Exhaled Volatile Organic Compounds of Infection: A Systematic Review

Waqar M. Ahmed, †,‡ Oluwasola Lawal, †,‡ Tamara M. Nijsen, † Royston Goodacre, § and Stephen J. Fowler* †,‡,∥

Division of Infection, Immunity & Respiratory Medicine, School of Biological Sciences, The University of Manchester, Oxford Road, Manchester, M13 9PL, United Kingdom

Philips Research, Royal Philips B.V., High Tech Campus 34, Eindhoven, 5656 AE, The Netherlands

School of Chemistry, Manchester Institute of Biotechnology, The University of Manchester, 131 Princess Street, Manchester, M1 7DN, United Kingdom

Manchester Academic Health Science Centre, University Hospital of South Manchester NHS Foundation Trust, Southmoor Road, Wythenshawe, Manchester, M23 9LT, United Kingdom

ABSTRACT: With heightened global concern of microbial drug resistance, advanced methods for early and accurate diagnosis of infection are urgently needed. Analysis of exhaled breath volatile organic compounds (VOCs) toward detecting microbial infection potentially allows a highly informative and noninvasive alternative to current genomics and culture-based methods. We performed a systematic review of research literature reporting human and animal exhaled breath VOCs related to microbial infections. In this Review, we find that a wide range of breath sampling and analysis methods are used by researchers, which significantly affects interstudy method comparability. Studies either perform targeted analysis of known VOCs relating to an infection, or non-targeted analysis to obtain a global profile of volatile metabolites. In general, the field of breath analysis is still relatively immature, and there is much to be understood about the metabolic production of breath VOCs, particularly in a host where both commensal microflora as well as pathogenic microorganisms may be manifested in the airways. We anticipate that measures to standardize high throughput sampling and analysis, together with an increase in large scale collaborative international trials, will bring routine breath VOC analysis to improve diagnosis of infection closer to reality.

KEYWORDS: breath, metabolomics, mass spectrometry, infectious disease, respiratory infection, volatile organic compounds, VOCs, microbial metabolites

Global health organizations and government agencies have expressed the need for accurate and timely diagnosis of microbial infection and reduced occurrence of antibiotic resistance.1,2 Analysis of exhaled breath has the potential to provide a highly informative and noninvasive alternative to current methods, many of which are based on invasive procedures paired with subsequent culture of the pathogen.3 Inert gases and water vapor are the major constituents of breath. Nonvolatile compounds are predominantly found in exhaled breath condensate.4,5 Volatile organic compounds (VOCs) and other low molecular weight compounds (such as hydrogen and ammonia) can provide a fingerprint of metabolite production and are released by various bodily routes, through urine, skin, feces, and breath.6 Exhaled breath VOCs have been studied extensively for pollutant exposure and disease diagnosis.7–9 Hundreds of VOCs have been identified from breath, which have both endogenous and exogenous origins.10,11 There are many advantages with analyzing breath VOCs, which include the following:

● Noninvasive and less strenuous sampling for subjects
● Theoretically unlimited samples of breath gas
● Free from biological and cellular material
● Direct access to the lungs and indirect access to the systemic circulation

In the past decade, several breath tests have been approved by the FDA for clinical use.12 For instance, exhaled nitric oxide is related to asthmatic airway inflammation and has been subsequently approved for clinical use.13,14 Other such breath tests include diagnosis of Helicobacter pylori15 and detection of a group of alkanes to screen patients with heart transplant rejection.16

A range of instruments are used to analyze breath VOCs and are discussed in more detail elsewhere.17,18 One of the most commonly used is gas chromatography—mass spectrometry (GC-MS), usually coupled with a preconcentration method...
such as thermal desorption or solid phase microextraction. Selected ion flow tube MS (SIFT-MS) and proton-transfer reaction MS (PTR-MS) instruments are also used to perform highly sensitive breath VOC analysis without the need for prior breath capture and preconcentration. Both instruments use soft ionization and, in comparison to electron ionization used in GC-MS, have the advantage that they analyze the intact product ion, which can then undergo controlled product ion fragmentation from collision reactions with reagent ions. Researchers have also applied secondary electrospray ionization-MS (SESI-MS) to breath analysis, avoiding preconcentration and requirement of specified precursor ions. Ion mobility spectrometry (IMS) and differential mobility spectrometry (DMS) instruments offer a more versatile alternative to bulky MS instruments. Here, ions are separated in the gas phase, where the drift time is proportional to the ion’s mobility through a controlled electrical field. Alternatively, electronic nose (e-nose) devices provide quick and simple classification between gaseous samples and have proven useful in differentiating disease from healthy subjects via exhaled breath.

Volatile organic compounds are prone to high temporal and diurnal variation and are inherently unstable depending on their chemical characteristics and the sampling method employed. Therefore, as for the majority of small molecule studies, enhanced data processing methods and standards are required to ensure robust and unbiased analysis. Technical standards and best practices are available and are continually developing within the breath research community.

Data preprocessing and machine learning techniques within breath research have also been reviewed.

**Scope of this Systematic Review.** A range of VOCs are emitted by microbes in vitro. It is therefore assumed that exhaled breath VOC profiles can be linked to volatile metabolite fingerprints emitted by microbes during infection pathogenesis within the host environment. The aim of this systematic Review is to assess research where breath VOCs are sampled and analyzed, toward diagnosis of microbial infections and infectious diseases. We will evaluate and compare breath VOCs related to infection and their sampling and analysis methods.

### Bibliographic Search Methods

**Database Search.** Studies were chosen on the basis of their description to discriminate infected subjects from control samples using breath VOC analysis. Specifically, this was based upon match with relevant keywords, syntaxes, and article type. This global search was performed using three bibliographic databases, namely, BIOSIS, Embase, and MEDLINE. Figure 1 shows an overview of the bibliometric process used including article attrition and study selection process.

