Taking your breath away: metabolomics breathes life into personalized medicine

Nicholas J.W. Rattray1*, Zahra Hamrang2*, Drupad K. Trivedi1, Royston Goodacre1, and Stephen J. Fowler3,4

1 School of Chemistry and Manchester Institute of Biotechnology, University of Manchester, Manchester, UK
2 Manchester Pharmacy School, University of Manchester, Manchester, UK
3 University of Manchester, Manchester Academic Health Science Centre, NIHR Respiratory and Allergy Clinical Research Facility, University Hospital of South Manchester, Manchester, UK
4 Respiratory Medicine, Lancashire Teaching Hospitals NHS Foundation Trust, Preston, UK

Breath-based metabolomics (breathomics) is an exciting developing area of biotechnology that centers on the capture, identification, and quantification of volatile organic compound (VOC) patterns in human breath and their utilization as tools in the diagnosis of a broad spectrum of medical problems. With the age of personalized medicines demanding rapid bespoke diagnosis and treatment, this area of molecular diagnostics is beginning to see an upsurge in biotechnological advancement. Here, we discuss recent improvements and directions in the development of breath VOC analysis and diagnosis platforms that offer the potential for disease biomarker discovery and disease prognosis.

VOCs as an unmapped source of biomarker potential within the breath
Breathing is something we all do approximately 20,000 times a day and is a medium that contains a wealth of potential for non-invasive disease detection. Currently, components of breath studied include breath condensate [1], exhaled breath particles [2], and in some instances, combined with carbon isotope-labeled substrates to investigate drug metabolism [3]. Given the broad nature of this field, we highlight current trends and the future direction of the potential of gas-based VOCs (see Glossary) as a source of disease diagnosis.

The modern era of breath gas analysis started in 1971 when Linus Pauling’s pioneering paper [4] to the National Academy of Sciences communicated that a typical human breath signature comprised over 250 spectral features containing information potentially originating from VOCs within the breath. These compounds are the products of numerous highly dynamic and regulated metabolic processes not only locally in the lung, but also throughout the body, whereby VOCs generated peripherally are transported to the breath via the pulmonary circulation and subsequent alveolar blood–gas exchange.

The relatively slow pace of breath biomarker research since is in part due to technological limitations associated with the complexity of reliably capturing breath, the analytical intricacy of extrapolating potential biomarkers from endogenous signals, and the inability of the chemical detectors used to achieve sufficient sensitivity and analyze specificity for the many low concentration compounds.

With current human biomedical monitoring focusing on tissue samples and bodily fluids as a source of diagnosis [5], breath-based biomarkers remain among the least developed despite having huge potential as a non-invasive diagnostic source.

Over the past few years, 3400 individual VOCs have been detected from deep alveolar breath, and breath diagnostics has garnered increasing attention from the media (www.bbc.co.uk/news/science-environment-22013700) and governing bodies [6]. The renaissance in breath diagnostics is fuelled by advancements in adaptive sampling methodology and an explosion in the diversity, versatility, and sensitivity of associated detection platforms. It is important that this upward trend in biotechnological advancement is continued in order to develop non-invasive real-time disease diagnostics that will be available in the clinic and at the bedside, at the point of care (PoC) [7,8].

The breath analysis pipeline
The breathomics analysis pipeline (Figure 1) is comparable to other metabolomics-based procedures for other substrates, such as biofluids and tissues [9,10]. A sample is collected in a standardized manner and chemically profiled by a suitable platform to gain the largest representative (in terms of reproducibility and robustness) metabolomic signature [11]. This can be performed on larger systems, such as gas chromatography–mass spectrometry (GC-MS). Subsequent data preprocessing and statistical or chemometric analysis of the platform data are carried out to identify significant metabolic panels that can be linked to

Corresponding authors: Rattray, N.J.W. (Nicholas.Rattray@manchester.ac.uk; nikrattray@gmail.com; Fowler, S.J. (Stephen.Fowler@manchester.ac.uk).
Keywords: breathomics; volatile organic compounds; personalized medicine; point-of-care; eNose; mass spectrometry.
*These authors contributed equally to this work.
0167-7999/ © 2014 Elsevier Ltd. All rights reserved. http://dx.doi.org/10.1016/j.tibtech.2014.08.003
the biology of the disease in question. The final step is the advancement of targeted, miniaturized systems that can be used in a clinical setting by healthcare professionals who are not necessarily experts in bioanalysis.

