chemosensors

MDPI

Review

Building a Sensor Benchmark for E-Nose Based Lung Cancer Detection: Methodological Considerations

Justin D. M. Martin * and Anne-Claude Romain

Sensing of Atmospheres and Monitoring, Departement of Environmental Sciences, Arlon Campus Environnement, University of Liège, 6700 Arlon, Belgium
* Correspondence: jdm.martin@uliege.be

Abstract: Lung cancer is one of the deadliest form of cancer in Europe, characterized by a lack of obvious symptoms until the terminal stages of the illness. Electronic noses are a rising screening technology to detect early-stage lung cancer directly in the homes of people at risk. Electronic noses need to be tested using samples from patients. However, obtaining numerous samples from cancer patient turns out to be a difficult task in practice. Therefore, the development of a sensor benchmark able to evaluate the performance of sensors without direct breath sampling is of high interest. This paper focuses on the methodology for developing such a benchmark, in the case of a breath sampling electronic nose. The setup used is introduced and general recommendations based on literature and undergoing experiments is detailed. The benchmark can be used for a variety of sensors and a variety of target illnesses. It is also possible to apply it to other types of medical gaseous samples or environmental VOC monitoring. The benchmark is currently still undergoing tests, and results will be published in a following article.

Keywords: electronic nose; cancer; breath; sensor array; lung; benchmark

Citation: Martin, J.D.M.; Romain, A.-C. Building a Sensor Benchmark for E-Nose Based Lung Cancer Detection: Methodological Considerations. *Chemosensors* **2022**, *10*, 444. https://doi.org/10.3390/chemosensors10110444

Academic Editor: Simonetta Capone

Received: 9 September 2022
Accepted: 24 October 2022
Published: 26 October 2022

1. Introduction

According to records from 2015 to 2017 in the United States by the National Cancer Institute, approximately 6% of men and women will be diagnosed with Lung Cancer (LC) at some point during their lifetime [1]. LC was the leading cause of death for men and the second leading cause of death for women in 2018 in Europe (estimated fatalities: 388,000 Europeans). LC is the most common form of cancer [2], and if the number of cases is slowly dropping in developed countries, the burden of LC is on the rise in emerging countries, resulting in an increase in the global number of cases [3].

Early detection serves well against cancer. For example, thanks to screening and cancer awareness, declines in breast cancer mortality rates in Europe have been reported [2]. This is crucial for a lot of cancer types, and especially LC, as early detection increases survival chances in the next five years by a factor of five (early stage against late stage) [4]. The usual diagnostic methods (low-dose computerized tomography, PET-scan, IRM) are hard to use for early lung cancer diagnostics in a wide asymptomatic population, as their cumbersome nature confines them to hospital usage only. To reduce pressure on medical care centers, a simple, portable, inexpensive, non-invasive, new early screening method is required.

Among last years' proposals, one promising method resides in the analysis of volatile organic compounds (VOC) naturally emitted by the human body. The abundance and composition of the VOCs in the breath is tightly linked with the human body's metabolic activity. These VOCs (or metabolomes) therefore act as biomarkers and have been extensively studied in the literature, which is especially plentiful for LC studies [5].

Of all the ways to analyze VOCs, sampling breath is a straightforward and easy one. Because of the logical proximity of lung cancer and the airways, a lot of projects used breath sampling for lung cancer detection. Several methods are popular in the literature:

gas chromatography coupled with mass spectrometry (GCMS) [6,7], breath condensate analysis [8,9], and the use of an Instrumental Odor Monitoring System (IOMS, also called electronic nose or e-nose). This last method has several advantages that make it interesting as an answer to the remote screening problem: ease of use, portability and low costs would be the main ones. Therefore, this paper will mainly focus on the IOMS approach.

The literature regarding the various projects tackling the creation of custom sensor arrays for lung cancer detection is starting to become plentiful. The varying approaches and considerations call for reflection on which good practices should be followed, and how to build a sensor benchmark for sensor selection in breath sensing. The aim of the present document was to share the reflection on the construction of a thorough benchmark procedure based on current literature in breath sensing and IOMS.

The article is structured into two main sections. First, a literature review on the use of the electronic nose, breath, the creation of gas mixtures, the target gases, sample handling, the data treatment and the validation using a reference method is performed. Reproducibility aspects are considered for all experiments. Then, the methodological proposal will be explained in regard to the findings in literature and the original experiments are realized. A short discussion concludes the article.

2. Literature Review

This literature review was not made according to a specific protocol, except for the part on target biomarkers. The methodology of the systematic literature review on biomarkers is detailed on point 2.3.2 "On the selection of target VOCs".

The goal was to cover this very plural subject in a manner that helps people working on tech aspects and device testing. The effort was to use the most useful articles collected during three years of experimentation. Articles for this review were found on Scopus, PubMed and University of Liège databases. The following keywords were used: "e-nose", "electronic nose", "sensor", "array", "breath", "lung cancer", "disease", "benchmarking", "merit", "machine learning", "standard", "drift", "calibration", "diagnostic", and "sampling". The references of the selected documents were examined to find more pertinent articles. A total of 659 articles published from 1980 to 2022 in the English language were considered for inclusion in this review. Meta-analysis articles were excluded. Other review articles were included.

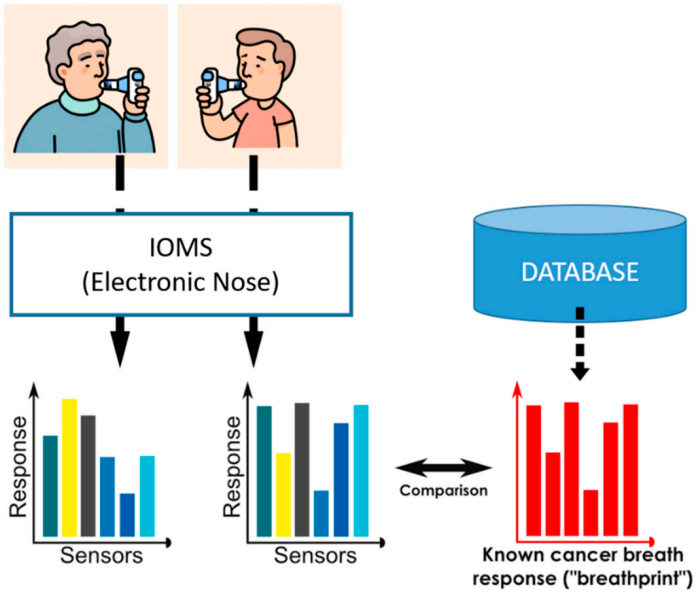
2.1. Instrumental Odour Monitoring System (IOMS)

An IOMS is a device mimicking the biological nose and able to tell gas mixtures apart from each other [10]. The basic principle of the e-nose was formed in 1982 in the work of Persaud et al. [11].

The particularity of an e-nose is that it relies on the "fingerprint" (or "breathprint" in the field of disease detection) of the gas mixture. Such a system will not identify chemical species in a gas sample (unlike a mass spectrometer), but can recognize mixtures of several hundreds of compounds at a time and tell each mixture apart in a very short timeframe [10].

As shown on Scheme 1, each measurement triggers several sensors. It is possible that a sensor reacts to a single compound, or to several at once. The strength of the signal obtained depends on each sensor's affinities. Each sensor being different, the vector containing the response of the whole array will be the mixture's "fingerprint".

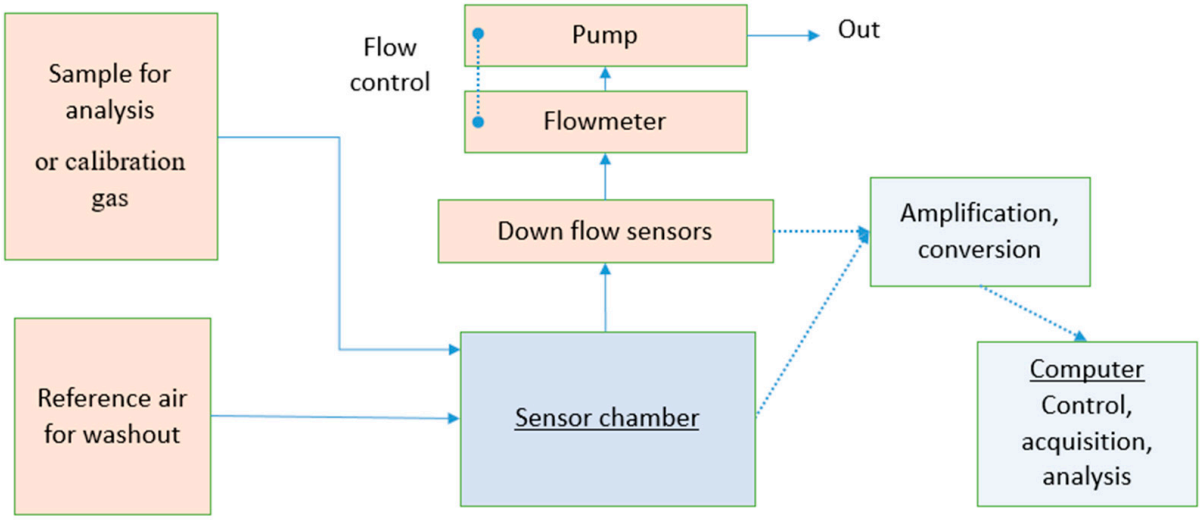
That vector can be compared with previously observed vectors (for example, those from the breath of cancer patients and healthy patients) using a clustering algorithm and dimensionality reduction [12], which allows for the classification of a new sample in a known category and therefore gives the user an indication on the need for further medical tests.



Scheme 1. Breathprint disease recognition principle using an IOMS.

IOMS are structured as shown in Scheme 2, and are as follows [12]:

- Sample storage, which includes piping and any form of sample containment before analysis. It can have a form of pre-concentration or pre-treatment against interfering compounds.
- The sensor array is usually composed of 4 to 32 sensors housed in one (or several) chamber(s). Sensors previously used in cancer breath detection are surface acoustic wave sensors (e.g., SAW, BAW, QMB/QCM) [13], polymer gas sensors [14,15] or carbon nanotube based sensors [16], but the most common are likely metal oxide semiconductor (MOS) sensors as the technology is well known and commercially available [17–19].
- A signal treatment system, which converts the analogic output of the sensors to a numeric output interpretable by the processing unit.
- A processing unit that will control the other units, collect and save the data from the sensors.



Scheme 2. Usual IOMS structure.

An IOMS seems to be able to identify cancer, and even isolate specific mutations of cancer [20]. It should however be noted that sample pools are usually small in these studies (usually, 20–100 patients, half of them being controls, with some exceptions having up to 300 patients) and further research is needed to confirm the IOMS as a diagnostic tool [21–26].

2.2. Establishing a Performance Metric for Gas Sensors

By taking computing power benchmark as an epitome [27], a good benchmark should be:

- Relevant (to the purpose of the device tested).
- Equitable (all sensors are tested and are compared on the same basis).
- Repeatable (results can be verified).
- As effective as possible in regard of cost and logistics.
- Should work for all kind of purposes and concentration ranges, in order to be usable. across devices with different purposes using the same concept.
- Transparent (metrics should be easy to understand).

By setting these goals, the aim was set to the conception of a general use breath sensor benchmark.

The first step to evaluate a sensor would be the evaluation of its metrological performances [28,29]: sensor response time, recovery time, operating conditions (e.g., working temperature, energy consumption), general behavior (i.e., linear or non-linear response), sensitivity to temperature and humidity variations, stability over time (drift), sensor lifetime and recommended stabilization period. While some of these aspects do not influence the quality of the information given by a sensor, they can, however, make the sensor integration and use difficult, depending on the setup.

