Preprints.org

Review

Breathomics: A Non-Invasive Approach to the Diagnosis of Breast Cancer

Hélène Yockell-Lelièvre [*] , Romy Philip , Palash Kaushik , Ashok Prabhu Masilamani , Sarkis H Meterissian

Review

Breathomics: A Non-Invasive Approach to the Diagnosis of Breast Cancer

Hélène Yockell-Lelièvre [1],*, Romy Philip [2] , Palash Kaushik [1], Ashok Prabhu Masilamani [1] and Sarkis H. Meterissian [2]

[1] Noze, 4920 Pl. Olivia, Saint-Laurent, QC, Canada H4R 2Z8

[2] Department of Surgery and Oncology, McGill University, Montreal, QC, Canada H4A 3J1

* Correspondence: hylelievre@noze.ca

Abstract: Breast cancer is the most commonly diagnosed cancer worldwide [1], underscoring the critical need for effective early detection methods to reduce mortality. Traditional detection techniques, such as mammography, present significant limitations, particularly in women with dense breast tissue, highlighting the need for alternative screening approaches. Breathomics, based on the analysis of volatile organic compounds (VOCs) present in exhaled breath, offers a non-invasive, potentially transformative diagnostic tool. These VOCs are metabolic byproducts from various organs of the human body whose presence and varying concentrations in breath are reflective of different health conditions. This review explores the potential of breathomics, highlighting its promise as a rapid, cost-effective screening approach for breast cancer, facilitated through the integration of portable solutions like electronic noses (e-noses). Key considerations for clinical translation - including patient selection, environmental confounders, and different breath collection methods will be examined in terms of how each of them affects the breath profile. However, there are also challenges such as patient variability in VOC signatures, and the need for standardization in breath sampling protocols. Future research should prioritize standardizing sampling and analytical procedures, and validating their clinical utility through large-scale clinical trials.

Keywords: breathomics; breath analysis; breast cancer screening; breast cancer diagnostics; electronic nose; volatile organic compounds

1. Introduction

Breast cancer represents a significant global health concern, as it is currently the most commonly diagnosed cancer worldwide, accounting for one third of cancer diagnoses in women [2]. With its incidence on the rise as the population ages, effective means of early detection are paramount in ameliorating interventions and decreasing cancer mortality. The current diagnostic paradigm for breast cancer typically begins with mammography as an initial screening tool. When abnormalities are detected, further investigations are conducted, including higher-resolution X-rays, ultrasound examinations, magnetic resonance imaging (MRI), and ultimately, biopsy.

These conventional diagnostic procedures present several challenges. They are often time-consuming, costly, and require the expertise of specially trained medical professionals; factors which collectively limit their accessibility. Also, the invasive nature of mammography as a screening tool can cause discomfort, pain and anxiety in patients, enough so to represent a significant motivator for avoiding screening altogether [3,4].

Furthermore, mammography, while being one of the most effective tools for early detection of breast cancer, presents serious limitations, particularly in women with dense breast tissue. About 40% of women have heterogeneously dense breasts, and 10% have extremely dense breasts, where the breast is almost entirely composed of fibroglandular tissue thus potentially obscuring cancers.

[5,6]. The sensitivity of mammography is therefore significantly reduced particularly for younger patients [6].

Given these limitations, there is a pressing need for alternative diagnostic approaches that are reliable, cost-effective, and less invasive. Such methods could potentially improve early detection rates, reduce healthcare costs, and increase accessibility to breast cancer screening, particularly in resource-limited settings. It is within this context that breathomics emerges as a promising avenue for breast cancer diagnosis, offering the potential for quick, easy, and noninvasive screening.

2. Breathomics as a Non-Invasive Diagnostic Tool for Breast Cancer Detection

Breathomics constitutes the comprehensive analysis of exhaled breath to identify and measure volatile organic compounds (VOCs) that may serve as biomarkers for various physiological and pathological conditions. The human breath contains hundreds of VOCs that originate from normal metabolic processes, pathological conditions, environmental exposure, or often many sources at once. These compounds can provide valuable information about an individual's health status [7,8]. Figure 1 illustrates the multiple sources of breath VOCs: many originate from blood, permeating the lung's alveoli, while some originate locally from the metabolism of lung cells. Characteristic ratios of different VOCs in the breath can distinguish between healthy and pathological states. For example, in the case of respiratory infections, pathogens emit specific VOCs that would not otherwise be present in exhaled breath, allowing their identification [9]. The human body also absorbs external VOCs, and this intake, like the emissions, can vary depending on the overall health of the individual. All these factors contribute to creating a complex breath signature that is reflective of an individual's overall health status.
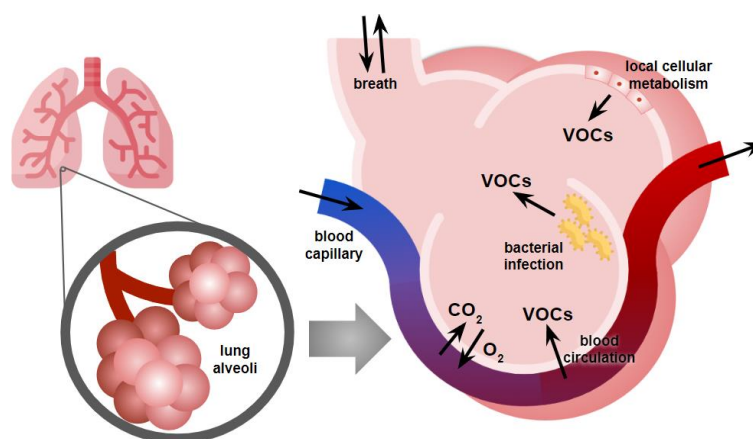

Figure 1. Endogenous sources for volatile organic compounds found in human breath.

Over the last decade, breathomics has emerged as a potential diagnostic tool, with research directed at its application to breast cancer diagnosis in particular [10]. The endogenous metabolic origin of breath VOCs associated with breast cancer primarily involves alterations in cellular metabolism due to cell activity. These metabolic changes lead to the alteration of the emission rate of VOCs, or to the production of specific VOCs. Cancer cells located anywhere in the body are in close proximity to blood vessels therefore VOCs from these cells can travel from the blood to the breath through the lungs.