Breath analysis is encountering several distinct challenges in the attempts to miniaturize inherently complex technology. The main areas of sample collection and analytical processing via MS analysis have been the foci of improvement to date. The cycle of development starts with the collection of a pilot sample set (preferably in duplicate).

Duplicate sampling allows for tandem analysis. A large MS platform can positively identify VOCs, concurrently permitting external validation and signal correlation of the ‘breathprint’ created by a smaller sensor-based device. The smaller devices have more prospect of acceptance at PoC within a clinical setting. Subsequent chemometric analysis, which must be appropriately validated [12,13] and reported so that others can reproduce this process [14], is used to classify potential patterns that can be linked to specific disease pathology and any potential biomarker relations.

Current VOC-capturing technologies
Measurement of exhaled-breath VOCs requires analytical methodologies that reproducibly capture analytes of interest while minimizing interference from the sample matrix. For metabolomic profiling, the sampling process should also be as chemically unselective as possible. Breath-capture methods range from directly breathing into an analysis platform [15] and the relatively simplistic collection within plastic Tedlar® and aluminumized Mylar bags or BioVOC™ bottles, to more complex systems that try to maximize sample quality and minimize the influence of the main challenges in breath analysis. Effective capture of a sample requires minimizing interference from exogenous environmental VOCs [16] or VOCs that do not originate from the area of interest, such as the lower respiratory tract or systemic compartment. VOCs from the upper respiratory tract and mouth are the targets of interest. Using a buffered-end tidal system [17] is one way to achieve efficient breath capture, along with using the breath collection unit developed by Ionicon. By contrast, uncontrolled end exhaled-breath sampling does not reliably and reproducibly measure concentrations of alveolar VOCs, and a steady breathing pattern should be established before analysis [18]. The breath-sampling apparatus must be acceptable to target patient groups (which may include those in respiratory distress or under ventilation) and to the operator. The apparatus must be safe and should comply with appropriate infection control requirements for use in the clinical environment. Current VOC-trapping technologies have their relative merits and drawbacks (Table 1).

Once the breath has been sampled, it needs to be introduced into the analytical platform. In the case of GC-MS, this is commonly via thermal desorption (TD) [19], solid-phase microextraction (SPME) [20], or porous monotraps (http://www.hichrom.com/product_range/existing_products/GLS/Monotrap.htm). These methods fall under the umbrella of active absorbent-based trapping and involve drawing the breath sample over a fiber or through a tube containing a combination of adsorbent polymers or activated charcoal. The adsorbent materials are heated to release trapped VOCs, which are then applied to a GC column. These methods benefit from being inherently designed to preconcentrate VOCs, can be tailored to target specific groups of VOCs, and initial sampling can be done by portable equipment before analysis via MS.

Current VOC identification platforms
Currently, the main analytical platforms for biomarker discovery and identification use time-of-flight (ToF) MS,
which provides accurate determination of mass at high resolution, therefore aiding the identification of unknown species. The resolving capabilities of GC have been coupled to this form of MS [21]. GC-MS remains the gold standard hybrid-analytical platform used and applied successfully in breath VOC biomarker identification over a diverse range of diseases [22,23]. However, GC-MS does not offer real-time analysis, has a large laboratory footprint (negrating portability), and suffers from the need for constant self-calibration for specific analyte as well as extended sample processing times.

In recent years, other hybrid MS techniques, such as selected ion flow tube MS (SIFT-MS), proton transfer reaction MS (PTR-MS) and switching reagent ion MS (SRI-MS) [24], have offered an alternative approach. These methods can require chemical ionization from larger reagent ions, such as H3O+, NO+, and O2+, but have been shown to produce real-time quantification of VOC profiles in healthy volunteers [25] along with biomarkers entrapped on human skin [26] and linked to diseases such as gastric cancer [27], lung cancer [28], and renal failure [29]. PTR-MS has also been used in the real-time evaluation of breathing patterns and how they influence VOC output [30,31]

Another technology that has seen advancements made with breath analysis is differential mobility spectrometry (DMS), which works through the identification of ions via their mobility in a carrier gas. DMS can be hybridized with ToF-MS as an ion filter, or combined with GC as a full detector. Importantly, these machines are small and have limited resource requirements. Once cross-validated with a full MS platform, they can potentially act as standalone sensors for used at the PoC. Analytical devices continue to be developed and are becoming more portable (Table 2).

