

Review

Not peer-reviewed version

---

# Real-Time Breath Diagnostics: Linking Molecular Pathways, Measurement Technologies, and Clinical Translation

---

[Velmurugan Thavasi](#)\*, [Nirmal Choradia](#), Naoko Takebe, Neal Naito, [Susan Yeyeodu](#), Peter W. Sadler, [Dean Hougén](#), Sanchith Velmurugan, Jordan P. Metcalf, Donna L. Tyungu, [Thirumalai Venkatesan](#)

Posted Date: 9 April 2026

doi: 10.20944/preprints202604.0672.v1

Keywords: breathomics; real time breath analysis; breath volatilome; volatile organic compounds; exhaled breath condensate; molecular pathways; metabolite biosynthesis; oxidative stress pathways; microbiome host co metabolism; proton transfer reaction mass spectrometry; selected ion flow tube mass spectrometry; ion mobility spectrometry; optical spectroscopy; wearable biosensors; machine learning; clinical validation



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a [Creative Commons CC BY 4.0 license](#), which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Review

# Real-Time Breath Diagnostics: Linking Molecular Pathways, Measurement Technologies, and Clinical Translation

Velmurugan Thavasi <sup>1,\*</sup>, Nirmal Choradia <sup>2</sup>, Naoko Takebe <sup>2</sup>, Neal Naito <sup>3</sup>, Susan Yeyeodu <sup>2</sup>, Peter W. Sadler <sup>4</sup>, Dean Hougen <sup>5</sup>, Sanchith Velmurugan <sup>6</sup>, Jordan P. Metcalf <sup>7</sup>, Donna L. Tyungu <sup>8</sup> and Thirumalai Venkatesan <sup>1,6</sup>

<sup>1</sup> Department of Physics & Astronomy, University of Oklahoma, Norman, OK, 73019, USA

<sup>2</sup> OU Health Stephenson Cancer Center, 800 NE 10th St, Oklahoma City, OK 73104, USA

<sup>3</sup> MVK Pharmaceutical LLC, 31800 N Capitol Ave, Indianapolis, IN 46202, USA

<sup>4</sup> School of Biological Sciences, University of Oklahoma, Norman, OK 73019, USA

<sup>5</sup> School of Computer Science, University of Oklahoma, Norman, OK 73019, USA

<sup>6</sup> School of Electrical and Computer Engineering, University of Oklahoma, Norman, OK 73019, USA

<sup>7</sup> Pulmonary, Critical Care & Sleep Medicine Division, Department of Medicine, University of Oklahoma Health Campus, 800 N. Research Parkway, Oklahoma City, OK 73104, USA

<sup>8</sup> Department of Pediatrics- Infectious Diseases, University of Oklahoma Health Science Center; Pediatric Infectious Diseases Lab, OU Children's Hospital, Oklahoma City, OK, USA

\* Correspondence: velmu@ou.edu

## Abstract

Diagnostic latency limits time-sensitive care and early detection, and exhaled breath provides a rapid, repeatable window into metabolic and inflammatory chemistry. We review real-time breath sampling and analytical technologies and evaluate their readiness for clinical adoption, with emphasis on molecular pathways reflected in the breath volatilome and in exhaled breath condensate. Real-time mass spectrometry enables kinetic VOC profiling and targeted quantification, while humidity-aware sensors and wearable condensate platforms extend monitoring beyond the laboratory. Pathway-anchored interpretation links breath readouts to ketone handling, isoprenoid metabolism, nitric oxide signaling, lipid peroxidation, uremic nitrogen handling, and microbiome-host co-metabolism, but performance remains vulnerable to confounding, drift, and non-representative comparators. Translation requires standardized breath fraction control, traceable features, robust quality systems, and governed device algorithm stacks so that breath outputs change decisions and outcomes.