Bioassays that are commonly used to diagnose different subtypes of breast cancers target specific biomarkers, mostly proteins such as hormone receptors (Human Epidermal Growth Factor Receptor 2 (HER2), Estrogen Receptor (ER), Progesterone Receptor (PR)). These bioassays function with the lock-and-key model, where one test targets one single molecule. Breath VOCs are typically small molecules that are the final product of a complex cascade of biochemical processes. Since the individual breath VOCs can be the product of many different biological origins, they can hardly be

targeted as single, specific biomarkers for a given health condition. What constitutes a biomarker in breathomics is the collective response of the body to a certain condition, which typically leads to a collective change in the concentrations of several VOCs in exhaled breath. This can lead to highly specific results for breath tests, since the collective response of dozens of individual biomarkers provides a disease-specific signature. While the field is very actively researched, the specific biological mechanisms that generate volatile metabolites remain under investigation.

3. Technologies Used in Breathomics Analysis

From research performed during the last decade, breathomics established that it can identify biomarkers for various diseases, including cancers, respiratory disorders, and metabolic conditions, allowing for early detection and intervention. The identification of specific VOCs in exhaled breath is central to establishing breath-based cancer diagnostics, and to date a lot of work has been done in an attempt to identify individual breath biomarkers specifically associated with breast cancer.

Currently, most research is conducted using intensive, lab-based instrumentation like gas chromatography coupled to mass spectrometry (GC-MS), allowing identification and potential quantification of breath biomarkers (Figure 2). GC-MS is renowned for its ability to separate and identify a vast array of VOCs present in exhaled breath down to trace levels, with capacity that is unmatched by other analytical techniques. Since the concentration of VOCs in the breath is still relatively low (in the ppb range), GC-MS requires sample preconcentration, often using sorbent tubes and thermal desorption (TD), or solid phase microextraction (SPME). Using GC-MS, a few research teams have already tentatively identified a few breath volatile biomarkers associated with breast cancer, which can reveal subtle changes in metabolic pathways associated with cancer. For instance, the research team of Michael Phillips, a pioneer in the field, has used TD-GC-MS early on to identify a series of methylated alkanes associated with oxidative stress, a complex series of biochemical processes that produces a characteristic blend of VOCs depending on the specific health condition it is linked with, such as lung cancer [11], tuberculosis [12], or aging [13].
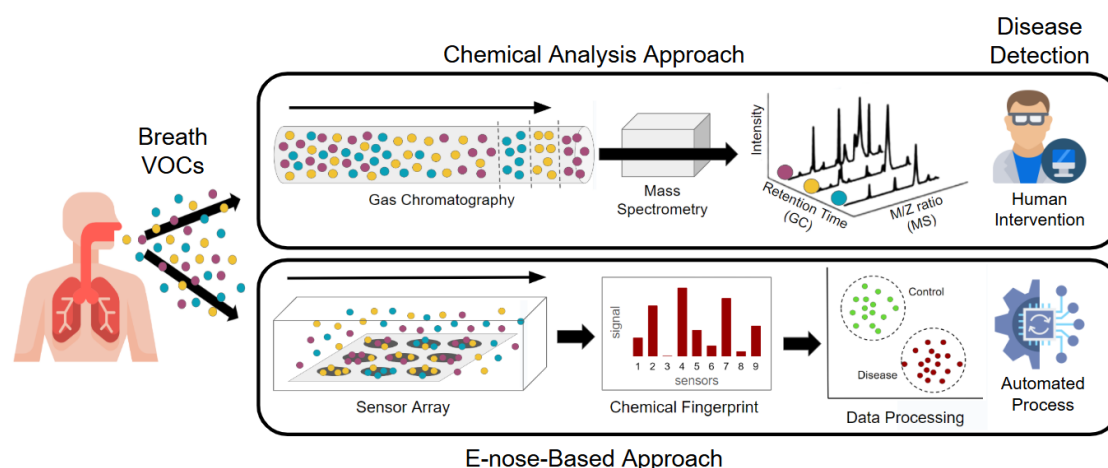

Figure 2. Illustration of the two main approaches currently used in breathomics. While fundamental research identifying individual biomarkers is conducted on complex, lab-confined instrumentation, new portable tools like e-nose focus on quick generation of breath patterns, involving the collective contribution of all breath VOCs.

While this fundamental research is crucial for advancing the field, there is growing interest in developing simpler, more accessible breath analysis tools for clinical use. The most popular candidate is the electronic nose (e-nose). They are portable, advanced devices designed to detect and identify complex mixtures of VOCs, mimicking the human sense of smell. At the heart of these devices is a sensor array that consists of multiple sensors, each engineered to respond to a broad spectrum of chemical compounds. These sensors are designed with different sensitivities, meaning they react to different aspects (chemical functions) of the VOCs present in breath. When a VOC mixture comes

into contact with the sensor array, each sensor generates a distinct electrical signal. The nature of this signal depends on the specific VOCs present, the concentration of each compound, and the interaction between the VOCs and the sensor's active material. The diversity within the sensor array is key to its functionality, as it allows the device to capture a wide variety of chemical signals. Its function mimics the human olfactory system, which comprises only about 300 sensory receptors but can differentiate over 10 000 different scents.

The most powerful aspect of a sensor array in an e-nose is its ability to generate a unique "fingerprint" or signature for each VOC mixture (Figure 2). This fingerprint is a composite of the individual sensor responses, forming a multi-dimensional data set that can be analyzed to identify a specific VOC mixture or an aroma profile. Once the sensor array generates the signal pattern, the data is processed using sophisticated pattern recognition algorithms. The e-nose must be trained with known samples to calibrate the sensor array. During training, the device learns to associate specific signal patterns with known VOC profiles. Once calibrated, the e-nose can recognize these patterns in new, unknown samples, effectively identifying different complex states and conditions.

Many preliminary research studies have already been performed using these promising devices for diagnostics purposes on diseases such as breast cancer [14] , lung cancer [15], as well as Alzheimer's and Parkinson's disease [16]. Through the use of compact tools like e-nose, portable breath analyzers could enable on-the-spot diagnosis/screening in various settings, including clinics, homes, or remote areas with limited healthcare access. Using e-nose solutions, breathomics could potentially allow patients with chronic conditions to monitor their health status regularly without the need for frequent hospital visits or invasive procedures. It could also help tailor treatments to individual patients by providing real-time information on metabolic processes and drug responses. Furthermore, because of the relatively low cost and rapid results it offers, breathomics represents a cost-effective method for large-scale health screenings and early disease detection programs.