**Statistical challenges in breathomics**

Each analytical platform used in breathomics has its own statistical challenges and limitations. However, aligning these outputs to search for a common map that captures and overlays different data types is essential. Tools are being developed to overcome the challenges in validating VOC data (Box 1). Raw data collected using eNose systems comprise variation in electrical resistance that can be coupled with off-line pattern-recognition software for analysis [32]. Data acquired from hybrid MS systems require many data treatments, such as deconvolution, baseline correction, interpolation, alignment, normalization, and scaling [33]. Various packages written for statistical software, such as Matlab, R, and Python, can perform these functions, along with splining, dynamic time warping, parametric time warping, or correlation optimized warping (COW) as well as Java-based applications, such as MZMine, MZMatch, or ProteoWizard [34]. Once the data have been preprocessed, a deterministic approach to the application of univariate analysis, such as one-way analysis of variance (ANOVA), Kruskal–Wallis, or an unpaired t-test, can used for comparing groups, with adjustments for multiple testing to minimize false discovery rates. At the same time, cross-validation techniques, such as bootstrapping, can be used to test the strength of the statistical test. Combination of this univariate strategy alongside multivariate analysis, such as principal component analysis (PCA) or genetic algorithms, gives a more comprehensive view of a data set.

Once data have been interpreted, a consensus has to be reached on the quality of experimental output. Accepted guidelines for this can be found within the Metabolomics Standards Initiative (MSI) [35], a community-driven collective that aims to achieve a global consensus for minimum reporting standards and allow for integration of data from independent laboratories. Curation of these data is also a vital, yet often overlooked aspect to metabolomics experiments. Adding final metadata sets to open-access, online repositories, such as MetaboLights [36], allows for greater experimental transparency, quicker knowledge transfer, and better design of future experiments.
<table>
<thead>
<tr>
<th>Sampling technology</th>
<th>Advantages</th>
<th>Disadvantages</th>
<th>Compatible technology</th>
<th>Refs</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Tedlar/Mylar Bag</strong></td>
<td>Cheap and inert</td>
<td>Susceptible to leaks and fragile</td>
<td>Can interface with either SPME or TD tubes for GC-MS analysis, or connected direct to ‘online’ analyzers, such as sensor or IMS-based systems</td>
<td>[57,58]</td>
</tr>
<tr>
<td></td>
<td>Reusable</td>
<td>Can suffer from UV degradation (resulting in inflated levels of hydrocarbons)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Water can condense within the bag and interfere with downstream GC-MS analysis</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Can connect to standard clinical respiratory equipment</td>
<td>Care must be taken with cleaning and storage to avoid contamination</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Complications in bag sampling, storage, and handling</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Flow Reactor</strong></td>
<td>Good sample reproducibility due to precise sample volumes being examined every time</td>
<td>Currently expensive</td>
<td>Can have thermal desorption tubes attached to sampling port or can be attached to PTR-MS or a similar analyzer</td>
<td>[24,59]</td>
</tr>
<tr>
<td></td>
<td>Inert</td>
<td>Requires a constant inert gas flow</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Constant flow minimizes the risk of water condensation</td>
<td>Not user friendly</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>High cost</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Storage of several samples is impractical</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Bio-VOC™</strong></td>
<td>Cheap and inert</td>
<td>Can only trap 150 ml of end-tidal breath and cannot preconcentrate</td>
<td>Can apply breath on to standard TD-tubes alongside SPME filters for further analysis</td>
<td>[60,61]</td>
</tr>
<tr>
<td></td>
<td>Bypasses upper respiratory/oral contamination</td>
<td>Proportion of ‘late expiratory’ breath sampled will vary by subjects’ lung volume</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>User friendly</td>
<td>Does not allow for breathing equilibrium to be reached</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Internal standard can be added to the device</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>BCA</strong></td>
<td>Preferentially samples the end-tidal breath but has two separate traps that can also pick up the front-end breath</td>
<td>Potentially expensive outlay</td>
<td>Compatible with GC-MS, GC-SAW, and GC-FID</td>
<td>[22]</td>
</tr>
<tr>
<td></td>
<td>Can link up to Breathlink™ for access to a collated database of previous breath profiles</td>
<td>Large in size</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>User friendly</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Digital Flow Control PoC ready</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 2. Current and developing analytical platforms for detecting breath VOCs