**Keywords:** breathomics; real time breath analysis; breath volatilome; volatile organic compounds; exhaled breath condensate; molecular pathways; metabolite biosynthesis; oxidative stress pathways; microbiome host co metabolism; proton transfer reaction mass spectrometry; selected ion flow tube mass spectrometry; ion mobility spectrometry; optical spectroscopy; wearable biosensors; machine learning; clinical validation

## 1. Introduction

In clinically urgent conditions, deterioration may unfold more rapidly than routine testing can assess. A patient with heart failure may worsen over the course of a few hours, while laboratory measurements are checked only once daily. Respiratory infections can also progress rapidly during an emergency visit, forcing treatment decisions before confirmatory results are available. At the other end of the spectrum, early malignant disease develops slowly, yet screening is often spaced far enough apart that subtle early changes pass unnoticed. The problem is more than practical. It reflects

a mismatch between the pace of biology and the pace of measurement. The consequences are familiar: delayed treatment, unnecessary hospitalization, and missed opportunities for early intervention. The issue also has a clear equity dimension. When diagnosis depends on centralized laboratories, stable supply chains, and specialized staff, the burden of delay falls most heavily on settings with the fewest resources. Under these conditions, a delayed result is not simply a technical inconvenience; it can shape clinical outcome. For this reason, recent global health initiatives have placed increasing emphasis not only on whether a diagnostic test exists, but on whether it can be performed at the time and place where the treatment decision is most critical [1].

Breath analysis offers a different kind of measurement because breath can be repeatedly sampled with minimal burden to the patient [2–4]. This analysis detects subtle physiological change that can be followed as a trajectory rather than inferred from isolated snapshots. The rationale for this approach is both biological and practical. Volatile organic compounds (VOCs) and small inorganic gases reflect systemic metabolism through blood-to-alveolar exchange, while airway-derived chemistry and the nonvolatile constituents recovered in exhaled breath condensate (EBC) reflect epithelial inflammation, oxidative stress, and processes at the airway surface. In this sense, breath analysis has the potential to narrow the gap between intermittent laboratory testing and continuous monitoring. From a molecular perspective, the breath matrix spans small inorganic gases such as CO<sub>2</sub>, NO, CO, and NH<sub>3</sub>, as well as a broad set of volatile organic compounds whose structures map to core biochemical pathways. Acetone reports ketone body metabolism, isoprene reflects flux through the mevalonate pathway, and diverse carbonyls, sulfur-containing volatiles, and aromatic species can accompany oxidative stress, airway inflammation, microbiome activity, xenobiotic biotransformation, and environmental exposure. Because the same molecular classes can arise from multiple sources, clinically useful breath signatures typically rely on multicomponent patterns and their time structure rather than a single molecule [11].

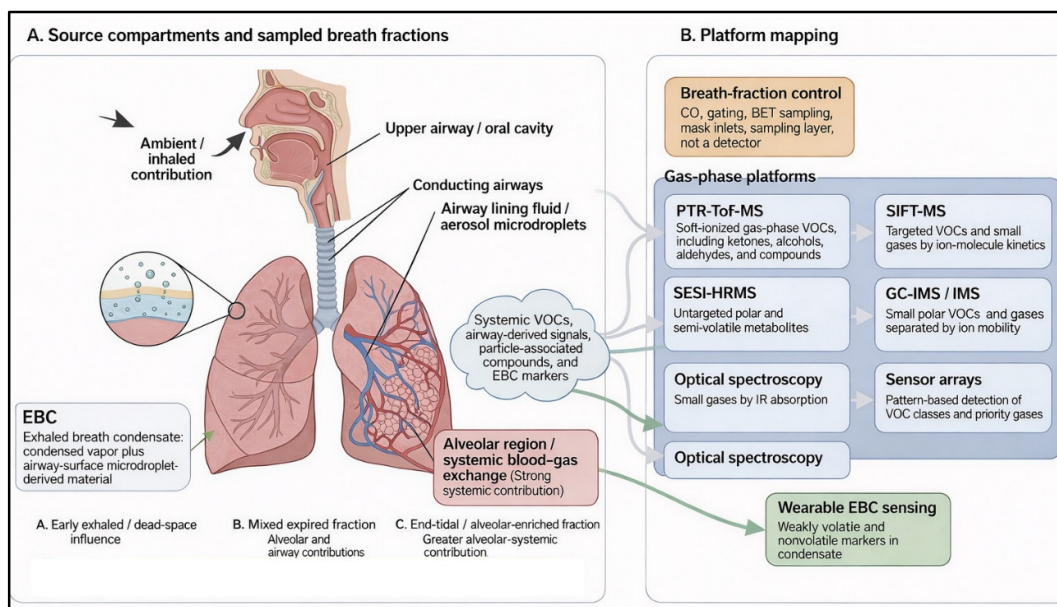
Early limitations in the breath analysis field were related to instrumentation and other factors [5,6]. Breath is a chemically crowded matrix, spanning a wide concentration range and shaped by diet, medication, microbiota, and ambient exposure [5,6]. Humidity, the portion of exhaled breath sampled (for example, mixed expired air versus end-tidal air), and residual effects from the sampling inlet or tubing due to prior samples or ambient exposure can all alter the measured signal. Many early discovery studies relied on case–control separation strategies that proved difficult to sustain once real clinical heterogeneity was introduced. In practice, comorbidity and background exposure are not nuisance factors that can simply be removed. They form part of the setting in which any useful diagnostic test must work. What has improved in recent years is the maturity of the analytical platforms and the supporting evidence base. Real-time methods can now measure breath chemistry on physiologically relevant timescales. Sampling can be stabilized by engineered interfaces, and modern inference pipelines can convert high-dimensional chemical data into outputs that fit more naturally into clinical workflow [7,8]. As a result, the field has moved away from presenting breath as a direct substitute for blood tests, swabs, or imaging. Its more credible role is as a repeated temporal layer of measurement, capable of tracking transitions, treatment responses, and early deviations that more intermittent testing may miss [9].