Some article types (specifically patents, letters, conference abstracts, and reviews) were excluded as they did not provide complete information about VOCs or methods and therefore were considered not relevant toward the scope of this Review. However, relevant articles were archived and reviewed as part of the quality assessment later on. In most cases, these
<table>
<thead>
<tr>
<th>study</th>
<th>publication year</th>
<th>subject size</th>
<th>target infection</th>
<th>headspace analysis performed?</th>
<th>instrument</th>
<th>potential breath biomarkers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hockstein et al.</td>
<td>2004</td>
<td>33</td>
<td>VAP&lt;sup&gt;b&lt;/sup&gt;</td>
<td>N</td>
<td>e-nose</td>
<td>N/A</td>
</tr>
<tr>
<td>Ruzsanyi et al.</td>
<td>2005</td>
<td>40</td>
<td>pulmonary (candida and pneumonia infections)</td>
<td>N</td>
<td>MCC-IMS</td>
<td>N/A</td>
</tr>
<tr>
<td>Hockstein et al.</td>
<td>2005</td>
<td>50</td>
<td>VAP&lt;sup&gt;b&lt;/sup&gt;</td>
<td>N</td>
<td>e-nose</td>
<td>N/A</td>
</tr>
<tr>
<td>Lechner et al.</td>
<td>2005</td>
<td>25</td>
<td>H. pylori</td>
<td>N</td>
<td>PFR-MS</td>
<td>hydrogen nitrate, hydrogen cyanide&lt;sup&gt;e&lt;/sup&gt;</td>
</tr>
<tr>
<td>Phillips et al.</td>
<td>2007</td>
<td>101</td>
<td>TB&lt;sup&gt;c&lt;/sup&gt;</td>
<td>Y</td>
<td>TD-GC-MS</td>
<td>1-methyl-naphthalene,&lt;sup&gt;e&lt;/sup&gt; 1,4-dimethyl-cyclohexane, and derivatives of heptane and benzene (aligned with in vitro results)</td>
</tr>
<tr>
<td>Chambers et al.</td>
<td>2009</td>
<td>56</td>
<td>IA&lt;sup&gt;c&lt;/sup&gt;</td>
<td>N</td>
<td>SPME-GC-MS/MS</td>
<td>2-pentylfuran&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td>Syhre et al.</td>
<td>2009</td>
<td>20</td>
<td>TB</td>
<td>N</td>
<td>SPME-GC-MS/MS</td>
<td>methyl nicotinate</td>
</tr>
<tr>
<td>Phillips et al.</td>
<td>2010</td>
<td>33</td>
<td>influenza A</td>
<td>Y</td>
<td>TD-GC-MS</td>
<td>2,8-dimethyl-undecane, 4,8-dimethyl-undecane, and 2-isopropyl-5-methyl-1-heptanol</td>
</tr>
<tr>
<td>Phillips et al.</td>
<td>2010</td>
<td>226</td>
<td>TB</td>
<td>N</td>
<td>TD-GC-MS</td>
<td>3-(1-methylethyl)-oxetane, 4-methyl-dodecane,&lt;sup&gt;e&lt;/sup&gt; hexyl-cyclohexane, bis(3,5,5-trimethylhexyl) phthalate, 1,3,5-trimethyl-benzene,&lt;sup&gt;e&lt;/sup&gt; 3,7-dimethyl-decane, tridecane,&lt;sup&gt;e&lt;/sup&gt; 4,6,8-trimethyl-1-nonen, 5-ethyl-2-methyl-heptane, 4-methyl-1-hexene</td>
</tr>
<tr>
<td>Rabis et al.</td>
<td>2011</td>
<td>53</td>
<td>P. aeruginosa</td>
<td>N</td>
<td>MCC-IMS</td>
<td>N/A</td>
</tr>
<tr>
<td>Ulanowska et al.</td>
<td>2011</td>
<td>29</td>
<td>H. pylori</td>
<td>Y</td>
<td>SPME-GC-MS</td>
<td>isobutane, 2-butanone,&lt;sup&gt;f&lt;/sup&gt; ethyl acetate, styrene,&lt;sup&gt;e&lt;/sup&gt; ethylbenzene,&lt;sup&gt;e&lt;/sup&gt; ethanol,&lt;sup&gt;f&lt;/sup&gt; and butane</td>
</tr>
<tr>
<td>Marsh et al.</td>
<td>2011</td>
<td>9</td>
<td>Influenza A</td>
<td>N</td>
<td>SIFT-MS</td>
<td>↑ nitric oxide, isoprene&lt;sup&gt;e&lt;/sup&gt;</td>
</tr>
<tr>
<td>Phillips et al.</td>
<td>2012</td>
<td>251</td>
<td>TB</td>
<td>N</td>
<td>TD-GC-SAW</td>
<td>camphene, l-beta-pinene, 1,3,5-trimethyl-benzene,&lt;sup&gt;e&lt;/sup&gt; 1-methyl-naphthalene,&lt;sup&gt;e&lt;/sup&gt; tridecane&lt;sup&gt;e&lt;/sup&gt;, 2-butyl-1-octanol&lt;sup&gt;e&lt;/sup&gt;, 4-methyl-dodecane&lt;sup&gt;e&lt;/sup&gt;</td>
</tr>
<tr>
<td>Lin et al.</td>
<td>2013</td>
<td>18</td>
<td>IA</td>
<td>Y</td>
<td>SPME-GC-MS</td>
<td>α-trans-bergamotene, β-trans-bergamotene, a β-vatirenene-like sesquiterpene, trans-geranylacetone</td>
</tr>
<tr>
<td>De Heer et al.</td>
<td>2013</td>
<td>46</td>
<td>IA</td>
<td>N</td>
<td>e-nose</td>
<td>N/A</td>
</tr>
<tr>
<td>Koo et al.</td>
<td>2014</td>
<td>64</td>
<td>IA</td>
<td>Y</td>
<td>TD-GC-MS</td>
<td>↑ 3-carene, n-butyric acid 2-ethylhexyl ester, nonanal,&lt;sup&gt;e&lt;/sup&gt; 2,6,11,15-tetramethyl-hexadecane ↓ ethanol,&lt;sup&gt;f&lt;/sup&gt; 2,6,10-trimethyl-dodecane, 5-methyl-5-propyl-nonane, longifolene, tetradecane&lt;sup&gt;e&lt;/sup&gt;, 2-butyland-1-octanol&lt;sup&gt;e&lt;/sup&gt;</td>
</tr>
<tr>
<td>Fowler et al.</td>
<td>2015</td>
<td>46</td>
<td>lower respiratory tract (H. influenza, S. aureus)</td>
<td>N</td>
<td>TD-GC-TOF-MS</td>
<td>dimethyl sulfoxide (and multiple microbe specific VOCs)</td>
</tr>
<tr>
<td>Filipiak et al.</td>
<td>2015</td>
<td>28</td>
<td>VAP (S. aureus, E. coli, Candida spp., P. aeruginosa)</td>
<td>Y</td>
<td>TD-GC-MS</td>
<td>2-methylbutane, ethanol,&lt;sup&gt;f&lt;/sup&gt; heptane,&lt;sup&gt;e&lt;/sup&gt; ethylbenzene,&lt;sup&gt;e&lt;/sup&gt; carane, tetradecane,&lt;sup&gt;e&lt;/sup&gt; tetradecanal ↓ acetone,&lt;sup&gt;f&lt;/sup&gt; tetrahydro-furan, dodecane</td>
</tr>
<tr>
<td>Schnabel et al.</td>
<td>2015</td>
<td>100</td>
<td>VAP (S. aureus, P. aeruginosa, E. coli, K. pneumoniae, H. influenzae, A. baumannii)</td>
<td>Y</td>
<td>TD-GC-TOF-MS</td>
<td>2-methylbutane, ethanol,&lt;sup&gt;f&lt;/sup&gt; heptane,&lt;sup&gt;e&lt;/sup&gt; ethylbenzene,&lt;sup&gt;e&lt;/sup&gt; carane, tetradecane,&lt;sup&gt;e&lt;/sup&gt; tetradecanal ↓ acetone,&lt;sup&gt;f&lt;/sup&gt; tetrahydro-furan, dodecane</td>
</tr>
<tr>
<td>Zetola et al.</td>
<td>2016</td>
<td>71</td>
<td>TB</td>
<td>N</td>
<td>e-nose</td>
<td>N/A</td>
</tr>
<tr>
<td>Gerritsen et al.</td>
<td>2017</td>
<td>3</td>
<td>IA</td>
<td>Y</td>
<td>TD-GC-MS</td>
<td>1,3-pentadiene, 2-ethyl-1-hexanol, 2-methyl-1-propanol</td>
</tr>
<tr>
<td>Van Oort et al.</td>
<td>2017</td>
<td>93</td>
<td>HAP/CAP&lt;sup&gt;f&lt;/sup&gt;</td>
<td>N</td>
<td>TD-GC-MS</td>
<td>acetone, carbon disulfide, 1-propanol&lt;sup&gt;f&lt;/sup&gt;,2-ethoxy-2-methyl-propone, cyclohexene, methylisobutylketone</td>
</tr>
</tbody>
</table>

<sup>a</sup>Tuberculosis. <sup>b</sup>Ventilator-associated pneumonia. <sup>c</sup>Invasive aspergillosis. <sup>d</sup>Hospital-acquired pneumonia/community-acquired pneumonia. <sup>e</sup>The same VOC identified in another study targeting the same infection (not necessarily same pathogen). <sup>f</sup>The same VOC identified in a study targeting a different infection or subject.
databases listed similar articles; therefore, duplicated articles were removed during the automated search.

Keywords were grouped into three tiers and were based on the scope of this Review:

- The first tier consisted of mandatory keywords within the title ("breath" or "exhaled").
- The second tier included keywords mandatory in the title or abstract ("volatile" or "VOC(s)" or "marker" and "infection" or "infectious" or "microbe" or "bacterial" or "fungal" or "viral" or "pathogen").
- The third and final tier included search terms that may have been present within the title, abstract, or text body ("chromatography" or "spectrometry" or "e-nose" or "spectrum" or "sensor"). After the bibliographic search, the resultant research articles were further filtered, in order to match the inclusion and exclusion criteria.

**Article Screening and Selection Criteria.** For this step, the resulting list of articles from the bibliographic database search was manually selected on the basis of the scope of this Review. This was accomplished by determining whether an article’s title and abstract description fit the predefined exclusion criteria: studies which sampled exhaled breath condensate only; studies using gas-specific analyzers (such as nitric oxide or CO₂); studies targeting upper respiratory tract disorders (for example, rhinitis or halitosis) or where breath VOC analysis was not performed.

Studies using animal subjects were included. Specifically, where studies performed investigations of zoonosis or used animal models for preclinical translational investigations. Similarly, studies investigating microbial infection with a chronic underlying condition, and descriptive feasibility studies, were also included.

During the database search, several articles were not assigned with an article type or were duplicated because of grammatical or punctuation differences. These articles were removed or assessed as part of the manual filtration process.