<table>
<thead>
<tr>
<th>Technique</th>
<th>Detection limit</th>
<th>Advantages</th>
<th>Disadvantages</th>
<th>Potential for point-of-care use?</th>
<th>Refs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spectrometric based</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GC-MS</td>
<td>ppb</td>
<td>Current gold standard: can identify unknowns; quantitative; automated</td>
<td>Expensive; time consuming; not currently portable; sensitivity not improved by preconcentration; requires dry samples</td>
<td>No</td>
<td>[37,38,62]*</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Highly sensitive; can preconcentrate samples to detect lower levels, automated</td>
<td>Complicated data deconvolution and compound identification processes</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>VOCs can be captured on different absorbent beds, such as SPME, TD, and Monotrap</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SIFT-MS</td>
<td>ppb</td>
<td>Real-time analysis; can achieve absolute quantification</td>
<td>Expensive; not ideal for broad profiling</td>
<td>No</td>
<td>[25,27]</td>
</tr>
<tr>
<td>DMS</td>
<td>ppt</td>
<td>Robust, compact, sensitive</td>
<td>Confident identification needs to be carried out on a MS system</td>
<td>Yes</td>
<td>[63]</td>
</tr>
<tr>
<td>PTR-MS/ PTR-ToFMS</td>
<td>ppt</td>
<td>Has high specificity and can detect very low mass compounds</td>
<td>Cross-signal interference; expensive</td>
<td>Yes</td>
<td>[18,59]</td>
</tr>
<tr>
<td>ESI-MS</td>
<td>ppb</td>
<td>Minimal need for adaptive sampling technology, rapid</td>
<td>Requires subject to be beside analytical platform for analysis; relatively expensive</td>
<td>No</td>
<td>[64,65]</td>
</tr>
<tr>
<td>FAIMS</td>
<td>ppb</td>
<td>Can be miniaturized; (+)ve and (−)ve ions can be detected simultaneously (Owlstone)</td>
<td>Requires preprogramming; not applicable to unknown compounds; reduced sensitivity in complex matrices; can suffer from signal suppression</td>
<td>Yes</td>
<td>[66]</td>
</tr>
<tr>
<td>Sensor based</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>eNose</td>
<td>ppt</td>
<td>Clinical PoC; data available in real time; ease of use; programmable; handheld</td>
<td>Requires preprogramming; calibration and signal needs to be compared with MS signal; database of disease signals needs to be created (Cyanose)</td>
<td>Yes</td>
<td>[67]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Different sensor design, such as quartz microbalance and conducting polymers (Cyanose), allows for large range of compound coverage</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gold Nano-Biosensor</td>
<td>ppt</td>
<td>Rapid; no need for preconcentration; highly sensitive; disease specific</td>
<td>In development: requires significant research for PoC</td>
<td>Potentially</td>
<td>[68]</td>
</tr>
<tr>
<td>Surface Plasmon Resonance</td>
<td>ppt</td>
<td>Highly selective; high throughput</td>
<td>Selective recognition needs to be preprogrammed on an appropriate chip surface (aqueous media)</td>
<td>Potentially</td>
<td>[69]</td>
</tr>
<tr>
<td>Piezoelectric Cantilever</td>
<td>ppt</td>
<td>Can be specifically tailored to individual compounds, not just classes</td>
<td>Possible issues with poisoning of binding ligands</td>
<td>Yes</td>
<td>[70]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>As lithographic techniques improve, more sensors can be applied to smaller chips</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Clinical applications of breathomics

Breath analysis offers the potential for biomarker discovery in an almost unlimited variety of clinical circumstances, ranging from disease diagnosis to stratification to treatment monitoring or prognosis (Table 3). Likewise, the breadth of disease groups would include airway and lung diseases, and even distal single-organ or systemic and multiorgan diseases that transmit VOC byproducts into the blood stream. Challenges that must always be addressed and accounted for in breath sampling for clinical studies include environmental contamination, patient comfort and safety, and infection control. Study-specific issues relate to the desire to sample preferentially the portion of the breath that arises from the area of interest (e.g., the mouth, airways, or alveoli). The study of volatiles arising in the mouth principally relates to the study of halitosis and is reviewed elsewhere [37].

