COMPARISON OF TARGET AND NON-TARGET APPROACHES IN BREATH ANALYSIS FOR DISCRIMINATION OF LUNG CANCER FROM OTHER PULMONARY DISEASES AND HEALTHY PERSONS

Michalis Koureas [1], Dimitrios Kalompatsios [1] and Andreas Tsakalof [1,2]

1 Department of Hygiene and Epidemiology, University Hospital of Larissa, Faculty of Medicine, University of Thessaly,

2 Department of Biochemistry, Faculty of Medicine, University of Thessaly, Larissa, Greece

The main goal of this study was to compare the efficiency of target and non-target GC-MS breath VOCs analysis in the diagnosis and discrimination of lung cancer (Ca+) patients from patients with other respiratory diseases (Ca-) and healthy controls (HC).

1. Materials and Methods

Population sample

Patients with pathological findings on chest CT (n=89) → Bronchoscopy-, cytological / histological analysis

- **Group 1**. Diagnosed with lung cancer. Ca+ (n=51)
- **Group 2**. Not diagnosed with lung cancer. Ca- (n=38)

Group 3. Healthy Controls, HC, (n=53)

Sampling and samples analysis

Breath samples were collected with Tedlar® bags. Extraction and pre-concentration of the analytes from breath samples was achieved by solid phase microextraction (SPME) using a 75μm Carboxen-polydimethylsiloxane (CAR/PDMS)-coated melt silicon fiber assembly. Instrumental analysis was performed with a Finnigan Trace GC Ultra/Polaris Q Quadrupole Ion Trap GC/ MSn system.

2. Targeted approach

The concentrations of 19 volatile organic compounds (VOCs) most often referred in the literature as lung cancer biomarkers were determined in the exhaled breath of study participants. Strong associations were identified only for 5 compounds indicated in the table (Koureas et al. *Metabolites* 10 (8), 2020).

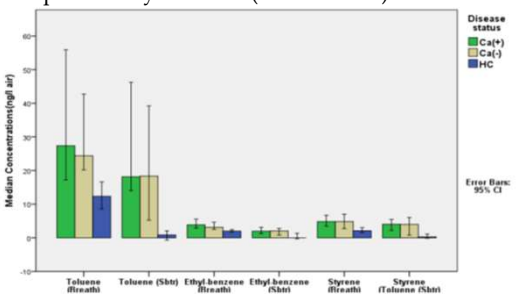
Concentrations of VOCs biomarkers (ng/L) in breath before and after ambient air correction (Sbtr) and significance of their differences between investigated groups.

		Patients Ca⁺		Patients Ca⁻		Healthy Controls		Ca⁺ vs. HC	Ca⁺ vsCa⁻
Substance		**%†**	**Median(IQR)**	**%†**	**Median(IQR)**	**%†**	**Median(IQR)**	***p****	***p****
2-propanol	Br	21.6%	528(324–804)	15.8%	490(382–702)	43.4%	315(218–497)	**0.002**	0.636
	Sbtr		neg.(neg.-neg.)		neg.(neg.-neg.)		neg.(neg.−130.97)	**0.041**	0.491
1-propanol	Br	11.8%	30.78(7.14–57.81)	21.0%	24.13(8.14–60.85)	35.8%	63.84(38.46–103.63)	**<0.001**	0.684
	Sbtr		neg.(neg.-neg.)		neg.(neg.-neg.)		neg.(neg.−14.09)	**0.005**	0.255
Ethyl_benzene	Br	76.4%	3.85(2.44–6.26)	76.3%	3.13(2.02–5.36)	49.0%	2.01(1.30–2.89)	**<0.001**	0.476
	Sbtr		1.99(0.18–4.15)		2.05(0.06–3.56)		neg.(neg.−1.78)	**<0.001**	0.442
Styrene	Br	86.2%	4.83(2.36–7.87)	89.5%	4.85(1.71–8.26)	62.3%	2.11(1.25–3.53)	**<0.001**	0.914
	Sbtr		4.02(0.53–6.93)		3.97(0.62–7.48)		0.28(neg.−1.57)	**<0.001**	0.888
Toluene	Br	86.2%	27.36(15.35–66.04)	76.3%	24.39(18.14–51.17)	58.5%	12.33(6.27–21.37)	**<0.001**	0.816
	Sbtr		18.16(2.32–56.58)		18.34(0.95–48.08)		0.87(neg.−4.21)	**<0.001**	0.592