4. Research Clinical Studies on Breathomics for Breast Cancer

About two dozen breathomics clinical studies have been conducted on women with various molecular subtypes and stages of breast cancer. The key highlights of these clinical studies have been summarized below in Table 1. While these studies show great variability in terms of how they were conducted, some consistent trends can be extrapolated from the results. These variations and trends are discussed further in the upcoming sections.

Table 1. Overview of the different studies on breathomics for breast cancer diagnosis.

1st Author	Year	Ref.	Study Population	Sensitivity	Specificity
Ebeler	1997	[10]	3 BC patients / 3 healthy controls		
Phillips	2003	[17]	51 BC patients / 50 women with abnormal mammogram but no cancer / 42 healthy controls	Model 1: 88.2% Model 2: 60.8%	Model 1: 73.8% Model 2: 82.0%
Phillips	2006	[18]	51 BC patients / 42 healthy controls	93.80%	84.60%
McCulloch	2006	[19]	31 BC patients / 55 lung cancer patients / 83 healthy controls	88%	98%
Lee	2009	[20]	17 BC patients / 24 healthy controls	75%	75%
Peng	2010	[21]	18 BC patients/ 30 lung cancer patients / 26 colon cancer patients / 18 prostate cancer / 82 healthy controls	90%	85%
Phillips	2010	[22]	54 BC patients / 204 healthy controls	75.3%	84.8%
Patterson	2011	[23]	20 BC patients / 20 healthy controls	72%	64%
Shuster	2011	[24]	13 BC patients / 16 patients with benign breast tumor / 7 healthy controls	94%	80%
Mangler	2012	[25]	10 BC patients / 10 healthy controls	80 to 100%	40 to 70%
Wang	2014	[26]	85 BC patients / 45 healthy controls		

Li	2014	[27]	22 BC patients / 17 breast benign tumors / 24 healthy controls	72.7%	91.7%
Barash	2015	[28]	169 malignant BC patients / 25 DCIS / 52 benign breast conditions / 30 controls	70 to 88%	71 to 87%
Herman-Saffar	2018	[29]	48 BC patients / 45 healthy controls	MK4: 89 to 93% Cyranose: 88 to 92%	MK4: 95 to 100% Cyranose: 78 to 85%
Phillips	2018	[30]	54 BC patients / 124 healthy controls	85%	85%
Díaz de León-Martínez	2020	[31]	262 BC patients / 181 healthy controls	100%	100%
Zhang	2020	[32]	71 BC patients / 54 gastric cancer patients / 78 healthy controls	93.59%	71.62%
Yang	2021	[33]	351 malignant BC / 88 healthy controls / 222 benign breast tumors	86%	97%
Nakayama	2022	[34]	45 BC patients / 51 healthy controls	86.3%	55.6%
Naz	2022	[35]	71 BC patients / 40 healthy controls	86.8%	75.0%
Liu	2023	[36]	465 BC patients / 4504 healthy controls	89.16%	87.70%
Zhang	2024	[37]	937 BC patients / 1044 healthy controls	85.9%	90.4%

BC: Breast Cancer; MK4 and Cyranose are two different models of e-nose devices.

4.1. Patient Selection and Study Design

From Table 1, it can be seen that the size of the participant cohorts vary substantially across the studies, from a very small size (3 BC patients vs 3 healthy controls, in the earliest study from Ebeler et al. [10]) to larger cohort sizes in later studies (Zhang et al. [37] tested 937 BC patients vs 1044 healthy controls). Overall, the average total cohort size in these studies range between 150 and 200 participants including both breast cancer patients and healthy control populations. The studies also predominantly focused on women with histologically confirmed breast cancer, ensuring robust validation of the presence of VOC biomarkers in exhaled breath. Most studies included breast cancer patients with varying disease stages, often biopsy-proven, before initiation of any treatment [22,25,26]. Healthy controls were carefully selected, often age-matched, with no history of breast cancer or related malignancies. For example, Phillips et al. [18] included 51 women with breast cancer, 50 with abnormal mammograms but no malignancy, and 42 healthy controls. The cohort of Barash et al. [28] included 169 patients with malignancy, 25 patients with ductal carcinoma in-situ (DCIS), 52 patients with benign breast conditions and 30 healthy controls, which allowed for the assessment of the diagnostic specificity of breath VOCs in distinguishing malignant from benign breast conditions, as well as their ability to differentiate breast cancer from healthy individuals.

The mean ages of participants were typically in the 50-60 range, reflecting the demographic most at risk for breast cancer. For instance, Wang et al. [26] included women aged 25-80, stratified by disease type. Age-matching was emphasized to reduce bias, as seen in Phillips et al. [18] and Shuster et al. [24]. This was done since age impacts metabolism, oxidative stress, and lung function, all of which influence the VOC generation in the organ, and subsequent exhalation in breath [13]. Older individuals may have higher levels of oxidative stress-related breath VOCs, regardless of disease status, leading to reduced specificity.

4.2. Confounding Factors and Exclusion Criteria

Exhaled breath VOCs and their origins have been extensively researched [38–40] in the recent decade where several non-disease contributing factors (or confounding factors) have been identified to affect breath volatiles. To minimize the influence of these confounding factors in a breath based clinical study, common practices for the participants include restricting exercise, smoking, and consumption of food or drink or alcohol for 2–3 hours before testing. However, the extent to which

these measures enhance reliability in breathomics remains to be established. In our tabulated list of breath studies on breast cancer, rigorous exclusion criteria were adopted to minimize the impact of confounding factors. The most common exclusion criteria include prior cancer diagnoses (except non-melanoma skin cancer) [17–19,26–28,30–32,37], acute pulmonary or systemic infections such as influenza, pneumonia or tuberculosis [26,28,30–32,36,37], and chronic respiratory diseases such as asthma and chronic obstructive pulmonary disease (COPD) [26,28,31,32]. In two cases, pregnant or breastfeeding women were excluded [26,32], and in one study those on medications whose metabolism could influence VOC levels were excluded [26]. A few of the studies also excluded participants with other active malignancies or significant systemic conditions such as diabetes [33], kidney or liver dysfunction [37], which are also known to impact the breath volatilome [41,42]. In almost all studies, breast cancer patients were required not to have undergone any treatment prior to the study in order to participate. Recent use of anesthesia [30,36] was grounds for exclusion in two studies, as they were suspected to introduce variability in breath VOC profiles. Other factors such as diet and smoking habits are known to have an impact over the breath volatiles. For example, individuals following a low carb, ketogenic diet have been demonstrated to produce higher levels of acetone on their breath, reflecting increased fat breakdown and ketone body production [43]. Smoking can introduce additional hydrocarbons, aldehydes, and other compounds into the breath [44], complicating the identification of cancer-specific biomarkers. However, only 2 studies excluded smokers from their cohorts [27,33].