The other important aspect is the sensor's sensitivity to a variety of target VOCs and confounding factors. This is the subject of the first sensor performance evaluation. By determining the sensor responses to increasing concentrations of chosen compounds, sensor behavior becomes apparent. It is often advised to favor diversity within an array of sensors, as it brings more information on samples and therefore provides good and robust discrimination [30]. In this case, diversity means that each sensor has a different sensitivity to each gaseous species of interest [31]. It is therefore desirable to compare sensors having similar sensitivities to the analyte between them, in order to evaluate which sensor provides the highest quality of information for the task at hand. Comparing sensors of radically different information would be counterproductive. The best sensor out of several similar sensors would be the most sensitive one, and with the lowest limit of detection (LOD), as breath VOCs are in parts-per-billion (ppb) to parts-per-million concentration (ppm) [5,32].

One way to favor diversity is to measure the amount of information given by a combination of the array's features (extracted from the raw sensor signal, such as the area under the curve, or the maximum peak height) using a statistical test (through Principal Component Analysis's (PCA) eigenvalues, interclass discrimination efficiency, correlation between features and variables, among others) is called a Filter method. The other way is by using a Wrapper method, which subsets the array's features to train a model and evaluate the resulting performances [33]. The algorithm starts by choosing a single combination of features and changes its composition by adding (Sequential Forward Selection, or SFS) or removing (Sequential Backwards Selection, or SBS) features progressively. Changes are kept if they improve the separation of the groups of interest. Both methods have a tendency to get stuck in local optimums, which should be kept in mind while using them [29].

Wrapper methods often achieve better predictive accuracy, but are computationally intensive [33]. They do not prevent irrelevant, redundant or correlated feature selection, however [29]. Filters are simpler and often find a general solution, where wrappers have an overfitting tendency [33].

A few very interesting methods have been explored in the literature. For example, it is possible to use PCA eigenvectors to select sensors, and therefore evaluate their quality

and complementarity. Sensors contributing the most to the highest ranking eigenvectors can be chosen, and it has been demonstrated that this technique improves classification accuracy [29]. It is also possible to produce uncorrelated features using PCA to achieve better classification performances [34]. It should be noted, however, that PCA captures directions of maximum variance, which do not always contain useful information for intergroup discrimination [33].

2.3. On the Use of Breath-like Gaseous Samples

Between healthy and sick people, the differences lie in the part-per-billion (ppb) concentrations for most VOCs, and in the absence/presence for some potential biomarkers [5,32]. Therefore, one of the main challenges of making breath-like mixes is the accurate and reproducible dilution of gases.

One approach consists in the insertion of a few microliters of the liquid volatile compounds in a gas sampling bag; the bag is equipped with a septum and pre-filled with analytical air, and with the help of a microliter syringe, it is possible to reach ppm-level concentrations [35,36]. Such "Static methods" work well for low concentrations and when small volumes are needed. In order to reach ppb level with liquid injection, however, a second dilution is necessary. The best method to do so is likely by using pumps and Mass Flow Controllers (MFC, a dynamic method, which means working with flow rates instead of volumes for dilution) [37]. Other methods have been developed [36], but only some of them are usually found in post-2000 literature.

An interesting alternative is the use of a permeation oven (dynamic method) as presented by Helwig et al. [38]. This device heats permeable tubes filled with the chemical species of interest, releasing small amounts of them into an air flow. This enables one to reach ppb-level concentrations with accuracy and enables in-line setups with machine-operated preprogrammed experiments. This allows for the easy calibration of sensors and reduces the sources of errors greatly. Coupled with MFCs, it was previously shown that this setup permitted a wide range of concentrations, with reasonable concentration errors (less than 1% typically, even if low permeation rates can give more than 10% error). With MFCs considered, Helwig et al. managed to obtain 12.1% error or less on output concentrations. The main drawback of this technique is the expensiveness of the oven and permeation tubes.

Dynamic methods also include the injection of droplets into the flow of diluting gas (usual error has been reported as 5–9% coefficient of variation). Other methods include diffusion (similar to a permeation oven in performances) and evaporation (bubbling). Evaporation is only viable for low vapor pressure compounds, otherwise the inaccuracy is too great (5–15%) [35].

The last usual possibility is to use commercial gas mixtures, sold in cylinders. They use precise gravimetry to reach part per million (ppm) concentrations (the lower limit varies depending on the compounds). These mixtures can reach high prices for the lower concentrations and have a limited lifetime. Due to stability issues, some gas mixtures cannot be stored for long before decaying. Akamatsu et al. used cylinders for some of their target gases (acetone, MiBK) and their "background mix" that simulated common interfering compounds [19]. For ppb-level concentrations, a dilution unit with MFCs is advisable to dilute cylinder gas further. MFCs are the standard gas-mixing tools for air pollutants [37], which are often in low concentrations—such as in breath VOCs. Among the notable uses, carbon dioxide-enriched air is easily available and is used as a diluent in order to recreate the real sampling conditions of breath [37].

2.3.1. Composition of Breath

Breath has several rather unusual characteristics as a gas sample: temperature close to the body's, increased carbon dioxide content (4–5%) and being saturated in humidity. These aspects influence MOS sensor behavior: water vapor and CO_2 interfere with sensors

as experienced by Gregis et al. [39]. These aspects are therefore important to consider when testing MOS sensors for breath sensing.

Variations in breath VOC mixtures have been shown to be linked with cancer. Therefore, it has been hypothesized that some of these VOCs are lung cancer biomarkers. Biomarker identification for lung cancer is a major research effort that has been going on for several decades. Even if several studies have shown similarities in their results, one cannot help but notice that there has been no consistent and validated list of biomarkers for both the clinical and in vitro studies. As stated by several sources, the use of biomarkers in breath for medical screening is a pretty new approach, and therefore suffers from a lack of standardized biomarker identification procedures. The complexity of the analyte and the variety of study objectives enhance the problem [40–43].

Regarding confounding factors, Tan et al. and Shlomi et al. showed that it was possible to differentiate people based on their smoking history, i.e., if they were heavy smokers, light smokers or never-smokers [20,44]. It has been widely reported that smoking history has a strong influence on the volatilome [45]. Previous studies by another team found no correlation between the sensor's output and smoking history, showing that it is possible to create a device that sees past usual confounders [46]. For more information on confounders, a short review of other confounding factors can be found in a previous publication [47] and in the work of Jia et al. [40].

2.3.2. On the Selection of Target VOCs

The first step for breath LC detection is to know the VOCs characteristic of a "healthy patient" or a "cancer patient", and in which concentrations they are found. Numerous gas chromatography mass spectrometry (GC–MS) studies aiming at cancer breath characterization can be found in literature.

A systematic literature review was conducted following the PRISMA method and the PRISMA-P checklist [48], on Scopus, PubMed and University of Liège databases. The following keywords were used: *("VOC" OR "biomarker" OR "breath" OR "chromatography") AND ("lung cancer") NOT ("condensate" OR "in vivo")*. The references of the selected documents were examined to find more pertinent articles. Papers about cancer cell cultures and breath condensate methods were not included, as it is believed that VOCs found in vivo are not similar to those found in vitro [49]. Methods using sorption tubes or SPME were included.

A total of 331 articles published from 1985 to 2021 in English language were evaluated. Meta-analysis and review articles were excluded. Sixty-six articles were found to be relevant to the subject and 21 were excluded for the following reasons: for not sampling cancer patients or having no control group (2), not identifying clear cancer biomarkers (18) or being inaccessible (1). A total of 45 articles were selected and compounds cited as significant for group discrimination were sorted by frequency of citation as potential biomarkers.

The most frequently cited (as biomarkers) compounds are as follows:

- With 12 occurrences, 1-propanol is the most cited compound.
- With 11 occurrences, 2-butanone (or methyl-ethyl-ketone).
- Isoprene has been cited 10 times in the reviewed papers.
- Hexanal, ethylbenzene and 2-propanol have nine occurrences each.
- Acetone was cited eight times.
- Pentane, benzene, and styrene were cited seven times each.
- Hexane, toluene and decane were cited sox times each.
- Propanal, nonanal, heptanal, undecane, 2-methylpentane, pseudocumene and ethanol were cited five times each.

Seven compounds were cited four times, eleven compounds were cited three times, and more than one-hundred and ninety-one compounds were cited two times or less. This last category was not studied any further, as the relevance of each compound was likely to be very low. This review was the basis for the selection of VOCs for use in IOMS evaluation.

It is however important to point out that sources of contamination for the biomarkers are frequently encountered in everyday life, adding incertitude to the conclusions.

The number of patients sampled for each study was recorded, and a sum of the number of controls and cancer patients involved in the discovery of each potential biomarker was calculated. Since studies tend to have a small sample pool, this method was done to enhance biomarkers that are backed by more statistical evidence. The compounds with a sample pool above 600 peoples are the following:

A variety of putative biomarkers was chosen with an educated guess from the list, as already reported in a previous article [47]. Main choosing criterions were the following: short half-life in the body, not reported as smoking-related, found as relevant for studies cumulating numerous test subjects, not found to be exclusively exogenous, not highly correlated with physical activity. However, there is no existing consensus on the quality of each compound as a biomarker. Choosing compounds with certitude is therefore impossible for now, and for many compounds in Table 1 our criterions were not met. Considering this, the chosen compounds for this study were the following: 2-butanone, 2-pentanone, and decane. Decane has been cited in six papers with a reported sample pool of 456.

Table 1. Ranking of compounds by sample pool size (each compound has been cited three times as potential biomarker or more).

Rank	Compound	Pool	Rank	Compound	Pool	Rank	Compound	Pool
1	1-propanol	2267	13	2-propanol	905	25	Cyclohexane	424
2	Isoprene	1840	14	2-pentanone	897	26	Hexane	408
3	2-butanone	1559	15	Benzaldehyde	861	27	Methyl-cyclopentane	408
4	Acetone	1488	16	Pentane	832	28	1,2,4-trimethylbenzene	403
5	3-hydroxy-2-butanone	1285	17	Ethanol	669	29	Ethylacetate	370
6	Pentanal	1151	18	Dimethylsulfide	668	30	Nonanal	339
7	Methanol	1023	19	Benzene	542	31	Octanal	316
8	Hexanal	1019	20	Styrene	540	32	Butanal	244
9	Propanal	999	21	Toluene	512	33	N-dodecane	241
10	Butane	980	22	Decane	456	34	Eicosane	233
11	Undecane	973	23	Heptanal	455	35	2-propenal	125
12	Ethylbenzene	962	24	2-methylpentane	455	36	Hexadecane	117

Several other compounds were acquired for use in the benchmark to better represent the variability and complexity of breath. These were chosen based on the previous list and studies on breath VOC confounding factors [45,50]. Among the compounds frequently found in breath are likely confounders (smoking-related compounds, for example) or biomarkers that are also likely to be exogenous. The compounds are:

- Pentane, 1-propanol, ethanol, dodecane, hexanal (potential biomarkers or behavioral contaminants).
- Acetone (unavoidable metabolism by-product).
- Toluene, 2-propanol (potential biomarker or smoking marker).

Calibration models of IOMS need to include the species of interest and all possible interferents. This enables multivariate analysis with interference detection, without, however, enabling interference compensation [33]. It is however possible to select sensors that have little or no sensitivity for interfering compounds, and therefore improve the performances of the system.

2.3.3. Sampling and Sample Storage

Once the atmospheres were synthesized (or breath has been sampled), they were used directly (inline) or are stored and used some time afterwards (offline).

Offline Sampling

Offline sampling requires some kind of container in order to transfer gas mixes to the sensing device. The most commonly used are sampling bags, whose volume varies a lot between breath sampling studies: 750 mL for end-of-breath alveolar air sampling [51] to 5 L for whole-breath sampling [52]. In a clinical context, the patient would blow in a gas sampling bag, and then the bag is immediately connected to the electronic nose [23,53]. The nose would be equipped with a downstream pump to suck the sample into the sensor chamber at a constant rate. This method, however, cannot isolate end-tidal breath, unless combined with a capnography-based or spirometry-based sampler. Some studies used a pump in an upstream position [24], which is not recommended because of the inherent contamination of the air stream by the pump.