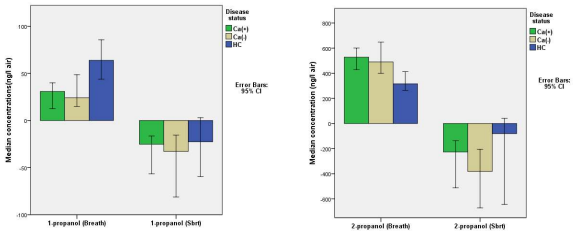
1. Observed variations in the concentrations of **monoaromatic compounds** may reflect the general impairment of pulmonary function (or more broadly the physiological and biochemical status of the patient) and are obviously not specific biomarkers of lung cancer.

2. 1-propanol & 2-propanol in human breath are endogenous and exogenous origin and are rapidly absorbed and metabolised to corresponding aldehydes - acetone and propanal. Thus negative corrected concentrations (Sbtr) are observed. The reaction is catalysed by alcohol dehydrogenase (ADH).

The breath concentrations of monoaromatic compounds were found to be significantly higher in the patients with impaired pulmonary function (Ca+ and Ca-) vs HC.



Difference in the breath concentrations of **1-propanol and 2-propanol** before and after ambient air correction (Sbtr).



3. Untargeted approach

Step 1. Data processing with XCMS online
Peak identification, alignment, retention time correction
Preliminary statistical analysis

Step 2. Detection of discriminative features
Criteria: p-value ≤ 0.01, fold-change ≥ 2, mzmed ≤ 150, rtmed ≤ 14, maximum intensity ≥ 10000
367 features detected corresponding to 110 peaks

Step 3. Evaluation of chromatographic peaks
Evaluation of shape of chromatographic peaks (Single Ion monitoring) → **28 peaks were excluded** due to not acceptable chromatographic characteristics

Step 4. Compound identification
Identification of substances with NIST library → **53 peaks were excluded** corresponding to siloxanes and Tedlar bag materials. **29 peaks selected**

Step 5. Reprocess of raw data with Thermo Xcalibur
Development of a dataset with the chromatographic area for to each peak

Step 6. Statistical analysis (29 peaks)

Identified compounds by NIST and their origin was suggested by the investigation of biochemical pathways (KEGG pathway database).
Exogenous: 2,5-dimethylfuran, 3-methylfuran, xylene isomers, 1-methoxy-2-propanol, **e**thylbenzene, toluene, eucalyptol
Endogenous/exogenous: acetaldoxime, p-benzoquinone, N-methyl acetamide, allyl methyl sulfide, methyl propyl sulfide, acetic acid, propionic acid, styrene, methyl lactate
Unknown compounds: 12

For 5 compounds retention time was verified with Pro EZGC® Chromatogram Modeler (Restek Corporation, U.S)
4 compounds were verified by analytical standards

4. Comparison of discriminant ability

Targeted analysis

Ca(+) vs HC	Ca(+) vs Ca(-)
Sensitivity 70,6%	Sensitivity 39,2%
Specificity 88,7%	Specificity 36,8%
Accuracy 79,8%	Accuracy 38,2%

Untargeted analysis

Ca(+) vs HC	Ca(+) vs Ca(-)
Sensitivity 82,3%	Sensitivity 69,8%
Specificity 91,3%	Specificity 63,3%
Accuracy 87,0%	Accuracy 66,9%

5. Conclusions

- Both approaches (untargeted and targeted) achieved adequate discrimination of Ca+ patients from HC with untargeted approach offering better discrimination power.
- Limited discrimination between Ca+ from Ca- patients was observed in target analysis while this was substantially improved in non-targeted approach.
- Monoaromatic compounds were identified as significant biomarkers by both approaches while 1- and 2-propanol were not identified by untargeted approach.

This publication is made in the context of a scholarship programme for postdoctoral research implemented by the University of Thessaly and funded by the Stavros Niarchos Foundation.