The shift from episodic testing to repeated or near-continuous measurement brings additional demands that are not captured by analytical performance alone. A wearable device or a routine breath-sampling system does not simply generate a chemical readout. Over time, it may also produce a detailed record of physiology and health behavior. Depending on how such data are handled, they may reveal aspects of a person's environment, daily habits, or future disease risk. For that reason, privacy protection, subgroup performance reporting, and data governance need to be built into validation from the outset rather than treated as issues to be addressed later. Recent international guidance on the ethics and governance of artificial intelligence (AI) in health points in the same direction [10]. This review is therefore organized around the question of how breath diagnostics can be used in practice rather than around instrumentation alone. We first examine the biological origins of breath signals and the main sources of noise and variability. We then consider the principal

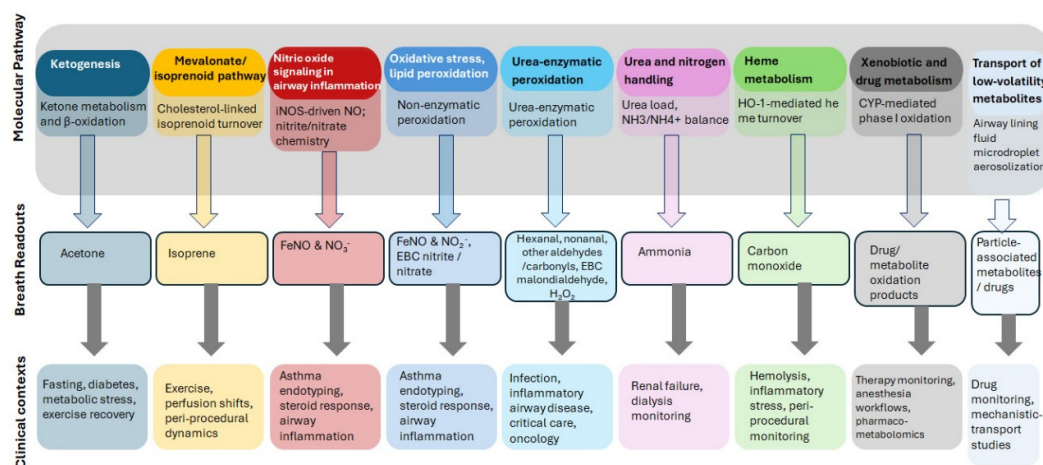
platform families, from reference-grade mass spectrometry to wearable sensors, and discuss validation pathways in relation to real clinical decisions and outcomes. Throughout, the emphasis is on the conditions under which breath measurements can be relied upon in routine care, particularly when sampling is repeated, low-burden, and performed outside a conventional clinical laboratory.