The aspect of breath analysis that relates to metabolomics is unlikely to reveal any single, unique biomarkers pertaining to particular diseases, organisms, or process. Specific combinations or classes of compounds (i.e., fingerprints) are more likely to form the basis of a ‘compound biomarker panel’ for a disease, as established from previous work in obstructive airway diseases. For example, in chronic obstructive pulmonary disease (COPD), a model VOC panel comprising 11 volatile compounds discriminated COPD from healthy controls, of which nine VOCs were aldehydes [19]. Similarly, the model that could

can indicate biomarkers linked to disease states and treatment progression.

Within this pipeline are many pitfalls and potential bottlenecks. Given that best methods for preprocessing cannot be made empirically, and due to the so-called ‘no free lunch theorem’ (where each chemometrics solution is different and one size does not fit all), an approach to each experiment has to be established deterministically. This includes how missing data are dealt with and what imputation methods are used; in addition, the optimal way to scale and normalize data also needs to be established, before the data are even analyzed by the plethora of possible univariate and multivariate methods.

Figure I. The three main stages for breathomics data treatment and analysis. Abbreviations: BMI, body mass index; TIC, total ion chromatogram.

differentiate asthmatic from healthy breath in a parallel study contained multiple methylated alkanes [38]. The latter classes of compound have received interest elsewhere as potential markers for lung cancer [22] and diabetes [23].

It may often be difficult to provide a biological explanation for the correlation between VOCs and pathology in particular disease states. However, such explanations would provide a degree of biological validity to any early discoveries, as would the discovery of similar classes of
volatile compound in different studies investigating related diseases. Methylated alkanes, such as 4-methyloctane and 3-methyldecane, arise in states of oxidative stress [23, 39]. Independent discovery of these methylated alkanes as discriminating compounds for different processes that involve oxidative stress provides reassurance that they are likely to prove useful. Although the nonspecificity of methylated alkanes may mean that they cannot be used for diagnosis, monitoring their levels within an individual over time might be useful for inferring how ‘oxidatively stressed’ the individual may be at any point. Another disease area that may be ideal for breath volatile monitoring is infection. The altered exhaled volatile profile of infected individuals relates theoretically both to the metabolism of the infecting organisms, and the host inflammatory response to that infection. This has been tested actively in patients with cystic fibrosis [40] and pulmonary TB [41].

One area in which there is the potential for specific single biomarker utility is drug metabolite detection. Such an application has already been investigated for the monitoring and dose titration of anesthetic gases on intensive care units [42]. Another ingenious approach to treatment monitoring is the detection of exhaled $^{13}$C following ingestion of an isotopically labeled drug [3]. The Helicobacter pylori test is one such example that measures the kinetics of labeled CO$_2$ from a pulse of $^{13}$C urea [43]. These methods have the advantage of reflecting the metabolism of a specific drug in a patient and could also aid in identifying patients who are either at risk of drug toxicity or treatment failure [44].

### Future advancements towards PoC detectors

Assays that minimize the number invasive diagnostic interventions (i.e., histological examination following invasive biopsy) are desirable. Elucidation of the breathome for specific diseases in large patient groups may one day become a direct and immediate solution in diagnostics and ultimately reduce the morbidity and costs associated with routine invasive procedures. In this section, we discuss more recent technologies supporting PoC applications.