**Quality Assessment and Additional Searches.** To reduce article selection bias, the final set of articles to be reviewed was agreed upon by two authors (W.M.A. and T.M.N.). The inclusion and exclusion criteria were developed with agreement between all authors.

To ensure only high-quality articles were assessed, the search was limited to peer-reviewed journals. In addition, only original research articles with descriptive methods (sufficient for replication of experiments) were chosen. Where patients were sampled, adherence to ICH-GCP guidelines was mandatory.

A high standard of quality was maintained during database searches and study selection. The database search was performed using the STN platform (Chemical Abstracts Service, American Chemical Society) by two specialists (Center for Information Services, Royal Philips B.V.). Keywords were subjected to syntax rules (e.g., keyword order, keyword distance, suffixes, and plurals, among others) within the English language. Both American and British spelling differences were considered.

Additional manual searches were performed through the literature search engine Scopus (Elsevier B.V., version 2016). Here, each article from the search results was entered, and the article’s reference list and citation list were searched for additional relevant articles, which may have not been included from the original database search.

**REVIEW OF EXTRACTED STUDIES**

**Principle Findings.** The bibliographic database search extracted 1526 articles. After article screening, study selection, and quality assessment methods were applied, the list reduced to 51 articles. The reviewing authors mutually agreed on the final set of articles to be reviewed.

The resultant articles had a publication date from January 2004 to July 2017, and the majority of studies were performed in European and North American countries. Fewer studies were performed in Asian and African countries. Several studies investigated breath VOCs of infection ($n = 24$), including pneumonia and nosocomial respiratory infection ($n = 9$), pulmonary aspergillosis ($n = 5$), pulmonary tuberculosis ($n = 5$), airborne influenza viral infection ($n = 2$), and infection of the circulatory and gastrointestinal systems ($n = 3$). These studies are summarized in [Table 1](#). Studies where subjects had an infection with an underlying noncommunicable disease were also reviewed ($n = 17$). Studies also investigated zoonotic infection or used animal models as preclinical in vivo experiments ($n = 14$). Across these categories, bacterial infections were the most studied ($n = 26$), followed by fungal ($n = 7$), viral ($n = 2$), and protozoan ($n = 1$). Others investigated mixed microbial infection. Some of these studies focused on the pathogenesis of the underlying disease and performed experiments on a subgroup with infection.

**Respiratory Tract Infections in Humans.** Pneumonia and Nosocomial Respiratory Infection. Respiratory infections are a common occurrence within hospitals due to the compromised nature of patients and their exposure to the clinical microbial community. Sampling from patients within a clinical setting requires investigators to adhere to regulations that govern the safety of clinical staff and patients. Therefore, the development and adaptation of devices to sample safely and efficiently patient breath is mandatory. With IMS-based and e-nose devices, it is possible to analyze immediately after sampling; for example, Ruzsanyi et al. and Rabis et al. utilize IMS-based technology with an overall test accuracy of above 82% from the latter study. Similarly, Hockstein et al. showed the ability of a commercial e-nose sensor to differentiate exhaled breath gas from mechanically ventilated patients diagnosed with ventilator-associated pneumonia (VAP). The authors compared breath analysis to computerized tomography scans and pneumonia scores, revealing an overall test accuracy of 60% for the latter.

A number of studies have investigated the capability of breath analysis to diagnose VAP and other respiratory infections within patients under mechanical ventilation ($n = 6$). Filipiak et al. linked VOCs extracted from the headspace to patient breath, using cultured swabs as a reference. For instance, 4-heptanone was found in the headspace of *Candida albicans*, and breath samples from patients with confirmed *C. albicans* infection. Similarly, Gao et al. found distinctive VOCs in the headspace of *Acinetobacter baumannii* culture and linked the findings to VAP patients with *A. baumannii* infection. Interestingly, VAP patients with and without *A. baumannii* infections could be differentiated. Both studies use a targeted approach where VOCs are known prior to in vivo analysis, thereby increasing confidence in chemical identification.

In contrast to targeted analysis, other studies performed nontargeted analysis. This type of study design theoretically allows for maximum coverage (depending on the sampling and analysis methods) of both known and unknown compounds.
Table 2. Studies Investigating Infection within Underlying Chronic Disease

<table>
<thead>
<tr>
<th>Study</th>
<th>Publication Year</th>
<th>Subject Size</th>
<th>Target Pathogen</th>
<th>Underlying Disease</th>
<th>Headspace Analysis Performed?</th>
<th>Instrument</th>
<th>Potential Breath Biomarkers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kamboures et al.</td>
<td>2005</td>
<td>43</td>
<td>P. aeruginosa, S. aureus</td>
<td>CF</td>
<td>N</td>
<td>custom-made system</td>
<td>N/A</td>
</tr>
<tr>
<td>Barker et al.</td>
<td>2006</td>
<td>40</td>
<td>P. aeruginosa</td>
<td>CF</td>
<td>N</td>
<td>GC-FID-MS</td>
<td>↑ pentane ↓ methanol</td>
</tr>
<tr>
<td>Syhre et al.</td>
<td>2008</td>
<td>21</td>
<td>A. fumigatus</td>
<td>CF</td>
<td>Y</td>
<td>SPME-GC-MS/MS</td>
<td>2-pentylfuran&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>Enderby et al.</td>
<td>2009</td>
<td>34</td>
<td>P. aeruginosa</td>
<td>CF and asthma</td>
<td>N</td>
<td>SIFT-MS</td>
<td>hydrogen cyanide&lt;sup&gt;d&lt;/sup&gt;</td>
</tr>
<tr>
<td>Robbrooks et al.</td>
<td>2010</td>
<td>105</td>
<td>P. aeruginosa</td>
<td>CF</td>
<td>N</td>
<td>TD-GC-TOF-MS</td>
<td>N/A</td>
</tr>
<tr>
<td>Scott-Thomas et al.</td>
<td>2010</td>
<td>46</td>
<td>P. aeruginosa</td>
<td>CF</td>
<td>Y</td>
<td>SPME-GC-MS/MS</td>
<td>2-aminoacetophenone</td>
</tr>
<tr>
<td>Shestivska et al.</td>
<td>2011</td>
<td>28</td>
<td>P. aeruginosa</td>
<td>CF</td>
<td>Y</td>
<td>SIFT-MS</td>
<td>methyl thiocyanate</td>
</tr>
<tr>
<td>White et al.</td>
<td>2013</td>
<td>14</td>
<td>P. aeruginosa, S. aureus, A. fumigatus</td>
<td>CF</td>
<td>N</td>
<td>PTR-TOF-MS</td>
<td>N/A</td>
</tr>
<tr>
<td>Gächrist et al.</td>
<td>2013</td>
<td>20</td>
<td>P. aeruginosa</td>
<td>CF</td>
<td>N</td>
<td>SIFT-MS</td>
<td>hydrogen cyanide&lt;sup&gt;d&lt;/sup&gt;</td>
</tr>
<tr>
<td>Dunner et al.</td>
<td>2013</td>
<td>26</td>
<td>P. aeruginosa</td>
<td>bronchiectasis</td>
<td>N</td>
<td>SIFT-MS</td>
<td>hydrogen cyanide&lt;sup&gt;d&lt;/sup&gt;</td>
</tr>
<tr>
<td>Sibila et al.</td>
<td>2014</td>
<td>50</td>
<td>S. aureus, P. aeruginosa, S. pneumoniae, H. influenzae, M. catarrhalis, Candida spp., S. viridans, S. epidermidis, Corynbacterium spp.</td>
<td>CF</td>
<td>N</td>
<td>e-nose</td>
<td>N/A</td>
</tr>
<tr>
<td>Kramer et al.</td>
<td>2014</td>
<td>11</td>
<td>P. aeruginosa, S. aureus, C. albicans, A. xylosoxidan</td>
<td>CF</td>
<td>Y</td>
<td>SPME-GC-TOF-MS</td>
<td>N/A</td>
</tr>
<tr>
<td>Joensen et al.</td>
<td>2014</td>
<td>106</td>
<td>P. aeruginosa</td>
<td>CF and PCD&lt;sup&gt;c&lt;/sup&gt;</td>
<td>N</td>
<td>e-nose</td>
<td>N/A</td>
</tr>
<tr>
<td>Whiteson et al.</td>
<td>2014</td>
<td>43</td>
<td>S. aureus, S. aureus, R. mucilaginosa</td>
<td>CF</td>
<td>N</td>
<td>GC-MS</td>
<td>2,3-butanedione</td>
</tr>
<tr>
<td>Slufek et al.</td>
<td>2015</td>
<td>185</td>
<td>S. aureus, P. aeruginosa, S. pneumoniae, H. influenzae, M. catarrhalis, Candida spp., S. viridans, Neisseria spp., S. epidermidis, Corynbacterium spp.</td>
<td>COPD</td>
<td>Y</td>
<td>e-nose</td>
<td>N/A</td>
</tr>
<tr>
<td>Gächrist et al.</td>
<td>2015</td>
<td>233</td>
<td>P. aeruginosa</td>
<td>CF</td>
<td>N</td>
<td>SIFT-MS</td>
<td>hydrogen cyanide&lt;sup&gt;d&lt;/sup&gt;</td>
</tr>
<tr>
<td>Neeriox et al.</td>
<td>2016</td>
<td>18</td>
<td>S. aureus</td>
<td>CF</td>
<td>N</td>
<td>TD-GC-MS</td>
<td>↑ 1,4-pentadiene, ethanol, acetone, 2-butanone, undecane, 2-methyl naphthalene, 3-hydroxy-2-butanone, isopropyl myristate, hexanal</td>
</tr>
</tbody>
</table>

