Use of experimental sensors for discrimination of artificial breath mixtures in a lung cancer screening context

Justin DM Martin[a], Marc Debliquy[b], Yifan Luo[b], Driss Lahem[c], Ahmadou Ly[c], Anne-Claude Romain[a].

[a] University of Liège - Campus Environnement, Arlon, 6700, Belgium
[b] University of Mons – Service Science des Matériaux – Faculté Polytechnique, Mons, 7000, Belgium
[c] Materia Nova ASBL – Material Science Department, Mons, 7000, Belgium
e-mail: jdm.martin@uliege.be

Lung cancer is still in the deadliest cancers in Europe for 2020. The European Cancer Information System (ECIS) [1] is freely available on the web for a clear view of the statistics per year and country: In 2020, lung cancer is the second type of cancer in men and third type for women in Belgium (13% of all cancers diagnosed). However, it is the first cause of death by cancer for both women and men for that year. This trend is seen all over Europe. The main cause being the late diagnosis of lung cancers, which are asymptomatic until the late stages of the illness. The need for an early screening method has driven research in several directions, one of which relies on the use of gas sensors to mimic dog's olfaction, as it was shown that dogs can detect lung cancer in the breath of patients [2].

The Pathacov project [3] aims at creating an electronic nose to detect lung cancer in the population at risk. Metal oxide sensors are being developed within the project to better detect cancer biomarkers, which are identified within a large-scale clinical study in university hospitals in the north-east of France. Before long and costly clinical trials, the performance of the prototype electronic nose has to be evaluated. It is important to test the system in a way that is relevant to the future usage conditions: breath is a warm, humidity saturated gas mixture with hundreds of volatile organic compounds (VOCs) that includes numerous confounders.

In order to do this, a novel approach has been implemented. Real breath samples from healthy and available volunteers are collected in Teflon FEP sampling bags (Figure 1). Half of each sample is transferred to another bag and spiked with a mixture of literature issued cancer VOCs in breath-like concentrations. This constitutes the lung cancer group. The other half of the sample is used as is and constitutes the healthy control group. Both are analysed by gas chromatography mass spectrometry (GCMS) and the prototype electronic nose in parallel. Each measurement by the electronic nose is repeated four to six times.

The results of the electronic nose are processed through a principal component analysis (PCA) in order to evaluate the contribution of each sensor to the separation of the group's clusters (healthy and cancer). This enables the selection of the best performing sensors to be included in the final prototype. The Euclidean distance between groups should be maximized and the Euclidean distance within a group (variance) should be minimized, similarly to a k-means clustering approach. This is used as an array performance metric: optimal sensors produce good clustering on PCAs, good clustering enables better classification performances – regardless of the classification method.

Results from this experiment show (Figure 2) that it is possible to differentiate breath with and without cancer VOC addition, using commercial metal oxide sensors (Table 1). However, some overlap is apparent between the two groups, and a good share of the commercial sensors do not seem to bring useful information for good clustering. GGS 8530T (Umwelt Sensor Technik™) and MP901 (Winsen™) seem to contribute most to the separation. One should, however, be reminded that this experiment is only an emulation of real samples, and that apparently not contributing sensors could be useful for real breath samples. The relevance of this approach will be confirmed by comparison with real patient breath analysis using a comparable procedure. Similar tests will be conducted using experimental sensors

[1] 'European Cancer Information System'. https://ecis.jrc.ec.europa.eu (accessed Jun. 03, 2022).
[2] C. Feil *et al.*, 'Sniffer dogs can identify lung cancer patients from breath and urine samples', *BMC Cancer*, vol. 21, no. 1, p. 917, Aug. 2021, doi: 10.1186/s12885-021-08651-5.
[3] 'Diagnostic de pathologies humaines par analyse de COV dans l'air expiré', *PATHACOV*, 2020. https://pathacov-project.com/ (accessed Jul. 06, 2022).