---

**Table 3. Selected examples of clinical studies of breath volatile analysis applied to human disease**

<table>
<thead>
<tr>
<th>Sample collection</th>
<th>Details</th>
<th>Analytical method</th>
<th>Disease</th>
<th>VOCs under investigation</th>
<th>Refs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bags</td>
<td>Tedlar bags; whole breath pulled into adsorbent tubes</td>
<td>TD/GC-ToFMS</td>
<td>Asthma/COPD</td>
<td>Untargeted</td>
<td>[58]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TD/GC-MS</td>
<td>Liver disease</td>
<td>Ethane, acetone, isoprene, alcohol, sulfur compounds</td>
<td>[71]</td>
</tr>
<tr>
<td></td>
<td>Tedlar bags; whole breath</td>
<td>PTR-MS versus</td>
<td>Lung cancer</td>
<td>Untargeted</td>
<td>[72]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SPME/GC-MS</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Tedlar$^*$ bags versus BioVOC$^TM$</td>
<td>SPME/GC-MS</td>
<td>Asthma</td>
<td>Untargeted</td>
<td>[74]</td>
</tr>
<tr>
<td></td>
<td>Nalophan bags; whole breath</td>
<td>SIFT-MS</td>
<td>Gastrooesophageal cancer</td>
<td>Alcohols, phenols, pentanoic acid, hexanoic acid, hydrogen sulfide, hydrogen cyanide, acetaldehyde, formaldehyde, acetone, acetic acid, isoprene, ammonia</td>
<td>[75]</td>
</tr>
<tr>
<td></td>
<td>Tedlar bags; washout; filtered air; whole breath</td>
<td>eNose (Cyranose)</td>
<td>Asthma</td>
<td>Unknown (sensor-based)</td>
<td>[76]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>COPD</td>
<td></td>
<td>[77]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Lung cancer</td>
<td></td>
<td>[78]</td>
</tr>
<tr>
<td></td>
<td>Mylar bags; washout; filtered air; breath sampled against resistance to</td>
<td>SPME/GC-MS, Gold nanoparticle sensors</td>
<td>Lung cancer</td>
<td></td>
<td>[79]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BioVOC$^TM$</td>
<td>BioVOC$^TM$; targeted late expiratory sampling; single breath</td>
<td>SPME/GC-MS</td>
<td>Lung cancer</td>
<td>Untargeted</td>
<td>[61]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Aldehydes</td>
<td>[60]</td>
</tr>
<tr>
<td></td>
<td>BioVOC$^TM$; single breath; adsorbent tubes</td>
<td>GC-MS</td>
<td>Liver disease</td>
<td>Ketones, methylsulfides, indole, dimethyl selenide, 1-propanol</td>
<td>[81]</td>
</tr>
<tr>
<td>Adsorbent tubes</td>
<td>Washout; filtered air; targeted late expiratory sampling; multiple breaths per sample</td>
<td>TD/GC-ToFMS</td>
<td>Asthma</td>
<td>Untargeted</td>
<td>[38]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TD/GC-MS</td>
<td>COPD</td>
<td>Untargeted</td>
<td>[19]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Unknown (IMS-based)</td>
<td>[63]</td>
</tr>
<tr>
<td></td>
<td>Targeted late expiratory sampling; adsorbent trap ‘Breath Collection Apparatus’</td>
<td>TD/GC-MS</td>
<td>Lung cancer</td>
<td>Untargeted</td>
<td>[82]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Oxidative stress</td>
<td>Alkanes</td>
<td>[23]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Tuberculosis</td>
<td>Untargeted</td>
<td>[83]</td>
</tr>
</tbody>
</table>

*Studies were selected to demonstrate the variability of techniques used for both breath sampling and sample analysis.

---

**References:**

[23, 39], [40], [41], [42], [43], [44].
Advancements in GC-MS
As far back as the 1970s, 2D GC-MS (GC×GC-MS) had been recognized as a powerful extension of standard GC-MS [45]. This comprehensive technique involves utilization of an orthogonal secondary short column that significantly improves the separating ability within the system by an order of magnitude. The resultant peaks are then displayed in a 2D plot that allows for more accurate identification compared with standard GC-MS, because it uses two specific coordinates instead of a single retention time, along with the standard mass fragmentation patterns. Use of this technology, such as the Shimadzu Comprehensive GCxGC and the Leco Pegasus 4D ToF-MS systems, for breath VOC analysis is in its infancy but has been successfully used to elucidate healthy breath VOC patterns [46] and certain diseases [45,47]. However, widespread use has previously been hampered by an expensive outlay, complications in handling cryogenic gases,

Box 2. Current technologies being adapted to portable breath analysis

| Portable and handheld field asymmetric ion mobility spectrometry (FAIMS) systems, such as the UltraFAIMS system (Owlstone) benefit from working at atmospheric pressure and chipset miniaturization (Figure I). Breath samples can be introduced into the central drift tube via the ion source where asymmetric tuning of the control voltage allows for the targeted detection of specific metabolite ions to reach a detector (in Figure I, this is the yellow trace). This information can be plotted and incorporated into a pattern recognition linked to specific disease biomarker fingerprints. Microfabricated piezoelectric cantilever arrays (Figure II) have the advantage of being able to be tailored to almost any molecular species while working at atmospheric pressures. Figure II A comprises a two-part blank electroactive cantilever that has a base level of electrical activity.