The first point of care while using offline sampling would be the release of VOCs from the bag's constituting materials. Several studies showed that Tedlar™ (Dupont de Nemours™, Wilmington, Delaware)(polyvinyl fluoride) itself emits detectable compounds at the ppb level, such as N,N-dimethylacetamide and phenol, hexane, 2,4-dimethylheptane, 4-methyloctane, as well as CS2, COS, acetonitrile and 1-methoxy-2-propyl acetate [54,55]. Other polymers like Kynar™ (Arkema™ SA, Colombes, France) or Flexfilm™ (SKC™ Ltd., Dorset, UK) also emit detectable species. The contamination for all bags can be reduced; pre-conditioning (nitrogen flushes and heating with an oven) and shorter storage times have a good mitigation power [55]. The bias caused by bags on sensors has been experienced in the literature, especially when switching to a different type of bag or switching to a different supplier; it can be big enough to make it possible for an e-nose to very clearly differentiate populations of samples based on these differences [23].

The second important point is about the bag's membrane is that it permits some diffusion, which is especially noticeable for ppb-level compounds [49]. It is usually advised not to store samples that way for more than 10 h before analysis to avoid significant drift. It is to be noted that water vapor permeates easily through Tedlar®, Kynar® and Flexfilm® bags and the humidity of the sample reaches an ambient level within one to a few hours [54,56]. To avoid sample alteration, storage time should therefore be as short as manageable, down to a few minutes if possible.

Sorption on sampling material is also a known problem. Even with material with low sorption capacity, there is still a possibility for the loss of analytes during sampling. This is one of the reasons why some studies recommend not to use sampling bags if avoidable [57]. Other methods include glass or stainless-steel containers, but these are less convenient and more expensive than bags. Washing bags reduces the problem of sorption by removing adsorbed compounds. It is to be noted that no study was found on sampling bag aging, nor on the evolution of background VOCs over the life cycle of reusable sampling bags. It is therefore possible that the age of the bag has some influence on the contained mixture.

Storage temperature is also a common concern for breath sampling. Since breath is saturated with water vapor, condensation will occur if the bag is not uniformly held at a temperature high enough (45 °C is often indicated for breath-like samples). This is an issue as water in its liquid state can concentrate polar compounds in the gas phase, altering the composition of the breath sample. It is therefore recommended to keep the sample above the dew point temperature [35,36].

Another common procedure in literature is adsorption, with Tenax® (Buchem™ B.V., Apeldoorn, The Netherlands) sorbent for example, followed by thermal desorption. This technique is often valued because it concentrates some compounds and raises the device's sensitivity [58]. However, sorbents are known to alter the samples significantly in a number of ways [39]. For example, some compounds are not well retained by sorbents, and therefore the amount of each of the compounds found after desorption may vary, depending on

each compound's volatility and affinity for the sorbent. Humidity is not highly retained by sorbents such as Tenax®, which is often seen as an advantage [39], but Tenax® does not retain light compounds (such as propanol for example) very well. Sorbents such as Carbograph® 5TD (LARA™ Srl., Olmetti, Italy) retain light compounds better, but tend to retain water more as well [59], and water has to be eliminated before GC analysis by dry purging. It is also worth noting that it has been reported that humidity impairs the adsorption of VOCs on sorbents [60]. One of the main advantage over the bags is the storage stability of sampling tubes over several weeks if refrigerated [61,62].

Inline Sampling

Aside from offline sampling, inline sampling also has some advantages, the most obvious ones being the simplicity of the system by comparison to offline sampling, and the avoidance of some interferences from the sampling apparatus. Several studies used this technique; some experiments in the literature, such as those of Kononov et al., used a rotameter as an indicator for the patient to blow at a defined flow [57]. However, relying on the patient alone to follow an instruction to obtain stable airflow sounds like an important source of variability without a way to obtain a stable flow. One way to do this would be by taking part of the breath with a pump perpendicularly to the flow [63], or by using the method described below.

For both online and offline sampling, the flow speed is of great importance for the sensors. Flow speed conditions regulate heat loss and sample consumption, and to some extent influence sensor responses and their reproducibility. The chosen flow speed varies between studies, but it is always stable. For example, Kononov et al. used a very high flow speed (3.5 L/min) for 10 s only. The sensor chamber was then kept without any flow for the time necessary for sensor signal stabilization [57]. This is most convenient method for inline sampling, as it permits normal breath flow speed (between 11 and 23 L/min in the testing conditions of [26]). The drawback is that the heat buildup might become a problem without active temperature stabilization. Similar methods are used by another workgroup [64,65]. For offline sampling, air flows are often lower, around 100–200 mL/min [16,66]. As large gas samples are inconvenient to handle and take up space (especially considering the need to keep them above body temperature), it is often advantageous to use a low volume of sample, which is adequate for alveolar breath sampling. For continuous flow systems, flow needs to be low to keep sample consumption moderate, since sensors need incompressible time to output a stable signal. A few hundred milliliters per minute are usually enough for heat dissipation and is an adequate value for many applications [28,67].

Between samples, sensors need to be zeroed by exposure to a reference air. For inline sampling, a "flush" of very high flow speed air to wash out the previous sample is often applied, but for offline sampling, a simple three-way valve is used to keep the flow speed constant and switch between samples and reference instantaneously [68]. The nature of reference air varies from room air to pure analytical air. For portable setups, using ambient air is appealing, as air canisters are often impractical, cumbersome, and heavy. In this case, a VOC filter is advisable to remove environmental contamination. Filters need to be combined with a humidifier, as most VOC filters also tend to dry the air. Reference air should therefore be humidified to an easily reproducible and stable level [69]. The use of pure distilled water is recommended to avoid contamination. Obtaining a level of relative humidity close to 90% at 38 °C is recommended, as it is closer to breath.

2.4. Data Treatment

Raw sensor data are rarely used directly to qualify a sensor or discriminate samples. Sensor output (conductance in µS) is often registered every second or several times per minute, and the information obtained by each measurement is far too plentiful. It has to be refined before use.

In most papers, the extracted features are the following, or a combination of the following [70,71]:

- The steady-state response (the mean, median or maximum value of the signal's plateau), usually with baseline subtraction to enhance response reproducibility and comparability. The baseline is the steady-state response of the reference air.
- The area under the curve (AUC), with baseline subtraction as well.
- The greatest ascending or descending slope, to give a measurement of the transient response). If the sensor is well-behaved (such as a polymer sensor or quartz microbalance sensor), it is possible to use the transient response as a predictor of the steady state response to reduce measurement time.

These extracted features are done after other data processing techniques, as necessary. These include sensor signal processing (e.g., noise filtering), normalization and standardization, which are the usual steps [12]. Normalization is useful to mitigate the effect of concentration, which often improves classification when the center of attention is a change in the ratios of concentrations, and not an identical increase of every compound's concentration. It is possible to find a same mixture in concentrated or diluted form, as ratios of concentration are not influenced by dilution effects. However, concentration is an influencing factor for all gas sensors. Suppressing the effect of concentration is done by using normalization, which is done using the following equation (Equation (1)) [12]:

$$x'_{ij} = \frac{x_{ij}}{\sqrt{\sum_{i=1}^{n} {x_{ij}}^2}} \quad (1)$$

where x_{ij} is a component of the array vector, x'_{ij} is the normalized component, j is the number identifying a sample, and i is the output (e.g., the steady state response) of the sensor number i.

Standardization (Equation (2)) is often used with normalization, as it centers and reduces the data of a sensor (between -1 and 1). It is notably necessary to perform before a PCA analysis, as standardization prevents the analysis giving more weight to sensors that have a higher baseline and response [12].

$$x_{std} = \frac{x_i - \mu}{\sigma} \quad (2)$$

where x_{std} is the standardized sensor output, x_i is the sensor output, μ is the mean of all the sensor's outputs, and σ is the standard deviation of all the sensor's outputs.

After this pre-processing step, other analyses can occur. Most of the time, data structure visualization is employed as a first step after pretreatment, using principal component analysis, for example, as stated earlier in this document (see Section 2.2). Visualizing data structures is useful for researchers that want to identify redundant sensors, marginal data, tendencies, and see clusters in the dataset [72]. In the case of seeing which sensor contributes the most in the discrimination, methods of feature selection come into play. Obtaining this information requires the creation of a testing system with all the sensors to be tested at the same time, and the analysis by that system of a few hundreds of mixtures to discriminate (healthy and cancer breaths). The sensors giving out the most relevant features in a dataset of healthy and cancer breath should be the most interesting to use [73]. From that point, there are several ways to proceed. For example, convex hull algorithms [74] can be used to compute sensor space and estimate the resolving power (i.e., ability to tell the difference between two gas mixtures) of an array of sensors. This can be used to select the best sensors by computing the effect of the removal of a feature on the resolving power.

3. Methodological Proposal

Based on this literature review, a method was developed for sensor selection in the domain of breath sensing for disease detection. Before delving into details, it is important to lay down some general guidelines.

The global idea behind the method is that experimental conditions should not stray away from the intended field application. Samples should keep as many of the breath's

characteristics as manageable. The testing device itself should be designed to be as close as possible to the final prototype (i.e., ensuring the sensors are tested in conditions close to those of the final prototype). Otherwise, relevancy might be lost and conclusions only valid in a laboratory environment might be drawn [75].

Making breath-like mixes is difficult due to the part per billion (ppb)-level concentration of biomarkers. Therefore, it requires a lot of rigor and care to avoid interferences; any source of contamination, no matter how slight, can become of great importance for the final results, even if it has been shown that a sensor array could keep good discrimination with confounding factors in play [13]. For good measure, the efficiency and reproducibility of the dilution has to be confirmed by a validated reference method, such as TD–GC–MS quantitative analysis. However, a larger variability in the atmosphere synthesis can be accepted, as electronic noses themselves are, in general, not as accurate or sensitive as a GC–MS. Error sources tend to pile up for most sample creation methods. Due to the unusually low concentrations, the total relative incertitude is often around 10 to 20% of the concentration value, depending on the method and the chemical species (for example, very light species that are difficult to adsorb on sorbents, or heavier species that tend to condensate because they have low vapor pressure, have been observed to be more troublesome during the trials made alongside this article).

There are several ways to transform a breath sample to suppress or modify problematic aspects—e.g., sorbents and thermal desorption to raise concentration, desiccants to remove water [76], filters to protect the inner parts of a device from bacterial and viral contamination. If a device has some of these specificities incorporated in its design, the most precious advice is to make it so synthetic atmospheres are processed in the same way as breath would. It is risky to try to foresee the impact of these modifications on the sensors' behavior and the quality of the discrimination without including them in all experiments from the start. Gaseous samples should undergo any treatments that future real breath samples would receive.

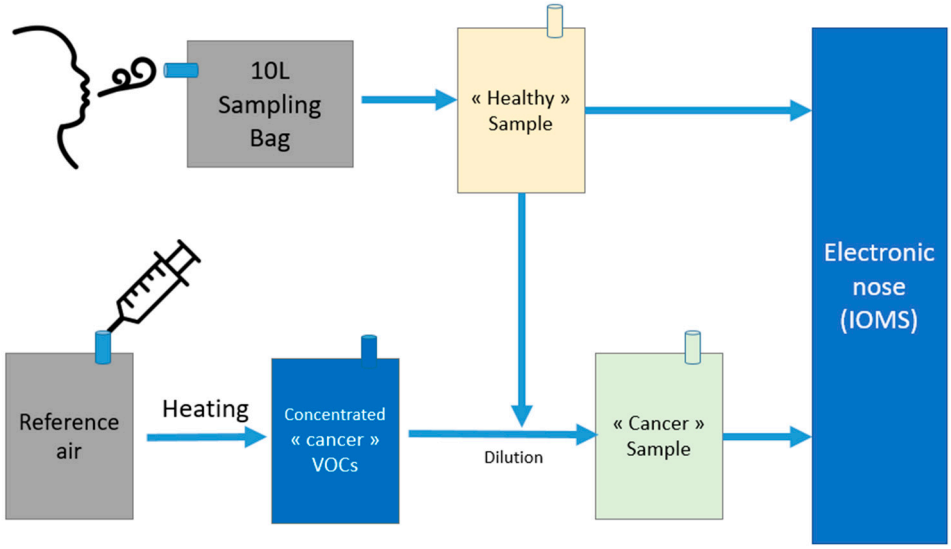
3.1. Choice of VOCs

From the most-cited compounds in literature, four have been picked as interesting (see Section 2.3.2). Their drawbacks (i.e., retention time in body, correlation to lifestyle aspects or non-cancer related aspects) are less important and their contamination sources are less common than others—however, the hypothesis that the sources are simply less known cannot be discarded, unfortunately. Within the existing body of knowledge, they look like interesting candidates.