While a lot of confounding factors have been identified and applied as exclusion criteria for the studies, their direct impact on the actual results is seldom explored. In a rare example, Yang et al., first tested an e-nose for the detection of breast cancer, excluding male patients, those with a history of asthma, diabetes mellitus, smoking, or having received chemotherapy in a vast study including 661 participants [33]. Under these conditions, they obtained a sensitivity of 86%, and a specificity of 97%. However, upon introduction of patients with comorbidities making up the exclusion criteria, they noted marked decreases in diagnostic accuracy. With their diagnostic odds ratio (DOR) originally near 11, the inclusion of those with diabetes pushed the DOR to 8.51. Another study conducted by Peng et al. studied specifically the impact of different confounding factors on the performances of their e-nose device, using a separate group of 59 healthy subjects carefully divided into different categories according to gender, age, ethnicity, family history of cancer, medication, diet or smoking habits [21]. Their e-nose analysis could not differentiate between individuals belonging to either of the different categories, demonstrating that none of these confounding factors affected their e-nose output. However, their e-nose could easily separate between the healthy controls and the patients with breast cancer. Furthermore, their study also included patients with lung cancer, colorectal cancer and prostate cancer, and the results showed a clear separation between each group, with the prostate cancer group being less defined [21].

Several studies opted to control their patient characteristics, though no actual consensus regarding exclusion or inclusion criteria was found. While many studies stipulated that participants could not have a history of any prior malignancy or have received cancer treatments such as chemotherapy, radiation, or surgery as those factors are likely to affect the VOC profile, there were numerous studies not excluding any comorbidities, and requiring only a clear mammogram within the past year as a measure of health for their healthy control group.

4.3. Breath Sample Collection and Environment

The volume of breath sampled varied from 0.5L to 2L, with most studies sampling 1L of alveolar breath. In nearly all of the studies, the exhaled breath was first sampled and stored either in a breath bag or in a TD tube containing a sorbent material. The stored breath was then introduced into the analytical equipment or the e-nose. Typically, these breath sampling procedures have been derived from standard research protocols when using lab-based analytical methods such as GC-MS, but point-of-care (POC) devices based on e-nose technology should ideally be able to sample the breath

directly from the patient. However, of the 6 reported studies [21,24,28,29,31,33] that used an e-nose device, only one [29] used direct breathing into the device.

Alveolar breath, constituting the end tidal volume, is the most relevant fraction of whole breath, given that it is where endogenous VOCs are found in greatest concentration [45]. The remaining, "dead space" portion of exhaled breath, comprising air from the mouth, throat, and large airways, contains compounds influenced by environmental contaminants, residual air, and non-diagnostic sources such as oral and nasal microbiota. Including this air can dilute or obscure the VOCs of interest, reducing the sensitivity and specificity of the analysis, therefore all studies except two [10,19] opted to only sample the alveolar part of the breath. However, the methodology to precisely collect the alveolar breath varies across studies. Some sampling is more manual in nature, with dead space being collected into one breath bag, and alveolar breath into a separate one [17,18,20,21,23–26,28,30,32], whereas other methods are more automated such as capnography, which uses CO_2 measurements to precisely mark the start of alveolar air and only then starts the breath sampling [33,36]. In one study [33] the team decided to perform the breath sampling while the patients were undergoing anesthesia (Sevoflurane 2%), with the Alveolar sampling being performed by the anesthesiologist by directly accessing the patient's endotracheal intubation prior to their breast tumor resection surgery. The healthy controls from this study were sampled using the exact same procedure in order to validate their results, but this sampling method is not a practical approach that could be more conveniently used in a doctor's office.

Nevertheless, exogenous VOCs are expected to appear in breath analysis, and research teams have used multiple approaches in order to mitigate their contribution. The most common is to sample the ambient air in the sample collection room to determine what constitutes "background" and what is endogenous. Another technique, though often found in tandem, is to control environmental and nutritional factors [20–22,27,28,36,45–47]. This has been done by having patients fast or refrain from certain activities (e.g. alcohol and tobacco consumption, practicing oral hygiene, etc.) for a given amount of time prior to sampling [18,20,28,31,32,36,37]. In other studies, patients have been asked to rinse their mouths with distilled water and/or breathe in clean air through filtered mouthpieces for several minutes as a form of "lung wash out" [21,22,28,36]. Though in comparing several collection methods, Di Gilio et al. found that ambient air sampling may imbue more benefit than pulmonary wash out [46]. They remark, as others have, that wash out may not actually remove environmental contaminants from exhaled breath as each VOC has specific kinetic properties that would affect its absorption, metabolism, and half-life, thereby determining the rate at which it is retained or released by the body. For example, longitudinal exposure to a pollutant could lead to certain lipophilic VOCs being taken into adipose tissue (for example, benzene-derivatives from car exhaust) They are later released into the blood, and subsequently in the breath at a slow constant rate, limiting the effectiveness of a wash out [46,48].

4.4. Breast Cancer Diagnosis: Sensitivity and Specificity

The performances obtained by each study for breast cancer diagnosis can be viewed in Table 1. Overall, sensitivity values range from 60.8% [17] to 100% [31], with most values found in the 80-85% range. Values for specificity range from 40% [25] to 100% [29,31], with most values found in the 75-80% range. As can be expected, the sensitivity and specificity values reported across the studies are heavily influenced by experimental protocols, participant selection, sample handling, analytical devices, and the statistical models employed. When it comes to breath sampling, studies that employed standardized alveolar breath collection methods typically achieved better results, as the alveolar portion represents the rich metabolic VOC biomarkers released from various organs across the body. Peng et al. [21] directly compared the performances of GC-MS against their e-nose sensor array on the same set of volunteers, with the e-nose showing both higher sensitivity (90% vs 85%) and specificity (85% vs 75%) when discerning between breast cancer patients and healthy controls.