![Image](TRENDS in Biotechnology)

**Figure I.** The flow path within a portable ion mobility system. The portability and directed application of such set-ups hold potential within clinical settings. Abbreviations: IM, ion mobility; MS, mass spectrometry.

**Figure II.** The flow path and feedback loops within a diode-based laser tuning system.

Functionalization by metabolite sensitive ligands (Figure II B) implies specificity to the device and the final step involves target metabolite binding to the functionalized ligand (Figure II C). This causes a change in the laser electronic or resonance signal that can be detected by the system. Bespoke panels of different levers can be designed around diverse biomarker panels, allowing for a truly tunable platform.

Laser absorption systems (Figure III) use the versatility of diode-based laser tuning in a highly targeted manner. A laser beam of specific energy is transmitted into a chamber filled with a breath sample and this beam interacts with specific metabolites. The laser beam then converges upon the photodetector, which translates the beam, and the information it conveys (such as metabolite presence and concentration), into an electrical signal that is returned to the signal processing unit.

![Image](TRENDS in Biotechnology)

**Figure III.** The three stages of fabrication within piezoelectric cantilever arrays.
and potentially large data files that can be cumbersome to process in a high throughput environment.

**Chemical arrays**

As mentioned above, limitations in the technology associated with the GC-MS setup, a routinely used tool in breath analysis, relate to acquisition and portability of the instrumentation and complexity of data analysis. Preconcentration of breath samples is often required that enhances the bias through detection enhancement of specific VOCs, whereas water vapor influences the detection of specific VOCs. In a study by Peled et al., VOCs were detected by cross-reactive chemical nanosensor devices in the presence of water vapor at a higher accuracy than by GC-MS to differentiate between benign and malignant pulmonary nodules. The principle VOCs responsible for differentiation between small cell lung cancer and non-small cell lung cancer were identified as decanal, acetonaphone, and 1,3-bis (1,1-dimethylethyl)-benzene with an *in vitro* separation of 100% sensitivity and 75% specificity [48]. Microfabricated piezoelectric cantilevers are another attractive alternative (Box 2).

**The bioelectronic nose (eNoses)**

The olfactory system has a major role in the differentiation between a diversity of low-molecular-weight VOCs. Olfactory receptors (ORs) are the largest multigene G-protein-coupled receptors, and their expression is restricted to the nasal epithelia. When odors are presented to the nasal cavity, they are selectively recognized by ORs, which trigger intracellular signal transduction pathways that transmit signals to the brain cortex for processing.

Historically, live animals have been used as olfactory biosensors in various disciplines ranging from disease detection to bioterrorism. In addition to live animals, various components of the olfactory system, such as tissues, cells, and olfactory-related proteins, have been considered for use as detection elements in olfactory-based biosensors. eNoses, which are constructed from highly sensitive biological olfactory components that are coupled to electronic transducers, have received increasing interest in the development of sensitive and precise assays for detecting VOCs [49]. One device comprises cells expressing ORs cultured onto surface plasmon resonance (SPR) chips [50]. SPR signals are generated as a result of signal transduction initiated by binding of odorant to ORs of cultured cells, and the response is dose dependent.

ORs and proteins offer promising prospects for application because they can be manufactured on a commercial scale and are more stable, thus are more attractive for application in field, than cell-based olfactory components. Following commercialization, eNoses could be used to fingerprint breath components as an initial step in differentiating between healthy and diseased breath samples, allowing subsequent analysis with GC-MS to quantify the concentration of specific VOCs, and ultimately assisting in novel biomarker detection [51].

**Laser-based technologies**

Laser systems are a sensitive means of targeting trace amounts of VOCs in the human breath and numerous analytical platforms fall under this category. Recent advances in laser and photodetector capabilities have rendered laser-based technologies sensitive to trace amounts of compounds with the scope for an increasing range of commercial PoC technologies becoming available. These advances are counteracted by complications in specificity of detection within complex media (such as breath), quantification, and multiplexing detection ability.