3.1.1. Using Real Breath as Dilution Gas

A way to test the efficiency of an IOMS to detect the subtle differences between cancer breath and healthy breath is to use healthy breath as the dilution medium for test samples, as shown on Scheme 3. Several breaths from healthy volunteers are sampled in gas bags, and from each breath sample two test samples are created: one is the breath without alteration (labelled "healthy") and the other is the same breath with cancer-related VOCs added. All samples are analyzed by the IOMS. If it is performing well enough, it should be able to tell the difference between breath with and without VOC additions. This enables to test classification performances prior to lengthy and costly clinical trials.

Of course, the amount of VOC added should be done in accordance to real cancer breath concentrations: it should be an amount that "transforms" a healthy breath into a cancer breath. The difference between median concentrations of compounds in both groups should be preferably used to determine the amount of VOCs to be added. Unfortunately, information about median concentrations in breath isn't plentiful. A research group [7,77–80] has been publishing several works where concentrations for healthy, cancer and smoker groups are available for a range of compounds. However, analysis of a larger amount of breaths would be needed to truly be able to establish the fundamental difference between groups, and give more confidence to this approach.



Scheme 3. Sample creation process using real breath for IOMS discrimination testing.

Breath sampling should follow a very strict procedure for such a trial. A good analysis can be found in the work of Amann et al. [81]. Here is a synthesis of the sampling procedure described, enriched with more recent findings in literature.

The sampling place is chosen to be as clean from VOCs sources as possible, with regulated temperature and sufficient ventilation. A 12 h "nothing by mouth" procedure (food, beverages except water, teeth brushing, chewing gums, and smoking are forbidden) is requested of the sampled people. Before sampling, a short questionnaire is filled to know more about the person's habits and lifestyle (diet, medication, sports frequency, smoking, body mass index [82]). The filling of the questionnaire should be done in a sitting position, enabling the person to rest for at least 10 min before sampling. Then, the person is asked to rinse their mouth and gargle three times with 50 mL of tap water, as a work by Ge et al., has shown the benefits of doing so on the quality of the classification [83]. The person is then sampled by using a FEP Teflon bag combined with a bacterial and viral filter (similar to a spirometry filter) to avoid cross contamination [81]. One complete blow at maximum lung capacity is sampled. Alternatively, one could use an end-of-breath sampler to select alveolar breath by capnography [84], as it was shown that the alveolar air is less contaminated by ambient air [85] and likely improves the quality of classification. With alveolar breath, several blows might be necessary to obtain a sufficient volume for analysis. As a side note, it seems that the use of a clean air supply for the patient to breathe in does not remove environmental contamination; due to VOC kinetics in term of their metabolism, half-life, diffusion or adsorption, breathing clean air for a few minutes does not remove the contamination [82,85].

The gas sample bag, once filled with a sufficient volume, should be used for dilution and analyzed as quickly as possible (see Section 2.3.3).

3.1.2. Contaminants

Diet and lifestyle are among the largest causes of exogenous VOC contamination. Almost every study involving breath asked the patients not to drink, eat, smoke, mouthwash or brush their teeth for 2 to 12 h before sampling, depending on the source. These products are VOC-emitting and/or cause a change in the VOCs emitted by bacterial flora in the mouth cavity, esophagus or stomach. Food intake has an impact on the metabolism that can be measured in the breath: acetone is known to vary if the subject is fasting or had a recent sugar intake, since acetone is a by-product of glucose metabolism. Fasting as a source of bias for e-nose detection of cancer has been assessed by McWilliams et al., which found that some combinations of sensors in an array can be significantly affected by fasting, and that fasting tended to make the separation between cancer subjects and control subject

clearer, but that result has to be verified on a larger sample pool [22]. It is therefore very important to include the most common confounders in the testing protocol to evaluate their influence on each sensor. An ideal sensor would be able to give useful information for classification regardless of the variety and concentration of confounders.

Sample storage is a common source of contamination, on which some control is possible. Amongst commonly available sampling bags, Teflon FEP bags have been found to be the most appropriate for VOC analysis due to their low background, low permeability to VOCs and low VOC retention [56,86,87]. Bags should be pre-conditioned and washed after each use. In both cases, the procedure involves filling the bag with dry clean air and heating the bag for some time before emptying it. The process is repeated several times for a cleaning to be complete. A study was conducted in the literature on Tedlar bags (similar to FEP on a lot of aspects) to track the efficacy of cleaning for each heating cycle; heating three times for 30 min is a good compromise between time consumption and cleaning efficiency [54], and was the choice for this method.

Samples in the bag should not be stored for more than a few hours if possible. If storage is unavoidable, we prefer storing under a lowered temperature to decrease the loss of compounds [88]. As condensation is to be expected, heating the bag until complete volatilization after storage is recommended. FEP bags and most sampling bags are permeable to humidity; expect at least 20% humidity loss in the first 30 min and at least 40% in the first 4 h [54].

3.1.3. Sample Creation Procedure

To reach ppb-level concentrations, it is easier to obtain ppm-level concentrations and dilute the mix further afterwards, as previously reported [47]. The insertion of a few microliters of the liquid compounds in a gas sampling bag filled with a known volume of analytical grade pure air, which is then heated (30 min at 60 °C) to ensure complete volatilization, works to reach ppm-level concentrations. It is to be noted that Bastuck et al., recommends the activated charcoal filtering of canister air, even of analytical grade [86]. Injected volumes were calculated in regard to the Ideal gas law to obtain the desired concentration for each compound.

The obtained mixture was diluted to ppb-level using Mass Flow Controllers (MFC) and canisters of 5%-CO_2-enriched air (or real healthy breath, see Section 3.1.1).

The final sample was stored for a short time before usage. Condensation is to be avoided at all costs, therefore all tubing and devices in contact with humidified air were heated (40 °C minimum). Bear in mind that heated gases might lower several aspects, such as:

- Flowmeter accuracy (and therefore dilution ratios and sample VOC concentrations).
- Sorption efficiency on sorption cartridges for TD–GC–MS.

Diluted air was humidified to the desired level using a bubbler filled with pure water, in combination with MFCs for dry and humid air, enabling precise humidity levels to be reached [28].

As it was reported before [47], canisters of CO_2-enriched air were used as the dilution air. Instead of real breath as background, pre-made VOC mix canisters (Westfalen®) can be used to provide interfering compounds (ppm level) to the dilution unit, where their concentration is reduced (ppb level). Each canister must have its own MFC, and only the diluting air (unless it is real breath) should be humidified before mixing.

The 90% relative humidity was picked as the appropriate level, as MOS sensors often saturate in humidity around these levels and do not behave differently when exposed to either 90% or 100% [12]. Less humidity reduces the risk of local condensation, which is why 90% is preferred instead of the breath-like 100%. The bubbler should be dimensioned to enable the complete saturation at the flow speed used in practice (up to 1 L/min in this case). The water should be heated at 40 °C as well to avoid cooling the sample air and reaching absolute level of humidity close to real breath. A certain time is needed for the bubbler to reach the required temperature.

Materials in contact with the samples should be non-emitting, and with no VOC adsorption capacity. Teflon (PTFE, FEP, PFA ...), glass and stainless steel are widely recommended, since they emit very low quantities of unusual and well-known compounds. Other types of material often have a higher VOC signature, or randomly adsorb and desorb compounds, altering the samples and causing cross-contamination. If it is not possible to avoid the use of these materials, proper testing should be made to ensure no sorption or alteration takes place.

For whether sampling would be done inline or offline, the choice fell on the latter. Direct inline sampling was first chosen for its simplicity, but quickly showed a lack of reproducibility, mainly due to unstable blowing rates, and proved itself impractical with regards to the time needed for sensors to stabilize, and the unattainable effort required of the patient. For increased reliability and patient comfort, a gas-sampling bag was chosen as a tampon between the patient and the device.

3.1.4. Sampling and Sensor Testing

The sensitivity of each sensor, be it experimental or commercial in origin, is often only known for a few compounds at most. By using a calibration approach, the objective is knowing the affinities of each sensor for the confounders and biomarkers. This allows for the improvement of experimental sensors (in the case of the Pathacov project [87], a project to detect diseases in breath, which is the frame for this paper) and the creation of arrays with sensors that complement each other, even before feature selection. To avoid creating a specific device (that would include all sensors at once), and also to test sensors in their final working environment, sensor testing was performed in the same setup that would be used for breath sensing. The number of sensors tested at a time was limited to six, and we prioritized sensors to be chosen for discrimination testing. The procedure was as follows:

- Each bag contained a single VOC in a concentration well above those observed in real breath: sensors are rarely sensitive enough to sense a single compound at 1–100 ppb. The collective signal of all VOCs is what is measured. To find the sensitivity to a single compound, a range of concentrations of 500 ppb to 5 ppm are often needed. Four different concentrations were used to obtain a four-point calibration line. This enables the verification that sensor response is indeed linear over the range.
- In addition to VOCs, on different repetitions, humidity was set at 40% RH or 90% RH to check the cross influence of water vapor on the sensor readings.
- Usual confounders were added on different repetitions, at the highest concentration of the tested biomarker VOC for cross influence assessment. Therefore, for a 0.5 to 5 ppm series of samples, the confounder was set at 5 ppm in all samples.

Results were used to create a linear regression model, the slope indicating the sensitivity to the compound in the specific conditions of the test (moisture, confounders ...) giving valuable data on the sensor's general behavior. This will also give insight on the linearity of the response in the commonly encountered sample concentration range.

For evaluation of sensor behavior alteration to varying environmental temperature conditions, the working device was set in a different temperature environment (for example, 10 °C or 35 °C) and their response to a standard gas mixture (e.g., Ethanol 5 ppm) was monitored after temperature acclimation of the device. The standard mixture was humidified to 40% RH at 20 °C.

Drifting tendencies were evaluated over time, after a preconditioning period of two weeks had passed; all calibration data were compared over time to assess the percentage of variation (by computing a ratio between the first calibration and the following ones). Calibration must occur every three months at least. Using principal component analysis, it is even possible to characterize the effect of drift across several compounds. Drift evaluation requires several months at least to be evaluated [68], with no known efficient workaround to speed up the process while remaining true to the actual working conditions of the sensors.

Sensor lifetime corresponds to the median time until sensors are out of order and need to be replaced. For MOS sensors, it can take years before replacement is needed, making evaluation slow and of limited priority at early-stage research. For other types of sensors, at least 5 sensors should be used in ambient conditions until sensors do not react to stimuli anymore. Then the median lifetime between sensors can be used as a metric.

The stabilization period (or pre-conditioning time) for new sensors was evaluated by exposing the sensor every day from startup to a standard gas (Ethanol 5 ppm and 90% RH at 20 °C) until response became reproducible (with less than 5% coefficient of variation for at least two weeks).

The sensors were also tested as an array, to evaluate the type of information they provide when exposed to breath VOCs. Samples were then split into populations of equal size:

- A mix using the concentrations found in the breath of cancer persons.
- A mix using the concentrations found in the breath of healthy persons.
- A mix using the concentrations found in the breath of healthy persons with some common smoking-related VOCs in usual concentrations.
- A mix using the concentrations found in non-cancerous persons having comorbidities frequently associated with lung cancer patients (such as Chronic Obstructive Pulmonary Disease, or COPD).

The composition of these mixtures (Table 2) is inspired from the works of Buszewski et al. [7,77–80] and a few others [89,90], which encompasses 384 lung cancer patients and 645 controls in total for Buszewski et al. Other publications had 70 lung cancer patients and 108 controls in total. The number of publications sharing concentration values is very low, and not all frequently cited compounds are quantified. The composition of the samples submitted for analysis by the electronic nose should be as close as possible to what is indicated in Table 2. The population size for COPD values is however much lower than for other groups, and values should therefore be considered with caution.