## 2. The Breath Metabolome and Volatilome: Molecular Sources, Compartments, and Confounding

VOCs from exhaled breath are more numerous and diverse than those from blood or urine [11]. Breath provides a composite chemical signal that is assembled from several biological and environmental sources that converge at the mouth or nose during exhalation. Some VOCs are derived from chemical reactions that have been exchanged from blood into alveolar air. Other VOCs are produced locally in the airways by epithelial metabolism, immune activity, oxidative reactions, and microbial processes. In addition, environmental VOCs are inhaled or ingested compounds, which can enter the breath directly or after partial metabolism. Exhalation also carries particles and microdroplets generated by the airway lining fluid. When these are condensed into exhaled breath condensate (EBC), the specimen becomes an aqueous matrix that contains ions, small metabolites, redox-active species, lipids, proteins, and nucleic acids. All of these can provide clinically useful information, but each is also open to dilution, contamination, and variability in breathing patterns and sampling geometry. The main contributing compartments and their corresponding content are shown in Scheme 1. To make the molecular origin of these breath signals explicit, Scheme 2 summarizes representative biochemical pathways and processes that generate common breath readouts and indicates the clinical contexts in which they are most often interpreted.



**Scheme 1.** Biological sources of breath signals and corresponding measurement platforms. The scheme summarizes the main source compartments contributing to exhaled breath, the sampled breath fractions, and the principal platform classes used to measure gas-phase VOCs, small inorganic gases, particle-associated compounds, and exhaled breath condensate (EBC) markers. Breath-fraction control is shown as a sampling layer rather than a detector.



**Scheme 2.** Representative molecular pathways contributing to exhaled breath signatures, linking key biochemical processes to measurable breath readouts and relevant clinical contexts.

### 2.1. Systemic Blood-to-Alveolar Exchange and Partitioning Physiology

Circulating metabolites enter exhaled breath through the blood to the alveolar exchange and by partitioning into the airway lining fluid. Whether a compound is detectable in breath is governed by its volatility, solubility in blood and airway surface liquid, blood-gas partition coefficient, ventilation-perfusion matching, and the fraction of breath sampled. Compounds with low blood solubility can equilibrate rapidly with alveolar air and often track the end tidal plateau, whereas more water-soluble species can be buffered by airway liquid and become sensitive to breathing pattern and recent wash-in and wash-out history. Regional recruitment and ventilation perfusion heterogeneity further shape the sampled fraction, so breath concentration reflects production, distribution, and exchange rather than circulating abundance alone [9,12,13].

This exchange link also supports a molecular reading of several canonical breath VOCs. Breath acetone is closely tied to hepatic ketone chemistry. During increased fatty acid beta-oxidation, acetyl-CoA accumulates and is routed into ketogenesis. HMG CoA synthase and HMG CoA lyase generate acetoacetate, and its decarboxylation produces acetone that partitions into blood and can be cleared through the lungs. Isoprene is commonly linked to the mevalonate pathway that supplies cholesterol and other isoprenoids, and it is often interpreted as arising from turnover of isoprenoid intermediates such as dimethylallyl diphosphate. Because exchange depends on perfusion and ventilation, isoprene and related small hydrocarbons can vary with circulation and respiratory dynamics as well as with biosynthesis. The circulating metabolite pool is also shaped by the gut microbiome. Microbial end products and microbially modified metabolites can enter the circulation and later appear in exhaled breath. Part of the breath signal, therefore, reflects host microbial co-metabolism rather than host metabolism alone, especially for pathways linked to fermentation, sulfur turnover, nitrogen handling, and aromatic compound degradation. Clinical and dietary studies have shown strong associations between breath volatiles and gut microbiome composition and function, and dietary intervention can shift breath VOC profiles in reproducible ways [14,15]. More recent mechanistic work has strengthened this view by showing that the microbiota shapes the breath volatilome across species [16]. The measured spectrum is also conditioned by the analytical front end. Humid breath promotes ion clustering and surface losses, and inlet memory can selectively distort reactive carbonyls and sulfur species unless materials and calibration are selected for humid matrices [9,12,13].

## 2.2. Airway Surface Chemistry and Exhaled Breath Condensate

Airway-derived signals represent a biological layer distinct from systemic exchange. The chemically active airway surface contains epithelial cells, immune cells, mucus, and resident microbes that reshape local redox balance and nitrogen chemistry. A well-defined example is exhaled nitric oxide. In inflammatory airway disease, cytokine-driven induction of inducible nitric oxide synthase in airway epithelium increases conversion of L-arginine to nitric oxide and L-citrulline, thereby raising FeNO during eosinophilic inflammation and altering its response to corticosteroid therapy [17,18]. Nitric oxide is also oxidized to nitrite and nitrate, and these anions can be recovered in exhaled breath condensate, linking a gas-phase signal to an aqueous proxy of airway lining fluid. Oxidative stress adds a second mechanistic axis. Reactive oxygen species generated by activated inflammatory cells and stressed epithelium can initiate lipid peroxidation in membrane polyunsaturated fatty acids. Breakdown of lipid hydroperoxides yields aldehydes and related carbonyl products, including malondialdehyde and 4-hydroxy-2-nonenal, alongside smaller aldehydes that are sufficiently volatile to appear in the gas phase. Because many of these products are reactive and short-lived, their detectability depends on the sampled fraction, the timing of collection, and the extent to which inlet chemistry and humidity alter recovery.