Except for a few studies [10,19,25,27], all the other studies have used some form of machine learning (ML) based data processing to generate the sensitivity and specificity metrics. However,

except for one study [35], all the studies have not performed any additional steps in ML data processing to clearly prove that their approach does not have a statistical overfit. For example, the study from Diaz de Leon-Martinez et. al. [31] is likely suffering from an overfit in data processing. In the study from Naz et. al. [35], SVM data processing method has been used and the authors went a step further and validated the sensitivity and specificity metrics using an approach called Leave-One-Out-Cross-Validation (LOOCV) which ensures that the reported metrics does not suffer from statistical overfit.

While the goal of most studies was to establish different volatile breath profiles that differentiate breast cancer patients from healthy individuals and patients with benign breast tumors, very few tried to investigate whether breath profiles could differentiate between patients with different types of breast cancer. The only one, to our knowledge, is the one conducted by Barash et al. [28] to examine differences in breath VOCs associated with the molecular subtype of breast cancer, using both GC-MS and e-nose. Their study explores the metabolic impact of specific mutations within cancer cells resulting in breast cancer subtypes such as Luminal 1, Luminal 2, HER2+ and triple negative, which can result in altered VOC profiles. Distinct VOC patterns associated with different cancer subtypes were observed, suggesting that specific metabolic pathways influenced by genetic mutations contribute to the volatile signature. These mutations drive changes in metabolic pathways (e.g., lipid peroxidation, oxidative stress), producing unique VOC patterns reflective of the tumor's genetic background. They found that each subtype could be discerned with accuracies of 81-88% for Luminal A, 78-86% for Luminal B, and 83-90% for triple negative cancer. For each of these, the sensitivity was approximately 80%. Additionally, HER2 status was identified independent of hormone receptor status, and interestingly, when HER2 was equivocal, it was found to be HER2+ 35% of the time during the VOC analysis, much like with reported Fluorescence In-Situ Hybridization (FISH) data.

5. Advantages and Limitations of the Approach, Future Prospects and Directions

One of the most significant advantages of breathomics in breast cancer screening is its non-invasive nature. Breath analysis involves collecting exhaled air, which is far more comfortable and less risky for patients compared to traditional screening methods such as biopsies, mammograms, or blood tests. This could lead to increased patient compliance, especially in populations that are hesitant to undergo more invasive procedures. Breath analysis can potentially provide rapid results, especially when advanced technologies such as e-noses are employed (Figure 3). This makes it a convenient option for early screening, especially in resource-limited settings where access to complex imaging technology may be scarce. Breathomics also has the potential to enable the detection of breast cancer at earlier stages, through the identification of breath VOC mixture patterns associated with micro-calcifications and benign lesions, which are potential breast cancer precursors [24].

However, because breathomics is still an emerging field, its applicability to the diagnosis of breast cancer needs further validation and development, since having a high sensitivity and low false-positive result rate (high specificity) is imperative. From the analysis in section 4, it is clear that across multiple clinical studies with different breath sampling devices, the exhaled breath from breast cancer patients seems to contain a unique and complex mixture of VOCs that can be clearly separated from the control population's breath samples. However, the complex VOC mixture can have variations based on a number of factors such as the cancer subtype [28], development stage [37] and the patient's overall health [33]. Furthermore, breath VOCs can be influenced by factors like the patient's menstrual cycle [49], as well as by external factors such as environmental pollutants, diet and smoking habits [43,50]. This variability needs to be taken into account to ensure the accuracy of results and generalizability of breathomics as a screening tool for breast cancer.

On the detection technology side, the spectrometry instruments, currently being used, lack portability and affordability and require time consuming data processing for identifying biomarkers from the spectrum of generated data. For e-noses, recent breakthroughs have enabled mass production, allowing both portability and affordability. It is also observed that e-nose devices are generally better performing as a breathomics tool for breast cancer in the detection metrics in

comparison to spectrometry-based devices. However, so far, all e-nose based studies suffer from non-standard breath sampling techniques and different mechanisms to capture the alveolar breath. Given the promise of e-nose technologies to be a low cost, real-time, Point-Of-Care breathomics solution, it is imperative that they integrate a standardized direct breath sampling technique and capnography logic to eliminate any variability from external factors and exogenous VOCs. E-nose devices are also leveraging advanced AI algorithms which are trained on large amounts of data, enabling them to handle confounders and help make more effective diagnostic decisions.
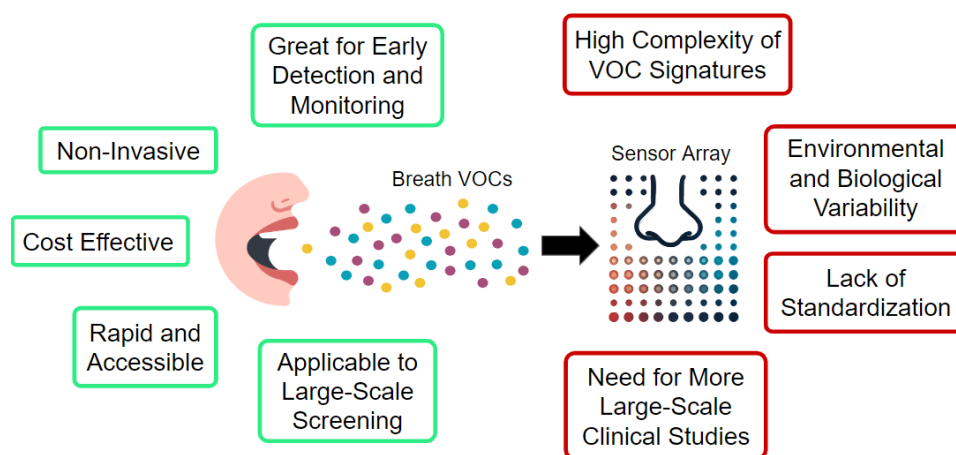

Figure 3. Schematization of the main advantages and challenges associated with the use of an e-nose for disease diagnosis using breath.

On the clinical side, large-scale clinical trials are necessary to validate the generalizability of breathomics in detecting breast cancer reliably across diverse populations, and further address the impact of confounding factors. Rigorously designed clinical studies, that take into account both the endogenous and exogenous factors that affect breath VOCs and adapt their sampling and analysis methods accordingly, can provide critical improvements in breast cancer breathomics.