<sup>a</sup>Cystic fibrosis <sup>b</sup>Chronic obstructive pulmonary disease. <sup>c</sup>Primary ciliary dyskinesia. <sup>d</sup>The same VOC identified in another study targeting the same infection (not necessarily the same pathogen). <sup>e</sup>The same VOC identified in a study targeting a different infection or subject.
within a given sample, usually hypothesis-free and with the use of multivariate analysis and statistical validation.\textsuperscript{38,39} Fowler et al. associated a range of VOCs with lower respiratory tract infection of patients in the intensive care unit.\textsuperscript{59} They discussed that the influence on breath VOC profile might not solely be associated with pathogens but also the inflammatory response. Van Oort et al. discovered 1-propanol as a potential biomarker of bacterial growth in patients with pneumonia.\textsuperscript{60} With regard to confirmed VAP diagnosis breath profiles, some similar VOCs were found and linked to metabolic pathways.\textsuperscript{61} Between the VAP or VAP-like studies described above, commonly reported VOCs included ethanol, heptane, nonanal, and carane.

**Pulmonary Aspergillosis.** Patients with a weakened immune system may also be prone to infection by *Aspergillus* spp., the causative agents of pulmonary aspergillosis. All studies extracted from the search investigated invasive aspergillosis (IA), a severe form of pulmonary aspergillosis. Terpene-based compounds were found in vitro and compared to breath samples from patients with confirmed IA\textsuperscript{11,42}. Koo et al. reported additional terpene-based VOCs, supported by reference standards and argued β-trans-betagermol was misidentified as β-farnesene in an older study.\textsuperscript{42} Gerritsen et al. also compared in vitro headspace to breath samples and aligned three VOCs between the two VOC profiles; however, the number of breath samples were low.\textsuperscript{62} Chambers et al. performed targeted analysis of 2-pentylfuran from culture headspace experiments and showed its potential as a biomarker for IA.\textsuperscript{44} However, 2-pentylfuran was not identified in other studies, although other low molecular weight chemical species were found.\textsuperscript{31–45} One study made use of e-nose technology, and together with robust feature classification methods, was able to show a distinction between the breath of prolonged chemotherapy-induced neutropenia patients with IA infection and healthy controls (cross validated accuracy of 90.9%).\textsuperscript{45}

**Pulmonary Tuberculosis.** As mentioned earlier, current diagnosis of microbial infectious diseases involves sampling methods which can be invasive. Samples are then cultured and typed (such as using proteomics-based diagnostics) and/or analyzed directly for specific genetic loci. It is important to note that microbial pathogens may remain dormant for a long period of time before symptoms of infection emerge and some may become latent; therefore, the presence of DNA alone does not mean that there is pathogenicity.

Early diagnosis of pulmonary tuberculosis (pTB) infection is important to prevent progressive illness and drug resistance; therefore, several studies have investigated breath VOCs of subjects infected with *Mycobacterium tuberculosis*. Philips et al. found a range of VOCs associated with pTB, notably methylated and nonmethylated hydrocarbons, with 1-methyl naphthalene, and 1,4-dimethyl cyclohexane aligned between both in vitro culture headspace and in vivo patient breath samples.\textsuperscript{46} The same authors found methylated and nonmethylated hydrocarbons (including derivatives of benzene) from larger international studies and with one study analyzing samples on-site (point-of-care, POC).\textsuperscript{47,48} Where targeted analysis was conducted, researchers found methyl nicotinate increased in patient breath, compared to routine sputum smear tests.\textsuperscript{49} This was based upon cited in vitro studies where nicotinic acid was found in the headspace of cultured *M. tuberculosis*.

Another study by Zetola et al. reported distinguishable pTB breath VOC profiles using an e-nose device.\textsuperscript{50} Interestingly, these authors reported unique profiles of pTB with and without drug treatment. They were also able to show differences between patients with human immunodeficiency virus and pTB, compared to pTB alone. Although this could include VOCs related to the response of the immune system, it also indicates a VOC profile can depend on the concentration and diversity of a microbial community within the host environment.

**Influenza.** An increased interest for effective treatment of airborne influenza virus following an influenza subtype A(H1N1) pandemic in 2009\textsuperscript{51} led to studies measuring the change in a subject’s breath signature over time, after vaccine inoculation.\textsuperscript{52,53} It was found that the concentration of breath VOCs, especially those emitted following an immune response (nitric oxide and oxidative stress-associated VOCs), was correlated with vaccine metabolism over time. Phillips et al. showed 2,8-dimethyl-undecane had a positive correlation with vaccine response over time.\textsuperscript{52}

**Infection within Underlying Chronic Lung Disease.** Non-communicable disease can be exacerbated by microbial pathogens, particularly for pulmonary diseases. Studies focused on obtaining breath VOC profiles of infected subjects together with an underlying chronic disease. A summary of these studies is listed in Table 2.

A few studies profiled breath VOCs from subjects with chronic obstructive pulmonary disease and mixed microbial infection and subsequently were able to distinguish between infected and noninfected subjects.\textsuperscript{54,55} It is worth mentioning that Shaiek et al. also found significantly different "breathprints" for subjects with pneumonia as a comorbidity (sensitivity 85%, specificity 86%).\textsuperscript{55}

Pulmonary infection in subjects with cystic fibrosis (CF) was found to be of particular interest among researchers. Feasibility studies have described efficient sampling and methods for high throughput and real-time analysis.\textsuperscript{56,57} In CF, the majority of studies focused on profiling pulmonary infection by *Pseudomonas aeruginosa* and targeted highly volatile compounds. One study reported an increased concentration of pentane and decrease of methanol; however, they also showed high variance in VOC intensity between breath samples.\textsuperscript{58} Cyanide compounds have been linked to *P. aeruginosa* infection within CF subjects, specifically hydrogen cyanide and methyl thiocyanate,\textsuperscript{59–61} using SIFT-MS together with in vitro experiments.

To show hydrogen cyanide as a potential breath biomarker of CF patients with infection, a follow up study with a larger sample size was performed. The study confirmed that HCN was indeed an early breath biomarker in children with CF; however, with low sensitivity, routine diagnostic use is unlikely to be useful.\textsuperscript{62} Hydrogen cyanide was also found not to be characteristic of patients with chronic supportive lung disease and infection with *P. aeruginosa*.\textsuperscript{63} The authors of the latter study also emphasized the importance of controlled sampling from the mouth, as hydrogen cyanide had high intersample variation when collected from the mouth.