Direct laser absorption spectrometry (DLAS) relies on the quantification of laser light absorbed by multicomponent gas mixtures to determine their respective concentrations. Challenges associated with the detection limit using this approach have related to its susceptibility to noise or interference, the requirement for transitions with larger strengths, and the need for a longer path length for enhancing detection sensitivity. Hence, laser-based absorption measurements have largely been superseded by technologies that modulate the laser source (e.g., quantum cascade lasers) [51], use transitions in unconventional wavelength regions, or increase the effective path length through inclusion of external cavities (e.g., cavity-enhanced comb and cavity ring-down spectroscopies) [52].

Cavity-enhanced comb spectrometry relies on signal enhancement through the cavity multipass effect in which the signal is recovered through comparison of light transmitted through a cavity in the presence or absence of intracavity absorption. This approach has been applied in a bid to detect minimum detectable concentrations of stable isotope ratios of CO₂, CO, and ammonia in a high-water background as ‘proof of concept’ [53].

Cavity ring-down spectrometry (CRDS) is a sensitive laser-based technique for absorption measurements. The radiation pulse is injected into a stable optical cavity (resonator) built of two highly reflective mirrors. Through multiple reflections between the mirrors, the radiation is trapped in the cavity and small portions of its intensity are transmitted after each encounter with a mirror, eventually giving rise to an exponential decay profile [52]. Ammonia and ethylene are biomarkers of interest to several pathologies, for which breath analysis was performed in real-time using CRDS [54]. Although CRDS has existed since 1988, advances in laser and photodetector technologies have contributed to several modifications of this technology, enhancing its sensitivity [55].

Laser photo-acoustic absorption spectrometry (LPAS) uses intermittent laser pulses to irradiate samples, and signals are generated as a consequence of converting electromagnetic energy into sound waves. Thus, the photoacoustic signal produced will directly depend on analyte concentration [55]. Quantification of ethylene by LPAS was used to assess ethylene as a biomarker for renal failure in older patients [56].

**Concluding remarks and future perspectives**

Exhaled VOCs have the potential to aid rapid disease detection, prognostication, and drug response at the point of clinical care through biomarker discovery (Box 3). However, a major challenge limiting the application of this approach is the lack of standardization in breath collection, profiling detection platforms, and robust statistical analyses. Use of more sensitive and portable approaches that
Box 3. Outstanding questions

- How can we standardize lab-based breath collection methods to ensure sample consistency for cross-validation of different data sets?
- What is the best approach to link high-confidence identification of biomarker patterns with sensor signals?
- How can we increase the complexity, sensitivity, and specificity of eNoses to compete with olfactory tissues?
- How can we develop algorithms for breath pattern detection (finger printing for specific disease groups)?
- How can we link breath metabolome data to corresponding proteomic and genetic information (ultraomic approach)?
- What can be done to improve the portability of MS-based technologies?

allow the selective identification and quantification of identified chemical groups will increase accessibility and application in the clinical setting in the forthcoming years. Current development and discovery of new tools within the breathomics pipeline may render it a reality in the clinic, not only for the diagnosis and monitoring of pulmonary diseases, but also for a whole host of systemic diseases.

Acknowledgments

N.J.W.R. and R.G. would like to acknowledge the UK Medical Research Council (MRC) and UK Biotechnology and Biological Sciences Research Council (BBSRC) for generous funding, and D.K.T. and R.G. also thank Cancer Research UK (CRUK) for funding. N.J.W.R. and D.K.T. would also like to thank Markes International alongside H. Poole and G. Lo from Hichrom Ltd for helpful discussions.

References

1. Ibrahim, B. et al. (2013) Breath metabolomic profiling by nuclear magnetic resonance spectroscopy in asthma. Allergy 68, 1050–1056
Review


Popa, C. et al. (2013) The level of ethylene biomarker in the renal failure of elderly patients analyzed by photoacoustic spectroscopy. Laser Phys. 23, 4


Sinues, P.M. et al. (2013) Human breath analysis may support the existence of individual metabolic phenotypes. PLoS ONE 8, e73274

Mochalski, P. et al. (2013) Near real-time VOCs analysis using an aspiration ion mobility spectrometer. J. Breath Res. 7, 11


Barash, O. et al. (2009) Sniffing the unique “odor print” of non-small-cell lung cancer with gold nanoparticles. Small 5, 2618–2624


548