Table 2. Concentrations found in several sources from literature for healthy, lung cancer, COPD and smoker groups.

VOC	Healthy Average (ppb)		Smoker Average (ppb)	COPD Average (ppb)		Cancer Average (ppb)	
	[78]	[77] [1]	[78]	[79]	[77] [1]	[78]	[77] [1]
2-butanone	5.1	7	10.6	1.45	6	8.8	9
Decane	-	11	-	0.23	7	-	9
Pentane	104.9	111	108.4	1.87	-	40	11
1-propanol	6.6	61	17	28.15	28	54.8	99
2-propanol	13.3	169	320.7	258.37	92	149.5	398
Ethanol	188.5	193	286.4	218.64	523	467	1203
2-pentanone	5	6	5.3	-	492	7.5	9
Acetone	226	580	330.2	-	19	359	1000
Hexanal	0	3	0	-	719	4.5	4
Toluene	30.9	13	46.8	0.63	4	12.9	7
Benzene	6.3	7	9.2	0.57	7	5.4	5

[1] Source gives concentration as a median instead of an average.

Other non-cancer related diseases can also influence the breath volatilome (e.g., diabetes); a more thorough testing including various mixes inspired by volatilome research literature can be made if it is relevant to the task at hand. The goal would mainly be to evaluate the risk of misclassifying another illness or condition as cancer. It will likely only be necessary if clinical trials reveal a specific source of misclassification. In that case, it is interesting to evaluate sensors on their ability to disregard that source.

The two populations were created from a number of concentrated bags, each having several "daughter" (diluted) bags which are differentiated only by slight variations in their concentrations (maximum 10% deviation from median concentration, ratios between compounds are preserved). This was to simulate variety inside a population. Every bag was submitted to the e-nose several times until completely empty. A fraction of each bag was adsorbed on Tenax, followed by Carbograph 5TD for TDGCMS analysis and composition checking.

Response time was usually measured as the necessary time to reach 90% of the steady-state response in response to a step change in concentration [12]. This was easily measured during calibration.

3.1.5. Gas Chromatography

A reliable and reproducible gas dilution procedure is essential for sensor testing. The validity of the dilution procedure should therefore be checked by a reference method. This also allows for comparison of results between the electronic nose and the mass spectrometry analysis.

The TD–GC–MS method is identical to what was reported before [47] and is based on usual methods for VOC analysis in breath literature.

Semi-quantitative analysis was made after calibration on all target VOCs (see Section 2.3.2). It was previously reported that only four compounds (pentanone, heptanal, decane, toluene) were used. This was mostly for logistical reasons. This set was chosen to better represent the diversity of breath VOCs and more accurately encompass the variety of behaviors, while keeping the number of analytes low. This way, methods were developed to ensure the GC separation and quantification of complex VOCs mixtures. However, for good practices, it is recommended to use all the selected target VOCs for calibration.

For sensor sensitivity evaluation, the calibration range was based on the concentrations used (0.5–5 ppm) instead of what is usually reported in breath (10 ppb−1 ppm). The calibration was made using a linear regression model on the peak area under the curve (AUC). For compounds without a calibration line, the quantification results are expressed in equivalent concentrations (the compound that is closest in retention time). For every gas bag, half was measured by IOMS and the other half was adsorbed on cartridges and analyzed in parallel.

3.2. Gas Sensor Benchmarking Apparatus

It is important to test the sensors in the actual environment they will be most likely used in. The device should be able to control and/or measure as many influencing parameters as possible. The benchmark prototype used for this article has the following features:

- Carbon dioxide measurement, by the mean of a CO_2 sensor. Compact sensors using infrared in the 0–6% range are available on the market. This is an important parameter as most MOS sensors behave differently with differences in the oxygen content of the sample, and it serves as a simple way to ensure no leaks occurred in the system. O_2 sensors can also be used alternatively, but often take more time than CO_2 sensors to reach a stable signal. CO_2 sensors and valves can also be used to select the most interesting part of each exhalation (alveolar air, for example).
- As breath is water-saturated at about 37 °C and sensors are influenced by both temperature and humidity, both parameters are crucial to monitor. A significant problem to consider is condensation. Liquid water in the system could remove some chemical species from the gas phase, which would alter responses from sensors. Accumulation of liquid water can also cause a range of problems (bacteria proliferation, VOC retention and contamination, material degradation, short circuits . . .). The whole system should therefore be heated to avoid condensation.
- The sensor chamber should be heated, and the temperature controlled with a feedback loop (i.e., Proportional-Integral-Derivative controller (PID)) in order to keep the am-

bient temperature around the sensors as stable as possible regardless of the device's surroundings. As sensors are very sensitive to heat loss, this enables reproducible sensing and constant sensor properties [12]. To avoid contamination, the sensor chamber should always be upstream from everything else if possible.

- Flow speed is also an important parameter. Flow speed and sensor chamber design should be chosen to enable laminar flow [28]. A flow control device (e.g., rotameter, Mass Flow Meter, MFC) and pump were placed downstream from the sensor chamber to ensure constant and reproducible flow. Offline analysis of collected samples gave more stable sensor signals than having a patient directly blow into the sensor chamber. Sorbent cartridges instead of gas sampling bags, while manageable, have not been chosen for the testing device. The increased complexity would make the device harder to develop, and it was chosen to not use sorption for the first versions of the device.
- All materials in contact with the samples should be non-emissive and resistant to chemical alterations to avoid sample alteration, as mentioned before [47]. Stainless steel, glass or PTFE are the most appropriate materials in this regard. Other materials should be placed downflow of the sensor chamber. If this is not possible, rigorous testing of their influence on samples should be realized, preferably by using IOMS and a reference method such as GCMS.
- A small volume sensor chamber is preferred, as this avoids the dilution of samples and provides quicker signal stabilization. However, without enough dead volume the heat dissipation of the sensors is likely to be too low, and this might result in lower reproducibility and damage to the sensors if the heaters are under constant voltage. There is no straightforward way to dimension a sensor chamber to have an adequate inner dead volume. The sensor chamber currently being used for this project has an internal dead volume of 7.5 cubic centimeters, has shown excellent heat dissipation properties (40 °C interior temperature while in function without external heating, at room temperature), and the sensors used displayed quick reaction times (40 s mean reaction time). To complement the setup, each sensor heater is controlled by an individual PID to ensure optimum sensor temperature is maintained even under changing or uneven conditions.
- To enable experiment monitoring in real time, measurements can be displayed on a computer showing the evolution of sensor outputs. To ensure all variables are under control and experiments proceed as expected, this feature is invaluable.
- Sensors should be placed so that they are in the same conditions in regard to the flow. In this case, the sensors were placed radially so that they were all perpendicular to the flow in the same way, and therefore were evaluated in the exact same environment.

The IOMS setup was already described in a previous publication [47]. As a reminder, the sensors used were as follows:

- TGS® 2603 (Figaro Engineering™, Osaka, Japan)
- G3530, G1430, G2530, G8530 (Umwelt Sensor Technik® GmbH, Thuringia, Germany)
- MP901 sensors (Winsen™ Electronics Technology Co., Ltd., Zhengzhou, China)
- BME680 (Bosch Sensortec™ GmbH, Reutlingen, Germany)

These sensors were chosen for their differences and apparent complementarity. It is worth noticing that metal oxide sensors need an initial stabilization period ("burn-in") of 3 to 10 days before use, either when starting from cold or for their first usage [12].

The setup includes a carbon dioxide sensor (GSS SprintIR®)(Gas Sensing Solutions Ltd., Cumbernauld, UK). Flow measurements were made using a small size flowmeter (Renesas® FS2012-1020)(Renesas Electronics Co., Tokyo, Japan), which enables flow control in combination with a small size diaphragm pump downstream (RS® PRO D200-03)(RS Group, London, UK).

For this setup, commercial MOS Sensors were used as a first step, but other types of sensors could be used as well with the same general recommendations. Be aware that the sensors' optimal operating conditions may vary: for example, polymer sensors tend to

operate better around room temperature, whereas metal oxide sensors need to be heated to 200–300 °C to function and require heat dissipation to avoid overheating.

Flow speed is an important variable to control in an e-nose, as higher air flows tend to cool down the sensors more. Constant flow speeds enhance comparability between experiments, and therefore should be aimed for.

Measurements on breath have a lot of similarities with those made on car exhausts for emission regulations (i.e., high humidity content, elevated temperature, many different compounds). The interesting approach of samplers for car exhausts comes from the de-dilution of samples by ambient air to prevent condensation from happening in the system [91]. The obvious drawback is that biomarker concentrations are reduced, and breath characterization could prove itself harder with less sensitive sensors. Even if this approach is noteworthy, such a setup has not been chosen for now.

For lung cancer breath detection using a sensor chamber under constant flow, response time should be as quick as possible. The main reason is that time influences the volume of sample necessary to reach steady state, along with flow speed and sensor chamber volume. A quick response time makes it possible to use smaller volumes, which is better logistically and in terms of patient comfort. Recovery time can be longer, however, as ambient air is likely to be used as zero-air. Alternatively, some methods enable the reading of sensor response before a steady-state is reached [70], but how this method performs in real discrimination conditions remains to be studied. Response time can be slower for systems using static air in the sensor chamber during measurement.

Operating conditions should be adequate for the task at hand (i.e., sensor temperature should be high enough to enable the detection of compounds of interest), but keep in mind that a high power consumption and high working temperature might impair the autonomy of remote devices and requires extra heat dissipation modules.

Linearity within the concentration range of the analyte is desirable, as it makes calibration simpler. The ideal sensor would be as insensitive to temperature and humidity shifts as possible, especially since breath has a high humidity content.

Evaluation of the effect of humidity is crucial, as reversible alterations of behavior can be observed, for example, in sensors at lowered temperatures [92].

Drifting tendencies would be kept to a minimum, as it improves reproducibility and reduces the need for recalibration, which is often logistically intensive [86]. For similar reasons, sensor lifetime should be as long as possible, but other factors such as sensor interchangeability and ease of replacement are to be considered in this case [67]. Lastly, a quick stabilization period after startup is preferred for the final user's ease of use.

3.3. Statistical Aspects

Using a certain number of breath-like mixtures, the objective was to characterize how well a sensor performs in a discrimination task, how much information it provides and how it complements other sensors.

When using samples, one important question is "how many?". Luckily, this question has drawn the attention of statisticians for a while and formulas have been developed to identify the optimal number of samples one should aim for. A screening study does not need as many samples as a diagnostic study: high sensitivity is required (sensitivity means correctly identify as many true positives as possible, i.e., low type I error) but low specificity can be tolerated (specificity means correctly identify as many true negatives as possible, i.e., low type II error) [93].

Several factors come into the determination of the necessary sample size: the prevalence of disease (number of positives in the tested population, the fewer positives, the bigger the needed sample) and the expected values of sensitivity and specificity for each outcome (healthy or sick i.e., null or alternative hypothesis). In Bujang et al.'s work, one can find tables to determine adequate sample sizes based on those parameters. It is underlined however that too small sample sizes are to be avoided to obtain a dataset with a sufficient statistical confidence, regardless of what the tables show [93]. Moreover, for most

multivariate analysis and discrimination approaches, large datasets are desirable to obtain robust models.

To know how many positives to expect, it is important to examine the correct population. Since the screening device will be used in people that are most at risk of developing a lung cancer (former smokers or actively smoking people, above 45 years old), the likeliness of encountering a person with lung cancer is higher than if the general population was sampled. Smoking is accountable for 80 to 90% of diagnosed lung cancer [94], but that does not mean that most smokers will develop lung cancer, specifically. It is therefore interesting to investigate Low Dose Computed Tomography (LDCT) trials that processed thousands of people in the population at risk and find how many of them had LC. During the UK Lung Cancer Screening Trial in 2016, 42 individuals were found to have LC amongst the 1994 participants (2.1%) [95]. During the Dutch–Belgian NELSON Trial between 2003–2009, 7557 participants underwent LDCT and 70 were found to have LC (0.9%) in the first round of tests and 54 in the second round (0.7%, or 1.6% in total) [96]. The 2015 NLST trial in the United States enrolled 53 454 participants and detected 1384 LC cases among them (2.6%) [97]. The approximation of the general incidence of lung cancer in the population at risk will therefore be considered to be around 2%.