Whiteson et al. demonstrated the presence of 2,3-butadione in both breath gas and microbial metabolism in sputum, an important outcome for further validating the use of breath biomarkers.\textsuperscript{64} Scott-Thomas et al. used GC-MS and detected 2-aminoacetophenone (a larger and less volatile compound than the cyanide compounds mentioned earlier) as a potential breath biomarker for CF subjects with *P. aeruginosa* infection (sensitivity 93.8%, specificity 69.2%).\textsuperscript{65} Using similar techniques, two studies alternatively took advantage of nontargeted
**Table 3. Studies Investigating Infection within Animal Subjects**

<table>
<thead>
<tr>
<th>Study</th>
<th>Year</th>
<th>Subject Size</th>
<th>Zoonotic or Model Study</th>
<th>Animal Subfamily</th>
<th>Target Infection</th>
<th>Headspace Analysis Performed</th>
<th>Analysis Method</th>
<th>Potential Breath Biomarkers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spinhirne et al.</td>
<td>2004</td>
<td>6</td>
<td>Zoonosis</td>
<td>Bovine</td>
<td>Respiratory tract</td>
<td>N</td>
<td>SPME-GC-MS</td>
<td>Acetaldehyde and decanal&lt;sup&gt;e&lt;/sup&gt;</td>
</tr>
<tr>
<td>Purkhart et al.</td>
<td>2011</td>
<td>18</td>
<td>Zoonosis</td>
<td>Caprine</td>
<td>ParaTB&lt;sup&gt;a&lt;/sup&gt;</td>
<td>N</td>
<td>MCC-DMS</td>
<td>N/A</td>
</tr>
<tr>
<td>Peled et al.</td>
<td>2012</td>
<td>27</td>
<td>Zoonosis</td>
<td>Bovine</td>
<td>Bovine TB&lt;sup&gt;b&lt;/sup&gt;</td>
<td>N</td>
<td>TD-GC-MS and e-nose</td>
<td>2,3-dimethyl-1,3-pentadiene, and 1,3-dimethylbutyl cyclohexane</td>
</tr>
<tr>
<td>Zhu et al.</td>
<td>2013</td>
<td>10</td>
<td>Animal model</td>
<td>Murine</td>
<td>Respiratory tract</td>
<td>Y</td>
<td>SESI-MS</td>
<td>N/A</td>
</tr>
<tr>
<td>Zhu et al.</td>
<td>2013</td>
<td>12</td>
<td>Animal model</td>
<td>Murine</td>
<td>Respiratory tract</td>
<td>N</td>
<td>SESI-MS/MS</td>
<td>N/A</td>
</tr>
<tr>
<td>Zhu et al.</td>
<td>2013</td>
<td>86</td>
<td>Animal model</td>
<td>Murine</td>
<td>Respiratory tract</td>
<td>N</td>
<td>SESI-MS</td>
<td>N/A</td>
</tr>
<tr>
<td>Barbour et al.</td>
<td>2013</td>
<td>24</td>
<td>Animal model</td>
<td>Murine</td>
<td>Borrelia hensii</td>
<td>Y</td>
<td>GC-FID</td>
<td>↑ Carbon monoxide&lt;sup&gt;f&lt;/sup&gt;</td>
</tr>
<tr>
<td>Bayn et al.</td>
<td>2013</td>
<td>38</td>
<td>Zoonosis</td>
<td>Bovine</td>
<td>Brucellosis</td>
<td>N</td>
<td>TD-GC-MS and e-nose</td>
<td>2-ethyl-1-hexanol, acetophenone, benzaldehyde, heptanal, octanal</td>
</tr>
<tr>
<td>Bean et al.</td>
<td>2014</td>
<td>11</td>
<td>Animal model</td>
<td>Murine</td>
<td>MRSA&lt;sup&gt;e&lt;/sup&gt;</td>
<td>N</td>
<td>SESI-MS</td>
<td>7 VOC m/z fragments</td>
</tr>
<tr>
<td>Ellis et al.</td>
<td>2014</td>
<td>23</td>
<td>Zoonosis</td>
<td>Bovine</td>
<td>Bovine TB</td>
<td>N</td>
<td>TD-GC-MS</td>
<td>Toluene, styrene, benzaldehyde, 2-ethyl-1-hexanol, α-acetophenone, 1,1-dimethyl 2-(1-methyl) cyclopropane</td>
</tr>
<tr>
<td>Langeroudi et al.</td>
<td>2014</td>
<td>8</td>
<td>Animal model</td>
<td>Murine</td>
<td>Sepsis-inducing</td>
<td>N</td>
<td>GC-FID/ TCD</td>
<td>↑ Carbon monoxide&lt;sup&gt;f&lt;/sup&gt;</td>
</tr>
<tr>
<td>Bergmann et al.</td>
<td>2015</td>
<td>42</td>
<td>Zoonosis</td>
<td>Caprine</td>
<td>ParaTB&lt;sup&gt;a&lt;/sup&gt;</td>
<td>N</td>
<td>NTME-GC-MS</td>
<td>↓ 2-butanone, benzene, 2-methyl-butanol</td>
</tr>
<tr>
<td>Fink et al.</td>
<td>2014</td>
<td>40</td>
<td>Animal model</td>
<td>Murine</td>
<td>Sepsis-inducing</td>
<td>N</td>
<td>MCC-IMS</td>
<td>1-propanol, butanal, acetophenone, 1,2-butandiol, 3-pentanone, acetone, 2-hexanone</td>
</tr>
<tr>
<td>Mellors et al.</td>
<td>2017</td>
<td>5</td>
<td>Animal model</td>
<td>Cercopithecine</td>
<td>TB&lt;sup&gt;b&lt;/sup&gt;</td>
<td>N</td>
<td>TD-GC-xGC-TOF-MS</td>
<td>1,1′-bicyclohexyl, 2,2-dimethylheptane, tridecane, 2-heptanone, allyl heptanoate, 4-methylene-1-(1-methyl)-bicyclo[3.1.0]hexane, 2-methylbutyl ester butanoic acid, n-amyl isovalerate, α-cymene, trans-α-cocimene</td>
</tr>
</tbody>
</table>

<sup>a</sup>Paratuberculosis. <sup>b</sup>Tuberculosis. <sup>c</sup>Methicillin-resistant *Staphylococcus aureus*. <sup>d</sup>The same VOC identified in another study targeting the same infection (not necessarily same pathogen). <sup>e</sup>The same VOC identified in a study targeting a different infection or subject.
metabolomic methods and found distinguishing VOC features from the breath, characteristic of CF subjects with infection.\textsuperscript{66,67} Notably, Neerincx et al. detected a distinctive VOC profile for \textit{Staphylococcus aureus} infection in children with CF (sensitivity 100\%, specificity 80\%).\textsuperscript{67} In contrast, Joensen et al. instead used an e-nose sensor and reported a significant difference for breath profiles of subjects with CF and \textit{P. aeruginosa} infection (sensitivity 71.4\%, specificity 63.3\%).\textsuperscript{68} Interestingly, the authors found no such separation for patients with primary ciliary dyskinesia and \textit{P. aeruginosa} infection. However, separation was found for other colonized pathogens such as \textit{Stenotrophomonas maltophilia}, although the study sample size was comparably low and detection by the e-nose was most likely masked by the excessive signal from VOCs associated with an inflammatory response.

The influence of multiple microbial coinfection contributing to CF has also been investigated.\textsuperscript{69} One study targeted three volatile sulfur compounds.\textsuperscript{69} As for the cyanide compounds mentioned before, the chosen sulfur compounds are highly volatile and similarly would show high intersample variation. Another study identified 2-pentylfuran from the breath of CF patients with \textit{A. fumigatus} infection.\textsuperscript{70} However, 2-pentylfuran was also identified in some patients with no microbial infection. Additionally, some patients infected with \textit{A. fumigatus} also had confirmed infection of \textit{Haemophilus influenzae} and \textit{P. aeruginosa}.

Infectious Diseases of the Circulatory and Gastrointestinal Systems in Humans. Aside from researching respiratory infection, studies have also investigated the potential of breath analysis in infectious diseases of the circulatory or gastrointestinal system (see Table 1 for a list of relevant studies). Urea breath tests for \textit{Helicobacter pylori} infection are routine practice, where urease intensity is measured (the urease enzyme pathway is used by \textit{H. pylori}) after digesting a labeled carbon isotope of urea, which after urease treatment generates labeled CO$_2$. With regard to breath VOCs, increased concentrations of hydrogen nitrate and hydrogen cyanide (by SIFT-MS) and ethanol and butane (SPME-GC-MS) were suggested as potential biomarkers of \textit{H. pylori} infection.\textsuperscript{71,72} However, these compounds are very volatile, are not species specific, are commonly found in breath, and therefore are difficult to translate into distinguishable biomarkers for \textit{H. pylori} infection. In addition, they may not be a suitable cost-effective alternative to established methods.