Exhaled breath condensate expands the accessible chemistry beyond VOCs by condensing microdroplets from airway lining fluid into an aqueous matrix. The matrix can contain ions, metabolites, redox markers, proteins, nucleic acids, and other airway surface constituents, but its translational value depends on rigorous control of dilution, oral carryover, and analyte stability. These challenges were formalized early and remain central to interpretation and cross-study comparability [19–21]. Methodological frameworks and technical standards emphasized the need to specify the collection device, temperature control, breathing maneuver, and normalization strategy, since device-dependent recovery and saliva contamination can dominate apparent biological differences if they are not measured and corrected [13,22,23]. EBC is therefore best positioned as a protocol-sensitive but high-yield matrix for longitudinal monitoring and phenotype stratification, where within-person trajectories can be interpreted under consistent collection conditions.

## 2.3. Particle Phase Transport of Low Volatility Metabolites and Nonvolatile Compounds

Breath is often discussed as a volatilome, but exhalation also carries suspended particles and microdroplets generated from airway lining fluid. These particles form when small airways and alveolar regions undergo transient closure and reopening, which creates thin liquid films that can rupture and aerosolize. The chemistry of these droplets is not random. Airway lining fluid is enriched in pulmonary surfactants, a lipoprotein-like mixture composed primarily of phospholipids and cholesterol, with associated surfactant proteins. Because surfactant components are amphiphilic, they stabilize the air-liquid interface and can also solubilize and concentrate hydrophobic metabolites, oxidized lipids, and drug molecules within a lipid-rich microenvironment. In parallel, the aqueous fraction of droplets carries dissolved salts, small polar metabolites, and peptides. This creates a transport route in which molecules with low intrinsic volatility can still appear in breath because they are exported in the particle phase rather than transferred via blood-gas equilibrium. Chen et al. [24] demonstrated this mechanism with unusual clarity by tracking venlafaxine after systemic administration in mice. Despite being predominantly ionized and nonvolatile, venlafaxine was detected in exhaled breath in association with particles, and its time course followed the blood pharmacokinetic profile. This observation matters beyond a single drug example. It shows that the boundary between gas-phase VOC analysis and airway fluid-adjacent sampling is less strict than it appears, and that some circulating compounds can reach the exhaled stream through aerosolization of airway surface material.

Particle phase readouts also introduce a distinct confounder class. Particle number and size distribution depend on breathing mechanics, airway hydration, and the closure and reopening behavior of small airways, all of which can shift with disease state and with the sampling maneuver itself. For that reason, particle-associated biomarkers should be interpreted together with respiratory

pattern metadata and, where feasible, direct particle metrics. Without this context, changes in droplet generation or sampling geometry can be misread as biochemical change, particularly for analytes that preferentially partition into surfactant-rich droplets.

#### 2.4. Molecular Confounders and the Value of Dynamic Measurements

Breath is chemically informative because it is produced at the boundary between internal metabolism and the external environment. The same property makes confounding difficult to avoid. Volatile organic compounds present in room air can be inhaled and later reappear in exhaled breath, producing a wash-in and wash-out behavior that can resemble endogenous generation. This is particularly relevant for common indoor volatiles and solvents, where the breath signal can be shaped by recent exposure history as much as by physiology. Diet and oral chemistry add another strong layer of variation. Ingestion studies show that everyday exposures can introduce dozens of reproducible features into the breath profile within minutes, even when those features are not related to disease biology unless the exposure itself is measured and accounted for [25]. Oral processes can also reshape breath composition through rapid chemical and microbial transformations at the mouth and upper airway surface, which can dominate parts of the spectrum if sampling timing and breath fraction are not controlled.