Here's an example. Considering a *p*-value of 0.05 and aiming for a power of 80% (probability of avoiding a type II error), and a prevalence of 5% (closest to 2%), one should need a minimum of 400 samples to detect a change in the sensitivity and specificity from 0.50 to 0.80, including 20 "sick" samples and 380 "healthy" samples [93]. If one wants the prevalence to be around 2% instead of 5%, it would be advisable to increase the number of "healthy" samples to 980, which of course becomes logistically intensive.

This number of samples is, however, intended for clinical trials, which is an important part of the validation process. In addition, the correspondence of artificial breath to real breath can only be confirmed by actual real breath sampling. On the other hand, the number of samples needed to test the array and individual sensors does not need to be as high as stated above, as artificial breath is not a replacement for clinical trials. For calibration/discrimination testing purposes, three readings of each mixture and each concentration/sample is enough, provided the creation and storage of mixtures has been proven to be reproducible enough. For array testing on complex mixtures, a 50% incidence is considered for ease of testing, which brings down the number of samples to 20 of each kind with the other parameters kept constant. Since this is the minimum value, a 50% increase was decided to ensure sufficient data were gathered for the data analysis, which brought the number of samples to 30 of each kind.

3.4. Data Treatment

3.4.1. Pre-Treatment

The sensor signal is recorded every second, and therefore needs to be pretreated to reduce the amount of information. RStudio®(RStudio, Boston, Massachusetts) was used to create a program in R language to do it automatically, as previously reported [47].

The first part of the code identifies where the signal of each sample starts, ends, and then computes the height of the plateau, the area under the curve, and the start and end slopes, taking into account the baseline signal. Since zero-air has less CO_2 than samples, tracking the CO_2 allows for easy sample tracking; the beginning and end of a measurement is found by looking at sudden and important variations in carbon dioxide levels. This was found to be more reliable than using the second derivative to find the beginning of the sample's signal using conductivity alone.

Normalization (signal of each sensor is divided by the square root of the sum of all signals squared) or standardization (subtraction of the signal's mean and division by the standard deviation) of the array's signals may also be applied to improve the following analyses.

This pre-processed data was recorded and used for Principal Component Analysis (PCA), data structure visualization and multivariate analysis.

Several functionalities are usually implemented at this stage, such as baseline drift compensation to compensate the drift of aging sensors. However, as the current experiments aimed to test the raw characteristics of sensors, and did not aim to obtain the best e-nose reliability over time, this step could be overlooked at this first stage.

3.4.2. Principal Component Analysis

Score plots and the associated loading plots of a PCA were considered. This is particularly useful for array testing, and less for calibration data. Using these, it was possible to highlight the ability of the e-nose to identify differences between samples, and the contribution of each variable to each component. This way, it was possible to obtain information regarding:

- The quality of the separation between groups, by looking at the location and spread of clusters of samples. The further away groups are from each other, and the less overlapping there is, the better the classification by multivariate analysis will be, and therefore the higher the quality of the array is.
- The effect of the modification of one variable on the results, appearing as a shift in data (variation in VOC concentrations or humidity, for example).
- The contribution of each sensor to the separation of groups. Using the loading plot, it is possible to see if a feature is either redundant, not contributing to the separation, or important for the separation. One can therefore identify which sensors are worth keeping and which should be replaced or improved. It is however important to keep in mind that artificial mixtures are not breath, and that an apparently redundant sensor might become useful using real breath. The hypothesis that mixtures give a good enough representation of actual breath must be verified and will be exposed in future work. If it is indeed verified, it will be possible to discard or keep sensors based on the mixtures' results alone, prior to any form of clinical trial.
- The effect of each feature (signal slopes, AUC, maximum value), the effect of each data treatment (e.g., with or without normalization, noise cancelling . . .).

Once the best contributing sensors have been identified, the configuration of the benchmark e-nose is modified by replacing redundant or less contributing sensors. Then the array testing method described in Section 3.1.4 is repeated until optimal sensor configuration is obtained. Then, clinical trials may begin and comparisons between lab and clinical results may give insight on the quality of the benchmark and ways to optimize it.

4. Discussion

Using an IOMS for lung cancer detection is not a novel idea, as tens of references can be found; various projects with varying approaches tried to tackle the problem. So far, none of them investigated the subject of gas mixture creation for testing a medical device.

This publication presents a way to create reproducible complex gas mixtures at the ppb level, using low-cost equipment, and the subsequent usage of these mixtures to test sensors and compare them with each other and with a reference method (TD–GCMS). This methodology is usable for a wide range of VOC-related applications but is mainly intended for the validation of IOMS built for medical diagnostics.

The method is currently being used to test a device designed to detect early-stage lung cancer. The quality of the classification using artificial atmospheres will be compared with the quality of the classification using real breath samples, to assess if conclusions drawn from testing with artificial atmospheres are similar to those drawn from testing with real breath samples. If correspondence is indeed confirmed, this would enable device conception and optimization from a laboratory environment and overall bettered performances from the very first field tests.

The complexity of the atmospheres created, and the composition based on volatolomics research from literature and project partners and the calibration approach all form potential improvements of what is usually performed during lab testing for medical electronic nose. Usual practices often lack relevance regarding the final use, often to simplify the setup and

improve reproducibility. By the present method, the goal is to prove that it is possible to obtain complex mixtures with good reproducibility while remaining true to the intended objective. Results and a detailed analysis of the method once applied will be published in another article.

Author Contributions: Conceptualization, J.D.M.M. and A.-C.R.; methodology, J.D.M.M.; software, J.D.M.M.; validation, J.D.M.M.; formal analysis, J.D.M.M.; investigation, J.D.M.M.; resources, A.-C.R.; data curation, J.D.M.M.; writing—original draft preparation, J.D.M.M.; writing—review and editing, A.-C.R.; visualization, J.D.M.M.; supervision, A.-C.R.; project administration, A.-C.R.; funding acquisition, A.-C.R. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Interreg V France-Wallonie-Vlaanderen, under the Pathacov project (funding number 1.1.297).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: The authors wish to thank Laurent COLLARD, Simon-Pierre LIEGEOIS Eduardo GONçALVES and Noémie MOLITOR for their help in the realization of the experiments. The authors wish to thank the partners of the PATHACOV project for the helpful discussions.

Conflicts of Interest: The authors declare no conflict of interest. This project is financed by Interreg France-Wallonie-Vlaanderen. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. National Cancer Institute SEER Explorer. Available online: https://seer.cancer.gov/statfacts/html/lungb.html (accessed on 22 February 2021).
2. Ferlay, J.; Colombet, M.; Soerjomataram, I.; Dyba, T.; Randi, G.; Bettio, M.; Gavin, A.; Visser, O.; Bray, F. Cancer Incidence and Mortality Patterns in Europe: Estimates for 40 Countries and 25 Major Cancers in 2018. *Eur. J. Cancer* **2018**, *103*, 356–387. [CrossRef] [PubMed]
3. Alberg, A.J.; Brock, M.V.; Ford, J.G.; Samet, J.M.; Spivack, S.D. Epidemiology of Lung Cancer. *Chest* **2013**, *143*, e1S–e29S. [CrossRef]
4. Silvestri, G.A.; Pastis, N.J.; Tanner, N.T.; Jett, J.R. Clinical Aspects of Lung Cancer. In *Murray and Nadel's Textbook of Respiratory Medicine*; Elsevier: Amsterdam, The Netherlands, 2016; pp. 940–964.e22; ISBN 978-1-4557-3383-5.
5. Zhou, J.; Huang, Z.-A.; Kumar, U.; Chen, D.D.Y. Review of Recent Developments in Determining Volatile Organic Compounds in Exhaled Breath as Biomarkers for Lung Cancer Diagnosis. *Anal. Chim. Acta* **2017**, *996*, 1–9. [CrossRef] [PubMed]
6. Bajtarevic, A.; Ager, C.; Pienz, M.; Klieber, M.; Schwarz, K.; Ligor, M.; Ligor, T.; Filipiak, W.; Denz, H.; Fiegl, M.; et al. Noninvasive Detection of Lung Cancer by Analysis of Exhaled Breath. *BMC Cancer* **2009**, *9*, 348. [CrossRef] [PubMed]
7. Buszewski, B.; Ligor, T.; Jezierski, T.; Wenda-Piesik, A.; Walczak, M.; Rudnicka, J. Identification of Volatile Lung Cancer Markers by Gas Chromatography–Mass Spectrometry: Comparison with Discrimination by Canines. *Anal. Bioanal. Chem.* **2012**, *404*, 141–146. [CrossRef] [PubMed]
8. Conrad, D.H.; Goyette, J.; Thomas, P.S. Proteomics as a Method for Early Detection of Cancer: A Review of Proteomics, Exhaled Breath Condensate, and Lung Cancer Screening. *J. Gen. Intern. Med.* **2008**, *23*, 78–84. [CrossRef] [PubMed]
9. Horváth, I.; Barnes, P.J.; Loukides, S.; Sterk, P.J.; Högman, M.; Olin, A.-C.; Amann, A.; Antus, B.; Baraldi, E.; Bikov, A.; et al. A European Respiratory Society Technical Standard: Exhaled Biomarkers in Lung Disease. *Eur. Respir. J.* **2017**, *49*, 1600965. [CrossRef] [PubMed]
10. Romain, A.-C.; Nicolas, J.; Andre, P. Three Years Experiment with the Same Tin Oxide Sensor Arrays for the Identification of Malodorous Sources in the Environment. *Sens. Actuators. B Chem.* **2002**, *84*, 271–277. [CrossRef]
11. Persaud, K.; Dodd, G. Analysis of Discrimination Mechanisms in the Mammalian Olfactory System Using a Model Nose. *Nature* **1982**, *299*, 352–355. [CrossRef] [PubMed]
12. Gardner, J.W.; Bartlett, P.N. *Electronic Noses, Principles and Applications*; IOP Publishing Ltd.: New York, NY, USA, 1999.
13. Gasparri, R.; Santonico, M.; Valentini, C.; Sedda, G.; Borri, A.; Petrella, F.; Maisonneuve, P.; Pennazza, G.; D'Amico, A.; Di Natale, C.; et al. Volatile Signature for the Early Diagnosis of Lung Cancer. *J. Breath Res.* **2016**, *10*, 016007. [CrossRef] [PubMed]
14. Janfaza, S.; Banan Nojavani, M.; Nikkhah, M.; Alizadeh, T.; Esfandiar, A.; Ganjali, M.R. A Selective Chemiresistive Sensor for the Cancer-Related Volatile Organic Compound Hexanal by Using Molecularly Imprinted Polymers and Multiwalled Carbon Nanotubes. *Microchim. Acta* **2019**, *186*, 137. [CrossRef] [PubMed]
15. Dragonieri, S. An Electronic Nose in Respiratory Disease. Ph.D. Thesis, University of Amsterdam, Amsterdam, The Netherlands, 2012.