As endogenous exhaled breath VOCs are directly linked to the circulatory system, bloodstream infections have also been investigated.\textsuperscript{73} Here, concentrations of some VOCs, most notably thioclorides, have shown to increase after colonization by the malaria causing protozoan \textit{Plasmodium falciparum}. After treatment with antimalarial chemotherapy, the concentration of thioclorides reduced significantly. Although thioclorides were not
found in the headspace of *P. falciparum*, they have been linked to many metabolic processes including lipid degradation. However, increased concentration of thioesters in breath may represent the VOC originating from the interaction between *P. falciparum* and the host immune system.

**Studies Investigating Zoonoses.** Early detection of zoonosis is important to both animal and human health. Zoonotic infections can ultimately spread between animals and from animals to humans. Studies investigating zoonosis, or using animal models, are shown in Table 3.

Bovine respiratory infections were predominantly researched compared to other zoonoses (60%).74−77 An early pilot study focused on efficient sampling methods.79 One study focused on brucellosis, a highly contagious zoonosis, and reported breath VOCs linked to *Brucella abortus* infection in bison.77 Early detection of bovine tuberculosis within cattle would prevent infection spreading across dairies. Studies were able to detect VOC profiles of *Mycobacterium bovis* from cattle breath samples.75,76 Interestingly, Peled et al. were able to show variation in breath VOCs for infected cattle and noninfected cattle, within dairies categorized as infected or infection-free.75 Another livestock infection, *Mycobacterium avium* subspecies *paratuberculosis*, was studied within ruminants.78,79 Tentative identification of a range of breath VOCs, or a group of VOC features,78 has been associated with both infected and noninfected ruminants.

**Studies Using Animal Models.** Preclinical *in vivo* models provide fundamental information about infection pathogenesis, the host response, and organ function, in pharmaceutical and medical device development studies. Researchers predominantly used murine animals for infection inoculation and subsequent breath analysis (see Table 3 for selected articles).80−83 although in a recent study Mellors et al. used a macaque animal model to perform nontargeted analysis investigating differences in breath profiles before and after TB infection.84 The authors reported several volatiles, of which dodecane, tridecane, and hexylcyclohexane have been associated with TB in previous human studies.47,48

In other studies by Zhu et al. using murine models, distinct breath VOC profiles between bacterial strains were reported.80 It is noteworthy that only an estimated 25% to 34% of VOC fragments can be directly linked to *in vitro* bacterial culture,83 illustrating that the metabolism of the bacterium inside the host is very different compared with when it is cultivated in an artificial environment. A further study by the same group with an increased subject cohort found lung infection could be discriminated by breath VOCs between *S. aureus* and *P. aeruginosa*.86

Bean et al. analyzed samples from mice inoculated with isogenic strains of methicillin-resistant *S. aureus*, and it was found that VOC fragments, shared with methicillin-sensitive *S. aureus*, showed different concentrations between the two strains.81 In another study by the same group, induced *S. aureus* and *P. aeruginosa* lysates provoked an immune response and lung damage after 48 h, which correlated with the breath VOC fingerprint.82 This is important as VOCs could be used to trace early onset and progress of infection. Where GC was used instead of SESI-MS, researchers sampled breath of mice inoculated with *Borrelia hermsii* and found an increased concentration in carbon monoxide,86 which reduced after antibiotic treatment.

Similarly, increased levels of carbon monoxide (normalized to carbon dioxide levels) were found in rats injected with lipopolysaccharide endotoxin purified from *Escherichia coli* to induce sepsis.87 Several VOCs relating to an inflammatory response (as a result of endotoxin injection) were identified in another study, also investigating sepsis in murine models.88

**REVIEW OF BREATH SAMPLING AND ANALYSIS METHODS**

**Breath Sampling.** Within the scope of this Review, we found that a wide range of breath sampling methods were used, as summarized in Table 4. This was however dependent on the target disease and the environment around the subject. Sampling devices were only used when transportation or storage of breath VOCs was necessary.

The portion of breath collected is important in disease diagnosis. For example, sampling alveolar breath may include a higher proportion of blood-borne VOCs. In some cases, breath was sampled from the nose to reduce the influence of VOCs originating from the mouth.61,63 Studies using animal subjects had to develop bespoke breath sampling systems depending on the animal’s anatomy.75,77,83 Breath from mechanically ventilated patients was sampled from the ventilator circuit line35−38 or by inserting a catheter into the endotracheal tube.39

Volatile organic compounds can be linked to other factors not directly associated with disease. These confounding factors were generally taken into account by most studies using various methods, including restriction of diet or excessive movement immediately prior to sampling and smoking restrictions, among others. Diet is an important consideration as VOCs of interest may change in concentration and therefore yield false-positive results.42,89 Where reported, studies defined dietary restrictions between 1 and 3 h before sampling.

The total volume of sampled breath ranged from a single breath40 to multiple breaths49,50 to predefined sample gas volumes.39 One study sampled a total of 100 mL per subject and concluded this volume was sufficient for discriminating CF subjects with infection (for SPME-GC-MS).50 Gas sampling bags varied in total volume between studies, up to 10 L.45,54 Some studies stored and transported breath in gas sampling bags,29 while others defined a time limit between sampling and analysis.57 Depending on the target infection, a portion of the breath may be sampled as opposed to total breath sampling. This is mandatory for blood-borne infection, where many VOCs may be found in the alveolar breath portion.

Gas sampling bags are known to introduce contaminants into breath gas and can change VOC concentration during storage as polar VOCs can adhere to the plastic surface. In addition, some gas sampling bags are known to be sensitive to amines and sulfur compounds. Some researchers described a sampling system where breath was sampled into evacuated steel containers thereby avoiding adsorption onto sorbent-packed tubes.58,69,86 This method was similarly used with a preconcentration system; however, it was limited to the number of samples that could be analyzed at any one time. Storing breath gas in evacuated steel containers has advantages over gas sampling bags as there is less risk of contamination and VOC degradation. In comparison, the purge and trap method (where breath gas is immediately purged onto sorbent tubes) would allow prolonged storage of breath VOCs (at least several days) and high-throughput analysis.90,91

With regard to sorbent tube use, some studies used sorbent material that were not appropriate for semi-quantitative analysis of highly volatile compounds but were identified and reported
anyway. Not all VOCs absorb to all sorbents. For instance, ethanol, with a carbon number of two, cannot be (semi-)quantitatively adsorbed onto many sorbent materials. Additional sample preparation may be required depending on the type of sorbent bed combination used. Volatiles with low vapor pressure require sorbent beds with a wide carbon absorption range, such as porous polymers.

**Sample Analysis.** Breath VOCs can be measured in real-time or off-line. The latter encompasses analysis off-site (where measurement was performed in a laboratory environment) or at the point-of-care (measurement in the same environment as the subject but not in real-time). The choice of instrument determined the type of analysis that could be performed. Sample analysis instruments are summarized in Table 5.

Gas chromatography coupled to MS is a popular analysis method throughout studies that use prior volatile extraction but requires that breath gas be preconcentrated prior to analysis. Many desorption methods are used for preconcentrating breath. Solid phase microextraction (SPME) is the most frequently used preconcentration method. However, SPME is highly biased toward hydrophobic VOCs, and use of multi-sorbent tubes, or directly purging breath gas (as possible with evacuated steel containers), is more suitable for most VOCs. Instrument settings varied between research groups, depending on the study design and if VOCs were targeted. Limited versatility and measurement time restricted GC-based methods to off-site analysis in most cases, although point-of-care analysis was possible in one study that used GC coupled with a surface acoustic wave detector. Here, it was reported that a breath sample could be measured within six minutes.

With regard to MS instruments, a range of mass analyzers was used. Studies using nontargeted profiling used time-of-flight mass analyzers. Many studies described the use of a single quadrupole mass analyzer (see Tables 1−3). In addition, studies took advantage of selected ion monitoring mode with triple quadrupole mass analyzers for targeted quantitative MS-MS analysis.