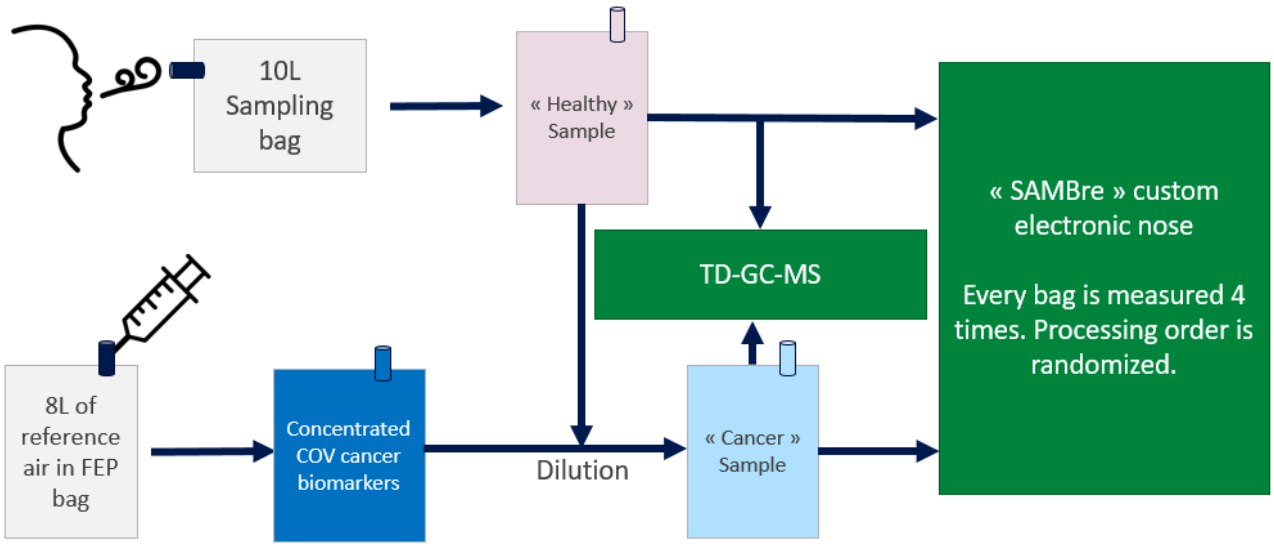

Figure 1. Sample processing diagram.

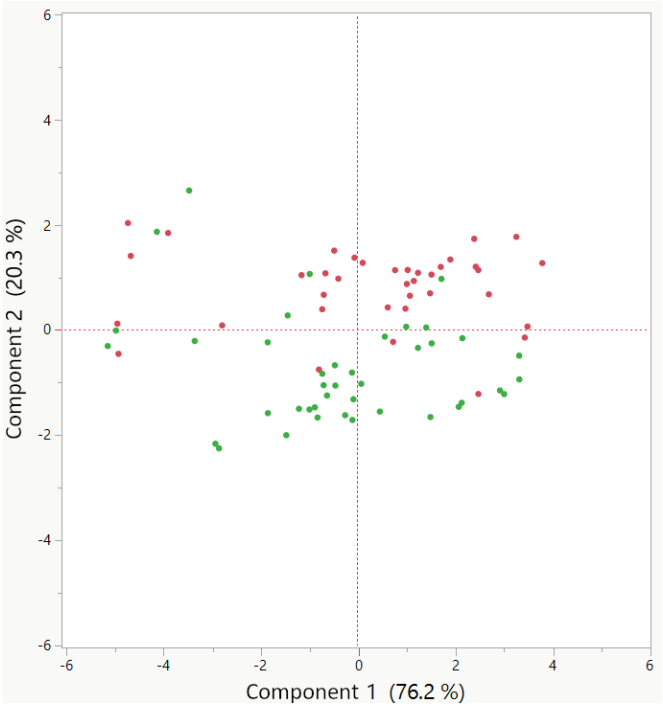

Figure 2. PCA analysis biplot for commercial sensors. Data comes from maximum conductance of each sensor, with baseline subtraction. Red dots are the "cancer" group samples, green are the "healthy" group samples. 79 measurements from 7 unique breaths were collected over 2 months (4 to 6 replicates per group and per unique breath). Two sensors (G8530T and MP901) seem to contribute more than others, as the loadings plot suggests.

Table 1. List of sensors used for the experiment

Builder	Sensors
Umwelt Sensor Technik™	1430T, 3530T, 8530T, 2530T
Winsen™	MP901
Figaro Engineering™	TGS2603
Bosh™	BME680