16. Chatterjee, S.; Castro, M.; Feller, J.F. An E-Nose Made of Carbon Nanotube Based Quantum Resistive Sensors for the Detection of Eighteen Polar/Nonpolar VOC Biomarkers of Lung Cancer. *J. Mater. Chem. B* **2013**, *1*, 4563–4575. [CrossRef] [PubMed]
17. Chang, J.-E.; Lee, D.-S.; Ban, S.-W.; Oh, J.; Jung, M.Y.; Kim, S.-H.; Park, S.; Persaud, K.; Jheon, S. Analysis of Volatile Organic Compounds in Exhaled Breath for Lung Cancer Diagnosis Using a Sensor System. *Sens. Actuators B Chem.* **2018**, *255*, 800–807. [CrossRef]
18. Jaeschke, C.; Padilla, M.; Turppa, E.; Polaka, I.; Gonzalez, O.; Mitrovics, J. Overview on Sniffphone: A Portable Device for Disease Diagnosis. In Proceedings of the 2019 IEEE International Symposium on Olfaction and Electronic Nose (ISOEN), Fukuoka, Japan, 26–29 May 2019; pp. 298–299.
19. Akamatsu, T.I.T.; Tsuruta, A.; Shin, W. Selective Detection of Target Volatile Organic Compounds in Contaminated Humid Air Using a Sensor Array with Principal Component Analysis. *Sensors* **2017**, *17*, 1662. [CrossRef] [PubMed]
20. Shlomi, D.; Abud, M.; Liran, O.; Bar, J.; Gai-Mor, N.; Ilouze, M.; Onn, A.; Ben-Nun, A.; Haick, H.; Peled, N. Detection of Lung Cancer and EGFR Mutation by Electronic Nose System. *J. Thorac. Oncol.* **2017**, *12*, 1544–1551. [CrossRef]
21. van de Goor, R.; van Hooren, M.; Dingemans, A.-M.; Kremer, B.; Kross, K. Training and Validating a Portable Electronic Nose for Lung Cancer Screening. *J. Thorac. Oncol.* **2018**, *13*, 676–681. [CrossRef]
22. Kort, S.; Tiggeloven, M.M.; Brusse-Keizer, M.; Gerritsen, J.W.; Schouwink, J.H.; Citgez, E.; de Jongh, F.H.C.; Samii, S.; van der Maten, J.; van den Bogart, M.; et al. Multi-Centre Prospective Study on Diagnosing Subtypes of Lung Cancer by Exhaled-Breath Analysis. *Lung Cancer* **2018**, *125*, 223–229. [CrossRef]
23. McWilliams, A.; Beigi, P.; Srinidhi, A.; Lam, S.; MacAulay, C.E. Sex and Smoking Status Effects on the Early Detection of Early Lung Cancer in High-Risk Smokers Using an Electronic Nose. *IEEE Trans. Biomed. Eng.* **2015**, *62*, 2044–2054. [CrossRef]
24. Kou, L.; Zhang, D.; Liu, D. A Novel Medical E-Nose Signal Analysis System. *Sensors* **2017**, *17*, 402. [CrossRef]
25. Di Natale, C.; Macagnano, A.; Martinelli, E.; Paolesse, R.; D'Arcangelo, G.; Roscioni, C.; Finazzi-Agrò, A.; D'Amico, A. Lung Cancer Identification by the Analysis of Breath by Means of an Array of Non-Selective Gas Sensors. *Biosens. Bioelectron.* **2003**, *18*, 1209–1218. [CrossRef]
26. de Vries, R.; Brinkman, P.; van der Schee, M.P.; Fens, N.; Dijkers, E.; Bootsma, S.K.; de Jongh, F.H.C.; Sterk, P.J. Integration of Electronic Nose Technology with Spirometry: Validation of a New Approach for Exhaled Breath Analysis. *J. Breath Res.* **2015**, *9*, 046001. [CrossRef] [PubMed]
27. Dai, W.; Berleant, D. Benchmarking Contemporary Deep Learning Hardware and Frameworks: A Survey of Qualitative Metrics. In Proceedings of the 2019 IEEE First International Conference on Cognitive Machine Intelligence (CogMI), Los Angeles, CA, USA, 12–14 December 2019; pp. 148–155.
28. Endres, H.-E.; Jander, H.D.; Göttler, W. A Test System for Gas Sensors. *Sens. Actuators B Chem.* **1995**, *23*, 163–172. [CrossRef]
29. (PDF) Data Analysis for Electronic Nose Systems. Available online: https://www.researchgate.net/publication/226437091_Data_analysis_for_electronic_nose_systems (accessed on 14 February 2020).
30. Johnson, K.J.; Rose-Pehrsson, S.L. Sensor Array Design for Complex Sensing Tasks. *Annual Rev. Anal. Chem.* **2015**, *8*, 287–310. [CrossRef] [PubMed]
31. Xu, Z.; Shi, X.; Lu, S. Integrated Sensor Array Optimization with Statistical Evaluation. *Sens. Actuators B Chem.* **2010**, *149*, 239–244. [CrossRef]
32. Gasparri, R.; Romano, R.; Sedda, G.; Borri, A.; Petrella, F.; Galetta, D.; Casiraghi, M.; Spaggiari, L. Diagnostic Biomarkers for Lung Cancer Prevention. *J. Breath Res.* **2018**, *12*, 027111. [CrossRef] [PubMed]
33. Hierlemann, A.; Gutierrez-Osuna, R. Higher-Order Chemical Sensing. *Chem. Rev.* **2008**, *108*, 563–613. [CrossRef]
34. Kermani, B.G.; Schiffman, S.S.; Nagle, H.T. Performance of the Levenberg–Marquardt Neural Network Training Method in Electronic Nose Applications. *Sens. Actuators B Chem.* **2005**, *110*, 13–22. [CrossRef]
35. Li, Y.; Täffner, T.; Bischoff, M.; Niemeyer, B. Test Gas Generation from Pure Liquids: An Application-Oriented Overview of Methods in a Nutshell. *Int. J. Chem. Eng.* **2012**, *2012*, e417029. [CrossRef]
36. Namies´nik, J. Generation of Standard Gaseous Mixtures. *J. Chromatogr. A* **1984**, *300*, 79–108. [CrossRef]
37. Walden, J.; Macé, T.; Haerri, H.-P.; Sutour, C.; Couette, J.; Niederhauser, B. *Guide on Dynamic Dilution Methods for NO, NO_2 and SO_2 at Limit Values*; European Metrology Research Program (EMRP): Braunschweig, Germany, 2014; p. 31.
38. Helwig, N.; Schüler, M.; Bur, C.; Schütze, A.; Sauerwald, T. Gas Mixing Apparatus for Automated Gas Sensor Characterization. *Meas. Sci. Technol.* **2014**, *25*, 055903. [CrossRef]
39. Gregis, G.; Sanchez, J.B.; Bezverkhyy, I.; Guy, W.; Berger, F.; Fierro, V.; Bellat, J.P.; Celzard, A. Detection and Quantification of Lung Cancer Biomarkers by a Micro-Analytical Device Using a Single Metal Oxide-Based Gas Sensor. *Sens. Actuators B Chem.* **2018**, *255*, 391–400. [CrossRef]
40. Jia, Z.; Patra, A.; Kutty, V.K.; Venkatesan, T. Critical Review of Volatile Organic Compound Analysis in Breath and in Vitro Cell Culture for Detection of Lung Cancer. *Metabolites* **2019**, *9*, 52. [CrossRef] [PubMed]
41. Tang, Z.; Liu, Y.; Duan, Y. Breath Analysis: Technical Developments and Challenges in the Monitoring of Human Exposure to Volatile Organic Compounds. *J. Chromatogr. B Anal. Technol. Biomed. Life Sci.* **2015**, *1002*, 285–299. [CrossRef] [PubMed]
42. Pennazza, G.; Santonico, M. *Breathprinting Roadmap Based on Experts' Opinions*; Elsevier Inc.: Amsterdam, The Netherlands, 2018; ISBN 9780128145623.
43. Herbig, J.; Beauchamp, J. Towards Standardization in the Analysis of Breath Gas Volatiles. *J. Breath Res.* **2014**, *8*, 037101. [CrossRef]

44. Tan, J.-L.; Yong, Z.-X.; Liam, C.-K. Using a Chemiresistor-Based Alkane Sensor to Distinguish Exhaled Breaths of Lung Cancer Patients from Subjects with No Lung Cancer. *J. Thorac. Dis.* **2016**, *8*, 2772–2783. [CrossRef]
45. Filipiak, W.; Ruzsanyi, V.; Mochalski, P.; Filipiak, A.; Bajtarevic, A.; Ager, C.; Denz, H.; Hilbe, W.; Jamnig, H.; Hackl, M.; et al. Dependence of Exhaled Breath Composition on Exogenous Factors, Smoking Habits and Exposure to Air Pollutants. *J. Breath Res.* **2012**, *6*, 036008. [CrossRef]
46. Peng, G.; Hakim, M.; Broza, Y.Y.; Billan, S.; Abdah-Bortnyak, R.; Kuten, A.; Tisch, U.; Haick, H. Detection of Lung, Breast, Colorectal, and Prostate Cancers from Exhaled Breath Using a Single Array of Nanosensors. *Br. J. Cancer* **2010**, *103*, 542–551. [CrossRef]
47. Martin, J.D.M.; Romain, A.-C. Experimental Evaluation of Gas Sensors Array for the Identification of Complex Voc Mixtures in Human Breath. *Chem. Eng. Trans.* **2021**, *85*, 199–204. [CrossRef]
48. PRISMA-P Group; Moher, D.; Shamseer, L.; Clarke, M.; Ghersi, D.; Liberati, A.; Petticrew, M.; Shekelle, P.; Stewart, L.A. Preferred Reporting Items for Systematic Review and Meta-Analysis Protocols (PRISMA-P) 2015 Statement. *Syst. Rev.* **2015**, *4*, 1. [CrossRef]
49. Pesesse, R. Contribution of Comprenhensive Two-Dimensional Gas Chromatography to Untargeted Volatilomics of Lung Cancer. Ph.D. Thesis, Université de Liège, Sart-Tilman, Liege, Belgium, 2019.
50. Hakim, M.; Broza, Y.Y.; Barash, O.; Peled, N.; Phillips, M.; Amann, A.; Haick, H. Volatile Organic Compounds of Lung Cancer and Possible Biochemical Pathways. *Chem. Rev.* **2012**, *112*, 5949–5966. [CrossRef]
51. Barash, O.; Peled, N.; Hirsch, F.R.; Haick, H. Sniffing the Unique "Odor Print" of Non-Small-Cell Lung Cancer with Gold Nanoparticles. *Small* **2009**, *5*, 2618–2624. [CrossRef] [PubMed]
52. Chen, X.; Cao, M.; Li, Y.; Hu, W.; Wang, P.; Ying, K.; Pan, H. A Study of an Electronic Nose for Detection of Lung Cancer Based on a Virtual SAW Gas Sensors Array and Imaging Recognition Method. *Meas. Sci. Technol.* **2005**, *16*, 1535–1546. [CrossRef]
53. Biehl, W.; Hattesohl, A.; Jörres, R.A.; Duell, T.; Althöhn, U.; Koczulla, A.R.; Schmetzer, H. VOC Pattern Recognition of Lung Cancer: A Comparative Evaluation of Different Dog- and ENose-Based Strategies Using Different Sampling Materials. *Acta Oncol.* **2019**, *58*, 1216–1224. [CrossRef] [PubMed]
54. Beauchamp, J.; Herbig, J.; Gutmann, R.; Hansel, A. On the Use of Tedlar®Bags for Breath-Gas Sampling and Analysis. *J. Breath Res.* **2008**, *2*, 046001. [CrossRef] [PubMed]
55. Mochalski, P.; King, J.; Unterkofler, K.; Amann, A. Stability of Selected Volatile Breath Constituents in Tedlar, Kynar and Flexfilm Sampling Bags. *Analyst* **2013**, *138*, 1405. [CrossRef] [PubMed]
56. Mochalski, P.; Wzorek, B.; Śliwka, I.; Amann, A. Suitability of Different Polymer Bags for Storage of Volatile Sulphur Compounds Relevant to Breath Analysis. *J. Chromatogr. B* **2009**, *877*, 189–196. [CrossRef]
57. Kononov, A.; Korotetsky, B.; Jahatspanian, I.; Gubal, A.; Vasiliev, A.; Arseniev, A.; Nefedov, A.; Barchuk, A.; Gorbunov, I.; Kozyrev, K.; et al. Online Breath Analysis Using Metal Oxide Semiconductor Sensors (Electronic Nose) for Diagnosis of Lung Cancer. *J. Breath Res.* **2019**, *14*, 016004. [CrossRef]
58. Gregis, G. Étude et Réalisation d'un Système Miniaturisé Pour l'analyse de Composés Organiques Volatils Considérés Comme Des Marqueurs Chimiques Du Cancer Du Poumon. Ph.D. Thesis, Université de Bourgogne Franche-Comté, Franche-Comté, France, 2017.
59. Strathmann, S. Sample Conditioning for Multi-Sensor Systems. Ph.D. Thesis, Der Fakultät für Chemie und Pharmazie der Eberhard-Karls-Universität Tübingen, Tübingen, Germany, 2001.
60. Wilkinson, M.; White, I.R.; Goodacre, R.; Nijsen, T.; Fowler, S.J. Effects of High Relative Humidity and Dry Purging on VOCs Obtained during Breath Sampling on Common Sorbent Tubes. *J. Breath Res.* **2020**, *14*, 046006. [CrossRef]
61. Harshman, S.W.; Mani, N.; Geier, B.A.; Kwak, J.; Shepard, P.; Fan, M.; Sudberry, G.L.; Mayes, R.S.; Ott, D.K.; Martin, J.A.; et al. Storage Stability of Exhaled Breath on Tenax TA. *J. Breath Res.* **2016**, *10*, 046008. [CrossRef] [PubMed]
62. van der Schee, M.P.; Fens, N.; Brinkman, P.; Bos, L.D.J.; Angelo, M.D.; Nijsen, T.M.E.; Raabe, R.; Knobel, H.H.; Vink, T.J.; Sterk, P.J. Effect of Transportation and Storage Using Sorbent Tubes of Exhaled Breath Samples on Diagnostic Accuracy of Electronic Nose Analysis. *J. Breath Res.* **2012**, *7*, 016002. [CrossRef]
63. Righettoni, M.; Ragnoni, A.; Güntner, A.; Loccioni, C.; Pratsinis, S.; Risby, T. Monitoring Breath Markers under Controlled Conditions. *J. Breath Res.* **2015**, *9*, 047101. [CrossRef] [PubMed]
64. Li, W.; Liu, H.; Xie, D.; He, Z.; Pi, X. Lung Cancer Screening Based on Type-Different Sensor Arrays. *Sci. Rep.* **2017**, *7*, 1969. [CrossRef] [PubMed]
65. Lu, B.; Fu, L.; Nie, B.; Peng, Z.; Liu, H. A Novel Framework with High Diagnostic Sensitivity for Lung Cancer Detection by Electronic Nose. *Sensors* **2019**, *19*, 5333. [CrossRef] [PubMed]
66. Chapman, E.A.; Thomas, P.S.; Stone, E.; Lewis, C.; Yates, D.H. A Breath Test for Malignant Mesothelioma Using an Electronic Nose. *Eur. Respir. J.* **2012**, *40*, 448–454. [CrossRef] [PubMed]
67. Romain, A.C.; Nicolas, J. Long Term Stability of Metal Oxide-Based Gas Sensors for e-Nose Environmental Applications: An Overview. *Sens. Actuators B Chem.* **2010**, *146*, 502–506. [CrossRef]
68. Burlachenko, J.; Kruglenko, I.; Snopok, B.; Persaud, K. Sample Handling for Electronic Nose Technology: State of the Art and Future Trends. *TrAC Trends Anal. Chem.* **2016**, *82*, 222–236. [CrossRef]
69. Montuschi, P.; Mores, N.; Trové, A.; Mondino, C.; Barnes, P.J. The Electronic Nose in Respiratory Medicine. *Respiration* **2013**, *85*, 72–84. [CrossRef]