Where MS was used, the mass-to-charge (m/z) acquisition range varied between studies. A broad m/z range was set by some studies thereby reducing the number of scans/second, while others set a smaller range (approximately from m/z 20 to 200) more suited to breath VOCs. To increase sensitivity

---

**Table 5. An Overview of Breath Sample Analysis Methods from Reviewed Studies**

<table>
<thead>
<tr>
<th>Analysis Method</th>
<th>Instrument</th>
<th>Sensitivity</th>
<th>Strengths</th>
<th>Limitations</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Gas Chromatography</strong></td>
<td>GC-MS, GC-TOF-MS, GCxGC-TOF-MS, GC-MS/MS, GC-SAW, GC-FID/TCD</td>
<td>ppb</td>
<td>• Large reference library&lt;br&gt;• Established methods for VOC analysis&lt;br&gt;• High sensitivity</td>
<td>• Pre-concentration required&lt;br&gt;• Relatively high cost&lt;br&gt;• Usually large instruments, and requires regular maintenance by a specialist&lt;br&gt;• Relatively slow analysis (per sample)</td>
</tr>
<tr>
<td><strong>Proton Transfer Reaction</strong></td>
<td>PTR-MS, PTR-TOF-MS</td>
<td>ppb to ppt</td>
<td>• No pre-concentration&lt;br&gt;• High sensitivity&lt;br&gt;• High specificity</td>
<td>• Relies on reaction with reagent ion&lt;br&gt;• Relatively high cost&lt;br&gt;• Usually large instruments, and requires regular maintenance by a specialist</td>
</tr>
<tr>
<td><strong>Selected Ion Flow Tube</strong></td>
<td>SIFT-MS, SIFT-MS/MS</td>
<td>ppb to ppt</td>
<td>• No pre-concentration&lt;br&gt;• High sensitivity&lt;br&gt;• High specificity</td>
<td>• Relatively high cost&lt;br&gt;• Relies on reaction with selected reagent ions</td>
</tr>
<tr>
<td><strong>Ion Mobility-based Spectrometry</strong></td>
<td>MCC-IMS, MCC-DM5</td>
<td>ppm to ppb</td>
<td>• High versatility&lt;br&gt;• On-line clinical analysis possible</td>
<td>• Chemical characterisation required&lt;br&gt;• Maintenance by specialist required</td>
</tr>
<tr>
<td><strong>Electronic Nose Sensors</strong></td>
<td>Cytosense, NA-Nose, Quartz Crystal</td>
<td>ppm to ppb</td>
<td>• Highly portable&lt;br&gt;• On-line clinical analysis possible&lt;br&gt;• Relatively low cost&lt;br&gt;• Highly adaptable to a clinical environment</td>
<td>• Binary classification&lt;br&gt;• Limited chemical selectivity&lt;br&gt;• Needs prior training&lt;br&gt;• Selective sensors are still under development</td>
</tr>
<tr>
<td><strong>Secondary Electrospray Ionisation</strong></td>
<td>SESI-MS, SESI-MS/MS</td>
<td>ppb</td>
<td>• No pre-concentration&lt;br&gt;• No soft ionisation reaction required</td>
<td>• Relatively high cost&lt;br&gt;• Characterisation required&lt;br&gt;• Requires regular maintenance by a specialist</td>
</tr>
</tbody>
</table>

*Parts per billion<br>• Parts per trillion<br>• Parts per million.*
of detection by GC for very volatile compounds, one study used a flame ionization detector,\textsuperscript{58} while another study additionally used a chemical ionization source with MS.\textsuperscript{65}

The mechanism behind PTR-MS and SIFT-MS allows highly sensitive real-time analysis of targeted breath VOCs.\textsuperscript{57,59,60,65} In some circumstances, a lack of suitable sampling inlets or study design meant these instruments were used at the point-of-care but not in real-time.\textsuperscript{53,62,71} Similarly, studies using SESI-MS, where developed for real-time analysis, measured breath VOCs at the point-of-care.

Mass spectrometry-based approaches are known to have high operational costs and lack portability, when compared to IMS-based detectors and e-nose sensors. To this end, some researchers measured breath VOCs using portable multicapillary column IMS instruments. The differences between IMS-based platforms used within studies are described elsewhere.\textsuperscript{23} Researchers were able to perform point-of-care or real-time measurements with IMS-based instruments, although some were proof-of-concept studies with a low sample size.\textsuperscript{33,78}

In comparison, e-nose devices were the most commonly used instrument for real-time analysis of breath VOCs. This is because they are comparatively simple to operate and adaptable depending on study requirements. There are many variations in e-nose device sensors, which included the commercial Cyranose 320,\textsuperscript{34,35,45,34,53,68} coated quartz microbalance sensors,\textsuperscript{50} and a nanomaterial-based sensor array.\textsuperscript{75,77} Consequently, specificity of volatiles and detection sensitivity will differ between sensing mechanisms.

**Quality Assurance and Data Analysis.** Quality assurance procedures are important for small molecule analysis.\textsuperscript{27,93} Compared to serum or urine metabolomics, quality assessment and control methods in breath research are relatively underdeveloped.

Research of sampling standardization has raised some important issues.\textsuperscript{94–97} With regard to breath sample quality, some studies sampled ambient air parallel to breath samples. Theoretically, VOCs present in environmental air samples are exogenous to breath and therefore could be subtracted from breath sample VOCs.

The diverse use and interpretation of reference standards and materials was observed. Analytical standards included use of target analyte standards,\textsuperscript{65,67} permeation tubes,\textsuperscript{60} retention indexing mixtures,\textsuperscript{48} and certified reference mixtures.\textsuperscript{39} However, not all studies described the use of analytical standards, depending on the instrument used. For example, reagent ions used in soft ionization methods were used to measure instrument precision.

Reporting standards have been published for metabolic studies.\textsuperscript{26,22} These standards allow necessary chemical and data analysis recommendations for studies, and these are particularly important in breath VOC research.

As breath VOC samples can have a high number of variables and confounding factors, most studies performed various univariate and multivariate analysis with the acquired VOC data. In most cases, exploratory unsupervised analysis was performed to gain an insight into how the data were correlated. Several unsupervised multivariate methods were employed, such as principal component analysis (PCA), fuzzy c-means, or hierarchical clustering analysis (HCA). To counteract high intersubject variability and reduce the dominant effect of high concentration VOCs, alternative data processing methods have been described.\textsuperscript{39,98}

In addition to unsupervised analysis, some researchers additionally employed supervised learning methods to further assess breath VOC data based on predefined hypotheses. Where data were deemed to have a linear relationship, linear discriminant analysis (LDA) and partial least-squares-discriminant analysis (PLS-DA) were used. Studies using nonlinear methods included k-nearest neighbor, random forest (RF), support vector machines (SVM), and weighted digital analysis (WDA). Further information on the use of these chemometric approaches is described in detail elsewhere.\textsuperscript{99} However, such multivariate methods, which are based on training a computational algorithm, risk data overcorrection, and it is important to avoid false discovery.\textsuperscript{100} Studies reduced the likelihood of statistical bias by robust validation of their statistical methods.

In some studies, metabolite pathway databases were used to evaluate VOCs for possible network linkage with host metabolism.\textsuperscript{36,76} Where targeted analysis of known VOCs was performed, multivariate methods were not used. However, in some nontargeted studies, acquired data was processed without additional statistical or hypothesis-free classification.\textsuperscript{36,41,42} Here, the authors explain that VOC profiles of subjects should be treated on an individual basis, thereby eliminating the effect of data overcorrection and reducing interference associated with confounding factors.\textsuperscript{36,42}

**DISCUSSION**

Breath contains a complex mixture of VOCs comprising many different chemical species from endogenous and exogenous sources. The host immune response to microbial invasion and pathogenesis can result in an enhanced VOC profile, thereby masking VOCs emitted directly from microbes. As perhaps expected, authors have reported the absence of VOCs from in vitro headspace when compared to breath samples,\textsuperscript{37} which is because of biochemical differences due to host–pathogen interactions that are missed in artificial culture. For example, a study found only four out of eight volatiles were aligned between in vitro and in vivo breath samples for A. baumannii.\textsuperscript{37} It has been reported that less than half of a VOC profile may be linked directly to microbial pathogens.\textsuperscript{83} This is evident when aligning bacterial metagenomics and breath VOC analysis.\textsuperscript{101} Additionally, cocultures have shown release of distinctive VOCs in headspace.\textsuperscript{102,103} Notably, P. aeruginosa has been shown to exacerbate the growth of A. fumigatus when cocultured together.\textsuperscript{103} This suggests distinctive VOCs may be produced for each pathogen–pathogen, pathogen–host, and host–disease interaction.