6. Conclusions

Breathomics is a promising, non-invasive approach to breast cancer screening that has the potential to offer rapid, cost-effective, and accessible screening for large populations. While breathomics shows great promise, it's important to note that the field is still developing. Many studies have been conducted, but the results have great variability due to techniques used for breath sampling and analysis. Future challenges include standardizing collection and analysis methods, developing detection technologies that are both sensitive and specific, and validating results through large-scale clinical trials in order to limit the effect of confounding factors. As research progresses, breathomics has the potential to revolutionize diagnostics by providing quick, easy, and non-invasive health assessments. Future research is likely to focus on refining the detection methods and integrating breathomics with other screening tools to maximize its utility in clinical practice. There is also a need to clearly and repeatedly identify breath biomarkers associated with breast cancer. Then, a more in-depth analysis addressing the metabolic origins of the observed breath VOCs associated with breast cancer will help to fully understand the complex relationship between the disease and the breath.

Author Contributions: Conceptualization, H.Y.L., R.P., P.K., A.P.M. and S.M.; methodology, H.Y.L. and R.P.; formal analysis, H.Y.L. and R.P.; investigation, H.Y.L. and R.P.; writing—original draft preparation, H.Y.L., R.P. and A.P.M; writing—review and editing, H.Y.L., P.K., A.P.M and S.M.; supervision, S.M.; project administration, A.P.M.; All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Sung, H.; Ferlay, J.; Siegel, R.L.; Laversanne, M.; Soerjomataram, I.; Jemal, A.; Bray, F. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J. Clin.* **2021**, *71*, 209–249, doi:10.3322/caac.21660.

2. Nolan, E.; Lindeman, G.J.; Visvader, J.E. Deciphering Breast Cancer: From Biology to the Clinic. *Cell* **2023**, *186*, 1708–1728, doi:10.1016/j.cell.2023.01.040.

3. Miller, D.; Livingstone, V.; Herbison, G.P. Interventions for Relieving the Pain and Discomfort of Screening Mammography. *Cochrane Database of Systematic Reviews* **2008**, *2009*, doi:10.1002/14651858.CD002942.pub2.

4. Alexandraki, I.; Mooradian, A.D. Barriers Related to Mammography Use for Breast Cancer Screening among Minority Women. *J. Natl. Med. Assoc.* **2010**, *102*, 206–218, doi:10.1016/s0027-9684(15)30527-7.

5. Kerlikowske, K.; Ichikawa, L.; Miglioretti, D.L.; Buist, D.S.M.; Vacek, P.M.; Smith-Bindman, R.; Yankaskas, B.; Carney, P.A.; Ballard-Barbash, R.; National Institutes of Health Breast Cancer Surveillance Consortium Longitudinal Measurement of Clinical Mammographic Breast Density to Improve Estimation of Breast Cancer Risk. *J. Natl. Cancer Inst.* **2007**, *99*, 386–395, doi:10.1093/jnci/djk066.

6. Sprague, B.L.; Gangnon, R.E.; Burt, V.; Trentham-Dietz, A.; Hampton, J.M.; Wellman, R.D.; Kerlikowske, K.; Miglioretti, D.L. Prevalence of Mammographically Dense Breasts in the United States. *Journal of the National Cancer Institute* **2014**, *106*, dju255, doi:10.1093/jnci/dju255/18614661/dju255.

7. Haworth, J.J.; Pitcher, C.K.; Ferrandino, G.; Hobson, A.R.; Pappan, K.L.; Lawson, J.L.D. Breathing New Life into Clinical Testing and Diagnostics: Perspectives on Volatile Biomarkers from Breath. *Crit. Rev. Clin. Lab. Sci.* **2022**, *59*, 353–372, doi:10.1080/10408363.2022.2038075.

8. Khoubnasabjafari, M.; Mogaddam, M.R.A.; Rahimpour, E.; Soleymani, J.; Saei, A.A.; Jouyban, A. Breathomics: Review of Sample Collection and Analysis, Data Modeling and Clinical Applications. *Crit. Rev. Anal. Chem.* **2022**, *52*, 1461–1487, doi:10.1080/10408347.2021.1889961.

9. Belizário, J.E.; Faintuch, J.; Malpartida, M.G. Breath Biopsy and Discovery of Exclusive Volatile Organic Compounds for Diagnosis of Infectious Diseases. *Front. Cell. Infect. Microbiol.* **2020**, *10*, 564194, doi:10.3389/fcimb.2020.564194.

10. Ebeler, S.E. Quantitative Analysis by Gas Chromatography of Volatile Carbonyl Compounds in Expired Air from Mice and Human. *Journal of Chromatography B* **1997**, *702*, 211–215.

11. Phillips, M.; Cataneo, R.N.; Cummin, A.R.C.; Gagliardi, A.J.; Gleeson, K.; Greenberg, J.; Maxfield, R.A.; Rom, W.N. Detection of Lung Cancer with Volatile Markers in the Breath. *Chest* **2003**, *123*, 2115–2123, doi:10.1378/chest.123.6.2115.

12. Phillips, M.; Cataneo, R.N.; Condos, R.; Ring Erickson, G.A.; Greenberg, J.; La Bombardi, V.; Munawar, M.I.; Tietje, O. Volatile Biomarkers of Pulmonary Tuberculosis in the Breath. *Tuberculosis* **2007**, *87*, 44–52, doi:10.1016/j.tube.2006.03.004.

13. Phillips, M.; Cataneo, R.N.; Greenberg, J.; Gunawardena, R.; Naidu, A.; Rahbari-Oskoui, F. Effect of Age on the Breath Methylated Alkane Contour, a Display of Apparent New Markers of Oxidative Stress. *J. Lab. Clin. Med.* **2000**, *136*, 243–249, doi:10.1067/mlc.2000.108943.

14. Barash, O.; Haick, H. Exhaled Volatile Organic Compounds as Noninvasive Markers in Breast Cancer. In *Omics Approaches in Breast Cancer*; Springer India: New Delhi, 2014; pp. 461–481 ISBN 9788132208426.

