.

Table 3: Diagnostic performance characteristics for the breath biomarkers, epidemiological risk model and combined model.

Method	Stage	Model	ROC-AUC (95% CI)	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)
M1	Early	VOC	0.523(0.114)	42.4	71.3	29.5	79.4
		LLP	0.702(0.109)	78.8	54.6	33.8	88.1
		VOC & LLP*	0.702(0.109)	78.8	54.6	33.8	88.1
	Advanced	VOC	0.606(0.0983)	81.3	40.7	37.3	81.5
		LLP	0.730(0.0906)	95.8	38.9	40.5	93.3
		VOC & LLP	0.682(0.0947)	75.0	59.3	44.3	83.1
	All Cases	VOC	0.598(0.0813)	50.0	71.3	56.9	64.2
		LLP	0.720(0.0738)	76.2	57.4	57.8	74.7
		VOC & LLP*	0.720(0.0738)	76.2	57.4	57.8	74.7
M2	Early	VOC	0.543(0.137)	92.3	26.4	37.1	82.4
		LLP	0.688(0.130)	100	30.2	40.3	94.1
		VOC & LLP	0.632(0.134)	42.3	79.2	47.6	72.4
	Advanced	VOC	0.561(0.157)	33.3	88.7	45.5	78.3
		LLP	0.597(0.157)	88.9	37.7	31.3	87.0
		VOC & LLP	0.627(0.155)	61.1	66.0	35.7	81.4
	All Cases	VOC	0.554(0.116)	84.1	35.8	51.4	70.4
		LLP	0.651(0.111)	90.9	37.7	54.2	80.0
		VOC & LLP	0.639(0.111)	90.9	34	52.7	78.3

Diagnostic performance metrics for multi-variate breath analysis. Values are reported at the Youden Index of the ROC-Curve. Values represent a breath biomarker only (VOC-only), Liverpool Lung Risk Project (LLPv2) based model and a combined VOC & LLPv2 model. Liverpool Lung Risk Project version 2. * For these models no VOCs were selected, model performance is thus identical to the LLPv2.

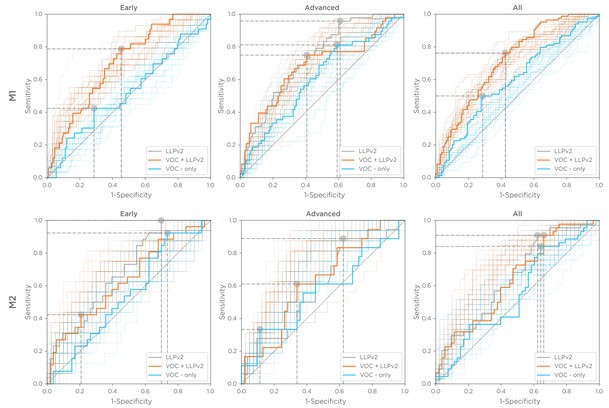


Figure 3: ROC-Curves for multi-variate classification models. ROC-Curves for the 3 reported models: 1. VOC-only; 2. Liverpool Lung Risk Model; 3. Combined. Empirical variation in ROC curve was simulated by repeating the model-building and the ROC construction process via a 20-split Monte-Carlo cross-validation applied to the training set.

4. Conclusion and Future Direction

- Univariate analysis identified 11 breath biomarkers that were significantly altered when comparing lung cancer cases and controls. In general lower concentrations of these biomarkers were observed with advanced stages of lung cancer. Acetoin and Allyl Methyl Sulfide were observed in separate cohorts with distinct analytical methodologies. Notably none of these differences were significant after correcting for multiple testing.
- Multivariate analysis demonstrated that the diagnostic performance of these markers did not exceed the Liverpool Lung Project risk model.
- The results of the LuCID study suggest that endogenously produced breath biomarkers have limited diagnostic potential in the relevant clinical population for the early detection of lung cancer. Targeted approaches amplifying the volatile biomarker signal of metabolic pathways altered in lung cancer are likely needed to develop a breath based screening test for lung cancer. Our current research focuses on exploiting one such pathway using an exogenous VOC probe for beta-glucuronidase in the tumor micro-environment.

Acknowledgments: Hannah Winter, Ben Taylor, all study subjects and all research staff of the participating hospitals.