An observation made between some studies was an increase in terpene-based compounds found in breath samples linked to invasive aspergillosis.\textsuperscript{41,42} Bacterial infection may also emit distinctive chemical groups depending on their individual virulence mechanisms. For example, many studies targeting P. aeruginosa reported nitrogen-containing compounds.\textsuperscript{36,60}\textsuperscript{69} Branchedly hydrocarbon VOCs, together with increased concentration of inert gases, were reported across many nontargeted studies and could indicate a host’s inflammatory response to infection rather than the presence of a particular microbial species.\textsuperscript{37,99,86} Breath samples in studies investigating influenza vaccination were not taken from subjects without prior vaccine treatment. Breath VOCs were therefore associated with oxidative stress and inflammatory response caused by the vaccine, which included branched hydrocarbons.\textsuperscript{72,73}
Bean et al. grouped similar mass spectral fragmentation patterns into one VOC identity, to reduce incorrect identification. Rounding exact mass to the nearest integer from high resolution instruments is a useful technique for small molecules, although we would stress that many metabolomics studies use accurate mass to infer molecular formulae of an analyte under analysis.

The ideal m/z range for breath analysis depends on the research hypothesis. Many VOCs characteristic have a mass spectral molecular ion peak below m/z 200, although this molecular weight is somewhat arbitrary. Some volatile and semivolatile chemicals in breath have distinguishable molecular ion peaks above m/z 200 (for example, pentadecane molecular ion peak m/z = 212), consequently reducing accuracy of identification within nontargeted experiments. In addition, the percentage probability of tentative identification may be lowered or less accurate.

Matching VOCs were identified across studies. Hydrogen cyanide was associated with both H. pylori infected subjects and P. aeruginosa infection in CF subjects. Two studies investigating bovine tuberculosis using the same analytical methods, however, found only two matching VOCs. Concentration levels for ethanol and heptane were contradictory between studies investigating respiratory infection within the intensive care unit. Hydrogen cyanide was well characterized in the breath of patients with infection and CF but was not suitable for diagnosis due to low sensitivity. 2-Pentylfuran has been proposed as a biomarker for invasive aspergillosis, although this has not been replicated by other groups in either in vitro headspace experiments or patient breath samples. 2-Pentylfuran was also quantified in foodstuffs by the same researchers. Similarly, 2-aminoacetophenone, a proposed biomarker of P. aeruginosa infection in CF patients, was also found in foodstuffs.

This raises the concern that breath VOC profiles are influenced by the environment or activity before or during sampling. Preference of sampling methods by research groups primarily relied upon the analysis capabilities, the target infection being studied, and the portion of breath collected (and this is reviewed in depth elsewhere). For example, access to intensive care units, site inclusion in under-developed countries, or facilities for animal studies may be dictory between studies investigating respiratory infection within the intensive care unit. This diversity makes breath sampling standardization efforts difficult across the breath research community. For example, if sampling breath from the nose, the presence of upper respiratory pathogens must be considered, for example P. aeruginosa nasal infection often precedes lung infection in cystic fibrosis. Where breath is sampled from a mechanical ventilator, it is important to consider exogenous VOCs from ventilator tubing and inhaled anesthetics contamination during sampling and analysis within the intensive care unit.

Another key concern is the lack of consistency of proposed breath biomarkers that were not found in other studies investigating the same infection. For example, out of three studies investigating lung infection in intensive care patients, only 4/26 potential breath biomarkers were matched between two or more studies (tetradecane, nonanal, ethanol, and heptane). Several breath biomarkers were also found for other diseases; for instance, nonanal was identified in the breath of patients with VAP and in patients with lung cancer from other studies. Another example from within the review is ethanol, which was also found in patients with H. pylori infection and tuberculosis infection. Although we have discussed the potential origin of these VOCs, the quest for finding highly selective breath biomarkers is without doubt very challenging, especially with the lack of defined chemical identification workflows and standardized methodologies in breath research. However, breath analysis alone will not constitute a complete diagnosis for a disease yet may prove useful when combined with other routine diagnostic methods, symptoms, and subject activity and medical history.

Clinical breath samples may include additional VOCs linked to drug treatment, for example, anesthetic agents found in patient breath after recovering from surgery. In addition, when VOC profiles are associated with a particular infection, the influence of noninfectious underlying conditions must also be taken into account. It is clear that sampling humans is very complex as they are exposed to many other factors, both intrinsic and extrinsic, as reviewed in Goodacre et al., which might affect the microbiome and, hence, the ability to predict infection status accurately.

With many studies, the accuracy of breath VOC analysis is relied upon by the positive identification of infection by comparative invasive diagnosis. If false-positive or -negative identification is given by these methods, the overall accuracy of breath VOC analysis may be reduced and hence unreliable. The most widely used statistical validation method was the leave-one-out technique; however, it is often more appropriate to use resampling methods such as bootstrapping. This is especially important as many studies use small sample sizes and statistically processed data can be interpreted in different ways. By way of example, studies have processed data with and without statistical validation and have found differences in overall test accuracy, which suggests that the data generated are perhaps not appropriate for predictive analyses. It is worth stressing that these invasive diagnostic methods are used as the primary reference data and if such data contain errors then these errors will be propagated throughout analysis; the adage “garbage in—garbage out”, coined by Zupan and Gasteger, is particularly relevant here for chemical data calibrated by “gold standard” data that are not so “golden”.

Breath VOC analysis for detecting infection requires further research and development. For cases where VOCs are well-characterized, applications for fast on-site diagnosis are being developed. However, no current VOC biomarkers of infection have been clinically approved as of yet. Increased research output and understanding of VOC metabolic pathways would increase confidence diagnostic approval.

With regard to breath VOC analysis for diagnosis, cost-effectiveness is very important. In order for a test to be cost-effective, a compromise between diagnostic accuracy, the prevalence of an infection, and monetary cost per sample is required. This is especially important when aiming to replace current practices or propose new breath biomarkers. Current low cost e-nose devices perform binary classification between multiple samples, limiting their VOC measurement ability and requiring prior sample training. However, some researchers have argued a binary classification is all that is needed for diagnosis and sensors selective for volatiles are under development.

The use of e-nose and IMS technology within the clinical environment is particularly interesting, as devices can be tailor-made and easy-to-use for clinical staff. In addition, sensor electronics can be mass-produced making integration into routine diagnosis less demanding than other technologies, resulting in low cost infection disease diagnostics.
International multicenter studies remove experimental bias from tentative biomarkers identifications and increase confidence for application in clinical diagnosis. We have noticed that fewer studies were performed in continents generally considered to have higher occurrence of infectious diseases. Future research could include sites in countries with a higher prevalence of a specific infection. Investigators new to breath analysis may find recent technical standards useful; for a review of standardization efforts across sampling and analysis of breath, see ref 28.

The continued interest in breath analysis for noninvasive and fast identification of infection is evident by the exponential rise in research publications over the past decade. With concerns of breath, see ref 28.

CONCLUSIONS

Studies show that diagnosis of infection and infectious diseases using breath VOC biomarkers is possible. External validation is required for reliable measurements, before biomarkers can be tested within clinical protocols. However, sampling and analysis methods are heavily dependent on the targeted disease, surrounding environment, and instruments used, which can affect standardization efforts, and consequently, a "one size fits all" analysis is not yet apparent. Metabolic pathways of VOC production are still not yet fully understood, and this is important as this would enhance infection diagnostics and help understand the complex host–pathogen crosstalk; in addition, microbial infection and associated inflammation may produce distinguishable VOC fingerprints or chemical groups. Increased high-throughput research of breath VOCs associated with infection is needed to further validate biomarkers and to realize the potential of noninvasive diagnostic methods fully.

AUTHOR INFORMATION

Corresponding Author

*E-mail: stephen.fowler@manchester.ac.uk.

ORCID®

Waqar M. Ahmed: 0000-0003-1490-6391
Roisyn Goodacre: 0000-0003-2230-645X

Author Contributions

W.M.A., T.M.N., O.L., R.G., and S.J.F. planned and carried out the systematic search and reviewed the studies. W.M.A. wrote the manuscript.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We would like to thank our funding body (European Commission FP7 Marie Curie Actions), under the Industry-Academia Partnerships and Pathways (IAPP) programme (MC-IAPP BreathDx 611951). We are also grateful for fruitful collaborations within the BreathDx Consortium (University of Manchester, Royal Philips B.V., Academic Medical Centre Amsterdam, and University of Twente), Center for Information Services, Chronic Disease Management Department, (Philips Research B.V.), the Laboratory of Bioanalytical Spectroscopy (University of Manchester), and the Manchester Breathsomics Group (University of Manchester).

REFERENCES

Review

acs infectious diseases