15. Peng, G.; Tisch, U.; Adams, O.; Hakim, M.; Shehada, N.; Broza, Y.Y.; Billan, S.; Abdah-Bortnyak, R.; Kuten, A.; Haick, H. Diagnosing Lung Cancer in Exhaled Breath Using Gold Nanoparticles. *Nat. Nanotechnol.* **2009**, *4*, 669–673, doi:10.1038/nnano.2009.235.

16. Tisch, U.; Schlesinger, I.; Ionescu, R.; Nassar, M.; Axelrod, N.; Robertman, D.; Tessler, Y.; Azar, F.; Marmur, A.; Aharon-Peretz, J.; et al. Detection of Alzheimer's and Parkinson's Disease from Exhaled Breath Using Nanomaterial-Based Sensors. *Nanomedicine* **2013**, *8*, 43–56, doi:10.2217/nnm.12.105.

17. Phillips, M.; Cataneo, R.N.; Ditkoff, B.A.; Fisher, P.; Greenberg, J.; Gunawardena, R.; Kwon, C.S.; Rahbari-Oskoui, F.; Wong, C. Volatile Markers of Breast Cancer in the Breath. *Breast J.* **2003**, *9*, 184–191, doi:10.1046/j.1524-4741.2003.09309.x.

18. Phillips, M.; Cataneo, R.N.; Ditkoff, B.A.; Fisher, P.; Greenberg, J.; Gunawardena, R.; Kwon, C.S.; Tietje, O.; Wong, C. Prediction of Breast Cancer Using Volatile Biomarkers in the Breath. *Breast Cancer Res. Treat.* **2006**, *99*, 19–21, doi:10.1007/s10549-006-9176-1.

19. McCulloch, M.; Jezierski, T.; Broffman, M.; Hubbard, A.; Turner, K.; Janecki, T. Diagnostic Accuracy of Canine Scent Detection in Early- and Late-Stage Lung and Breast Cancers. *Integr. Cancer Ther.* **2006**, *5*, 30–39, doi:10.1177/1534735405285096.

20. Lee, K.S.; Abouelnasr, M.F.; Bayer, C.W.; Gabram, S.G.; Mizaikoff, B.; Rogatko, A.; Vidakovic, B. Mining Exhaled Volatile Organic Compounds for Breast Cancer Detection. *Adv Appl Stat Sci* **2009**, *1*, 327–342.

21. 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, doi:10.1038/sj.bjc.6605810.

22. Phillips, M.; Cataneo, R.N.; Saunders, C.; Hope, P.; Schmitt, P.; Wai, J. Volatile Biomarkers in the Breath of Women with Breast Cancer. *J. Breath Res.* **2010**, *4*, 026003, doi:10.1088/1752-7155/4/2/026003.

23. Patterson, S.G.; Bayer, C.W.; Hendry, R.J.; Sellers, N.; Lee, K.S.; Vidakovic, B.; Mizaikoff, B.; Gabram-Mendola, S.G.A. Breath Analysis by Mass Spectrometry: A New Tool for Breast Cancer Detection? *Am. Surg.* **2011**, *77*, 747–751, doi:10.1177/000313481107700632.

24. Shuster, G.; Gallimidi, Z.; Reiss, A.H.; Dovgolevsky, E.; Billan, S.; Abdah-Bortnyak, R.; Kuten, A.; Engel, A.; Shiban, A.; Tisch, U.; et al. Classification of Breast Cancer Precursors through Exhaled Breath. *Breast Cancer Res. Treat.* **2011**, *126*, 791–796, doi:10.1007/s10549-010-1317-x.

25. Mangler, M.; Freitag, C.; Lanowska, M.; Staeck, O.; Schneider, A.; Speiser, D. Volatile Organic Compounds (VOCs) in Exhaled Breath of Patients with Breast Cancer in a Clinical Setting. *Ginekol. Pol.* **2012**, *83*, 730–736.

26. Wang, C.; Sun, B.; Guo, L.; Wang, X.; Ke, C.; Liu, S.; Zhao, W.; Luo, S.; Guo, Z.; Zhang, Y.; et al. Volatile Organic Metabolites Identify Patients with Breast Cancer, Cyclomastopathy and Mammary Gland Fibroma. *Sci. Rep.* **2014**, *4*, 1–6, doi:10.1038/srep05383.

27. Li, J.; Peng, Y.; Liu, Y.; Li, W.; Jin, Y.; Tang, Z.; Duan, Y. Investigation of Potential Breath Biomarkers for the Early Diagnosis of Breast Cancer Using Gas Chromatography-Mass Spectrometry. *Clin. Chim. Acta* **2014**, *436*, 59–67, doi:10.1016/j.cca.2014.04.030.

28. Barash, O.; Zhang, W.; Halpern, J.M.; Hua, Q.-L.; Pan, Y.-Y.; Kayal, H.; Khoury, K.; Liu, H.; Davies, M.P.A.; Haick, H. Differentiation between Genetic Mutations of Breast Cancer by Breath Volatolomics. *Oncotarget* **2015**, *6*, 44864–44876, doi:10.18632/oncotarget.6269.

29. Herman-Saffar, O.; Boger, Z.; Libson, S.; Lieberman, D.; Gonen, R.; Zeiri, Y. Early Non-Invasive Detection of Breast Cancer Using Exhaled Breath and Urine Analysis. *Comput. Biol. Med.* **2018**, *96*, 227–232, doi:10.1016/j.compbiomed.2018.04.002.

30. Phillips, M.; Cataneo, R.N.; Cruz-Ramos, J.A.; Huston, J.; Ornelas, O.; Pappas, N.; Pathak, S. Prediction of Breast Cancer Risk with Volatile Biomarkers in Breath. *Breast Cancer Res. Treat.* **2018**, *170*, 343–350, doi:10.1007/s10549-018-4764-4.

31. Díaz de León-Martínez, L.; Rodríguez-Aguilar, M.; Gorocica-Rosete, P.; Domínguez-Reyes, C.A.; Martínez-Bustos, V.; Tenorio-Torres, J.A.; Ornelas-Rebolledo, O.; Cruz-Ramos, J.A.; Balderas-Segura, B.; Flores-Ramírez, R. Identification of Profiles of Volatile Organic Compounds in Exhaled Breath by Means of an Electronic Nose as a Proposal for a Screening Method for Breast Cancer: A Case-Control Study. *J. Breath Res.* **2020**, *14*, 046009, doi:10.1088/1752-7163/aba83f.