70. Osorio-Arrieta, D.L.; Muñoz-Mata, J.L.; Beltrán-Pérez, G.; Castillo-Mixcóatl, J.; Mendoza-Barrera, C.O.; Altuzar-Aguilar, V.; Muñoz-Aguirre, S. Reduction of the Measurement Time by the Prediction of the Steady-State Response for Quartz Crystal Microbalance Gas Sensors. *Sensors* **2018**, *18*, 2475. [CrossRef]
71. Marco, S.; Gutierrez-Galvez, A. Signal and Data Processing for Machine Olfaction and Chemical Sensing: A Review. *IEEE Sens. J.* **2012**, *12*, 3189–3214. [CrossRef]
72. Zuppa, M.; Distante, C.; Siciliano, P.; Persaud, K. Drift Counteraction with Multiple Self-Organising Maps for an Electronic Nose. *Sens. Actuators B Chem.* **2004**, *98*, 305–317. [CrossRef]
73. Liu, B.; Yu, H.; Zeng, X.; Zhang, D.; Gong, J.; Tian, L.; Qian, J.; Zhao, L.; Zhang, S.; Liu, R. Lung Cancer Detection via Breath by Electronic Nose Enhanced with a Sparse Group Feature Selection Approach. *Sens. Actuators* **2021**, *339*, 129896. [CrossRef]
74. Fernandez, L.; Yan, J.; Fonollosa, J.; Burgués, J.; Gutierrez, A.; Marco, S. A Practical Method to Estimate the Resolving Power of a Chemical Sensor Array: Application to Feature Selection. *Front. Chem.* **2018**, *6*, 209. [CrossRef] [PubMed]
75. Barsan, N.; Koziej, D.; Weimar, U. Metal Oxide-Based Gas Sensor Research: How To? *Sens. Actuators B Chem.* **2007**, *121*, 18–35. [CrossRef]
76. Huang, C.-H.; Zeng, C.; Wang, Y.-C.; Peng, H.-Y.; Lin, C.-S.; Chang, C.-J.; Yang, H.-Y. A Study of Diagnostic Accuracy Using a Chemical Sensor Array and a Machine Learning Technique to Detect Lung Cancer. *Sensors* **2018**, *18*, 2845. [CrossRef] [PubMed]
77. Rudnicka, J.; Walczak, M.; Kowalkowski, T.; Jezierski, T.; Buszewski, B. Determination of Volatile Organic Compounds as Potential Markers of Lung Cancer by Gas Chromatography–Mass Spectrometry versus Trained Dogs. *Sens. Actuators B Chem.* **2014**, *202*, 615–621. [CrossRef]
78. Ulanowska, A.; Kowalkowski, T.; Trawińska, E.; Buszewski, B. The Application of Statistical Methods Using VOCs to Identify Patients with Lung Cancer. *J. Breath Res.* **2011**, *5*, 046008. [CrossRef]
79. Monedeiro, F.; Monedeiro-Milanowski, M.; Ratiu, I.-A.; Brożek, B.; Ligor, T.; Buszewski, B. Needle Trap Device-GC-MS for Characterization of Lung Diseases Based on Breath VOC Profiles. *Molecules* **2021**, *26*, 1789. [CrossRef]
80. Ligor, T.; Pater, Ł.; Buszewski, B. Application of an Artificial Neural Network Model for Selection of Potential Lung Cancer Biomarkers. *J. Breath Res.* **2015**, *9*, 027106. [CrossRef]
81. Amann, A.; Miekisch, W.; Pleil, J.D.; Risby, T.; Schubert, J. *Methodological Issues of Sample Collection and Analysis of Exhaled Breath*; Chapter 7; Maney Publishing: Leeds, UK, 2010; pp. 96–114.
82. Beauchamp, J. Inhaled Today, Not Gone Tomorrow: Pharmacokinetics and Environmental Exposure of Volatiles in Exhaled Breath. *J. Breath Res.* **2011**, *5*, 037103. [CrossRef]
83. Ge, D.; Zhou, J.; Chu, Y.; Lu, Y.; Zou, X.; Xia, L.; Liu, Y.; Huang, C.; Shen, C.; Zhang, L.; et al. Distinguish Oral-Source VOCs and Control Their Potential Impact on Breath Biomarkers. *Anal. Bioanal. Chem.* **2022**, *414*, 2275–2284. [CrossRef]
84. Herbig, J.; Titzmann, T.; Beauchamp, J.; Kohl, I.; Hansel, A. Buffered End-Tidal (BET) Sampling—A Novel Method for Real-Time Breath-Gas Analysis. *J. Breath Res.* **2008**, *2*, 037008. [CrossRef] [PubMed]
85. Di Gilio, A.; Palmisani, J.; Ventrella, G.; Facchini, L.; Catino, A.; Varesano, N.; Pizzutilo, P.; Galetta, D.; Borelli, M.; Barbieri, P.; et al. Breath Analysis: Comparison among Methodological Approaches for Breath Sampling. *Molecules* **2020**, *25*, 5823. [CrossRef]
86. Bastuck, M. *Improving the Performance of Gas Sensor Systems with Advanced Data Evaluation, Operation, and Calibration Methods*; Linköping Studies in Science and Technology. Dissertations; Linköping University Electronic Press: Linköping, Sweden, 2019; Volume 2009; ISBN 978-91-7685-003-9.
87. Diagnostic de Pathologies Humaines par Analyse de COV dans l'Air Expiré. Available online: https://pathacov-project.com/ (accessed on 6 July 2022).
88. Fugit, J.L.; Dutaur, L.; Simon, V.; Riba, M.L.; Torres, L. Generation and Storage of Standard Gas Mixtures Containing Traces of Analytes. *Fresenius Environ. Bull.* **1996**, *5*, 682–687.
89. Poli, D.; Carbognani, P.; Corradi, M.; Goldoni, M.; Acampa, O.; Balbi, B.; Bianchi, L.; Rusca, M.; Mutti, A. Exhaled Volatile Organic Compounds in Patients with Non-Small Cell Lung Cancer: Cross Sectional and Nested Short-Term Follow-up Study. *Respir. Res.* **2005**, *6*, 71. [CrossRef] [PubMed]
90. Schallschmidt, K.; Becker, R.; Jung, C.; Bremser, W.; Walles, T.; Neudecker, J.; Leschber, G.; Frese, S.; Nehls, I. Comparison of Volatile Organic Compounds from Lung Cancer Patients and Healthy Controls—Challenges and Limitations of an Observational Study. *J. Breath Res.* **2016**, *10*, 046007. [CrossRef]
91. Hood, J.F.; Silvis, W.M. *Predicting and Preventing Water Condensation in Sampled Vehicle Exhaust for Optimal CVS Dilution*; SAE Technical Paper; SAE: Warrendale, PA, USA, 1998; p. 980404.
92. Wang, C.; Yin, L.; Zhang, L.; Xiang, D.; Gao, R. Metal Oxide Gas Sensors: Sensitivity and Influencing Factors. *Sensors* **2010**, *10*, 2088–2106. [CrossRef]
93. Bujang, M.A.; Adnan, T.H. Requirements for Minimum Sample Size for Sensitivity and Specificity Analysis. *J. Clin. Diagn Res.* **2016**, *10*, YE01–YE06. [CrossRef]
94. Alberg, A.J.; Brock, M.V.; Samet, J.M. Epidemiology of Lung Cancer. In *Murray and Nadel's Textbook of Respiratory Medicine*; Elsevier: Amsterdam, The Netherlands, 2016; pp. 927–939.e5; ISBN 978-1-4557-3383-5.
95. Field, J.K.; Duffy, S.W.; Baldwin, D.R.; Brain, K.E.; Devaraj, A.; Eisen, T.; Green, B.A.; Holemans, J.A.; Kavanagh, T.; Kerr, K.M.; et al. The UK Lung Cancer Screening Trial: A Pilot Randomised Controlled Trial of Low-Dose Computed Tomography Screening for the Early Detection of Lung Cancer. *Health Technol. Assess.* **2016**, *20*, 1–146. [CrossRef]

96. van Klaveren, R.J.; Oudkerk, M.; Prokop, M.; Scholten, E.T.; Nackaerts, K.; Vernhout, R.; van Iersel, C.A.; van den Bergh, K.A.M.; van 't Westeinde, S.; van der Aalst, C.; et al. Management of Lung Nodules Detected by Volume CT Scanning. *N. Engl. J. Med.* **2009**, *361*, 2221–2229. [CrossRef]
97. Chudgar, N.P.; Bucciarelli, P.R.; Jeffries, E.M.; Rizk, N.P.; Park, B.J.; Adusumilli, P.S.; Jones, D.R. Results of the National Lung Cancer Screening Trial: Where Are We Now? *Thorac. Surg. Clin.* **2015**, *25*, 145–153. [CrossRef]