32. Zhang, Y.; Guo, L.; Qiu, Z.; Lv, Y.; Chen, G.; Li, E. Early Diagnosis of Breast Cancer from Exhaled Breath by Gas Chromatography-Mass Spectrometry (GC/MS) Analysis: A Prospective Cohort Study. *J. Clin. Lab. Anal.* **2020**, *34*, e23526, doi:10.1002/jcla.23526.

33. Yang, H.-Y.; Wang, Y.-C.; Peng, H.-Y.; Huang, C.-H. Breath Biopsy of Breast Cancer Using Sensor Array Signals and Machine Learning Analysis. *Sci. Rep.* **2021**, *11*, 103, doi:10.1038/s41598-020-80570-0.

34. Nakayama, Y.; Hanada, M.; Koda, H.; Sugimoto, M.; Takada, M.; Toi, M. Breast Cancer Detection Using Volatile Compound Profiles in Exhaled Breath via Selected Ion-Flow Tube Mass Spectrometry. *J. Breath Res.* **2022**, *17*, 016006, doi:10.1088/1752-7163/aca696.

35. Naz, F.; Groom, A.G.; Mohiuddin, M.D.; Sengupta, A.; Daigle-Maloney, T.; Burnell, M.J.; Michael, J.C.R.; Graham, S.; Beydaghyan, G.; Scheme, E.; et al. Using Infrared Spectroscopy to Analyze Breath of Patients Diagnosed with Breast Cancer. *J. Clin. Oncol.* **2022**, *40*, e13579–e13579, doi:10.1200/jco.2022.40.16_suppl.e13579.

36. Liu, J.; Chen, H.; Li, Y.; Fang, Y.; Guo, Y.; Li, S.; Xu, J.; Jia, Z.; Zou, J.; Liu, G.; et al. A Novel Non-Invasive Exhaled Breath Biopsy for the Diagnosis and Screening of Breast Cancer. *J. Hematol. Oncol.* **2023**, *16*, 63, doi:10.1186/s13045-023-01459-9.

37. Zhang, J.; He, X.; Guo, X.; Wang, J.; Gong, X.; Jiao, D.; Chen, H.; Liu, Z. Identification Potential Biomarkers for Diagnosis, and Progress of Breast Cancer by Using High-Pressure Photon Ionization Time-of-Flight Mass Spectrometry. *Anal. Chim. Acta* **2024**, *1320*, 342883, doi:10.1016/j.aca.2024.342883.

38. Beauchamp, J.; Davis, C.; Pleil, J. Breathborne Biomarkers and the Human Volatilome. **2020**.

39. Kuo, T.-C.; Tan, C.-E.; Wang, S.-Y.; Lin, O.A.; Su, B.-H.; Hsu, M.-T.; Lin, J.; Cheng, Y.-Y.; Chen, C.-S.; Yang, Y.-C.; et al. Human Breathomics Database. *Database (Oxford)* **2020**, *2020*, doi:10.1093/database/baz139.

40. Kuo, P.-H.; Jhong, Y.-C.; Kuo, T.-C.; Hsu, Y.-T.; Kuo, C.-H.; Tseng, Y.J. A Clinical Breathomics Dataset. *Sci. Data* **2024**, *11*, 203, doi:10.1038/s41597-024-03052-2.

41. Van den Velde, S.; Nevens, F.; Van Hee, P.; van Steenberghe, D.; Quirynen, M. GC-MS Analysis of Breath Odor Compounds in Liver Patients. *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.* **2008**, *875*, 344–348, doi:10.1016/j.jchromb.2008.08.031.

42. Saidi, T.; Zaim, O.; Moufid, M.; El Bari, N.; Ionescu, R.; Bouchikhi, B. Exhaled Breath Analysis Using Electronic Nose and Gas Chromatography–mass Spectrometry for Non-Invasive Diagnosis of Chronic Kidney Disease, Diabetes Mellitus and Healthy Subjects. *Sens. Actuators B Chem.* **2018**, *257*, 178–188, doi:10.1016/j.snb.2017.10.178.

43. Ajibola, O.A.; Smith, D.; Spaněl, P.; Ferns, G.A.A. Effects of Dietary Nutrients on Volatile Breath Metabolites. *J. Nutr. Sci.* **2013**, *2*, e34, doi:10.1017/jns.2013.26.

44. Buszewski, B.; Ulanowska, A.; Ligor, T.; Denderz, N.; Amann, A. Analysis of Exhaled Breath from Smokers, Passive Smokers and Non-Smokers by Solid-Phase Microextraction Gas Chromatography/mass Spectrometry. *Biomed. Chromatogr.* **2009**, *23*, 551–556, doi:10.1002/bmc.1141.

45. Lourenço, C.; Turner, C. Breath Analysis in Disease Diagnosis: Methodological Considerations and Applications. *Metabolites* **2014**, *4*, 465–498, doi:10.3390/metabo4020465.

46. 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*, doi:10.3390/molecules25245823.

47. Walsh, C.M.; Fadel, M.G.; Jamel, S.H.; Hanna, G.B. Breath Testing in the Surgical Setting: Applications, Challenges, and Future Perspectives. *Eur. Surg. Res.* **2023**, *64*, 315–322, doi:10.1159/000531504.

48. Haick, H.; Broza, Y.Y.; Mochalski, P.; Ruzsanyi, V.; Amann, A. Assessment, Origin, and Implementation of Breath Volatile Cancer Markers. *Chem. Soc. Rev.* **2014**, *43*, 1423–1449, doi:10.1039/c3cs60329f.

49. Sukul, P.; Schubert, J.K.; Trefz, P.; Miekisch, W. Natural Menstrual Rhythm and Oral Contraception Diversely Affect Exhaled Breath Compositions. *Sci. Rep.* **2018**, *8*, 10838, doi:10.1038/s41598-018-29221-z.

50. Krilaviciute, A.; Leja, M.; Kopp-Schneider, A.; Barash, O.; Khatib, S.; Amal, H.; Broza, Y.Y.; Polaka, I.; Parshutin, S.; Rudule, A.; et al. Associations of Diet and Lifestyle Factors with Common Volatile Organic Compounds in Exhaled Breath of Average-Risk Individuals. *J. Breath Res.* **2019**, *13*, 026006, doi:10.1088/1752-7163/aaf3dc.