Smoking adds confounding through two mechanisms that are hard to separate in routine cohorts. First, cigarette smoke contributes direct exogenous volatiles and reactive carbonyls that can persist in the airway and sampling pathway. Second, chronic smoking alters airway biology through sustained inflammation and structural remodeling, which changes endogenous breath chemistry even after recent smoking is excluded. Variation is also present in healthy individuals. Stable person-specific breath profiles and time-of-day effects can explain a substantial share of the measured variance and may obscure early disease-related changes when cohorts are sampled under different schedules or routines [26]. These observations are a reminder that confounding is not an occasional nuisance in breath studies. It is a baseline condition that must be measured and modeled. Microbiome-related variation should be treated in the same way. It is not simply background noise, because it can reflect the genuine host and host microbial biology [14–16]. A clear molecular example is the production of short-chain fatty acids by gut bacteria. Acetate, propionate, and butyrate arise from fermentation of dietary substrates through conserved bacterial pathways, including acetate formation from acetyl CoA, propionate formation through succinate or propanediol routes, and butyrate formation through butyryl CoA intermediates. The underlying capacity is encoded at the gene level in the microbiome through pathway-specific enzymes, and differences in pathway abundance can shift the circulating pool of these metabolites and related derivatives that may be detected in breath under appropriate analytical conditions. In practical terms, this means that diet-driven shifts in microbiome metabolism can create structured and repeatable breath changes that are biologically real, yet clinically misleading if the diagnostic question is unrelated to nutrition or gut microbial activity.

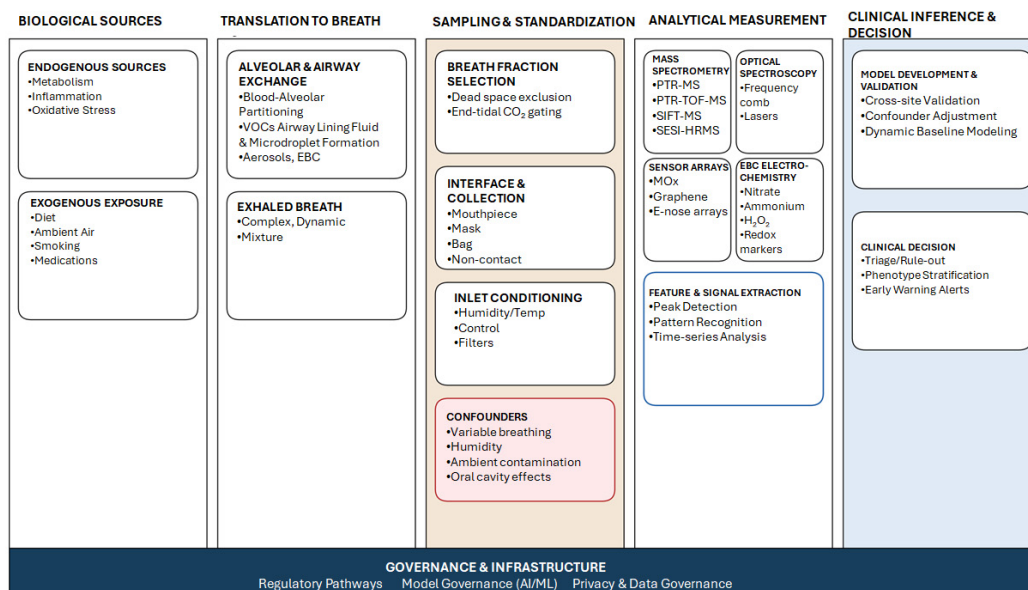
For translation, the key question is whether a model remains reliable when this degree of real-world variability is present [6]. Repeated measurement offers one defensible way to address the problem. Within-person trajectories can be more stable and more informative than population cut-offs, particularly when the goal is early warning, treatment adjustment, or recovery monitoring. This aligns with the growing interest in VOC dynamics, where kinetic information helps distinguish endogenous production from recent exposure and can reveal responses to intervention that are missed by single-time point sampling [9]. At the same time, disease specificity is rarely achieved at the level of one molecule, because many breath signals converge on shared biochemical processes such as oxidative stress, lipid peroxidation, and microbiome-linked metabolism. What is most likely to prove useful in practice is a multivariate biochemical pattern that remains interpretable in the presence of confounding, reproduces across sites, and is stable enough to support clinical action. Seen in this way, the analytical platform is not only a technical choice. It determines which part of chemical space is captured and whether that information can be stabilized sufficiently for routine clinical use.

### 2.5. Molecular Classes and Biochemical Origins of Breath Analytes

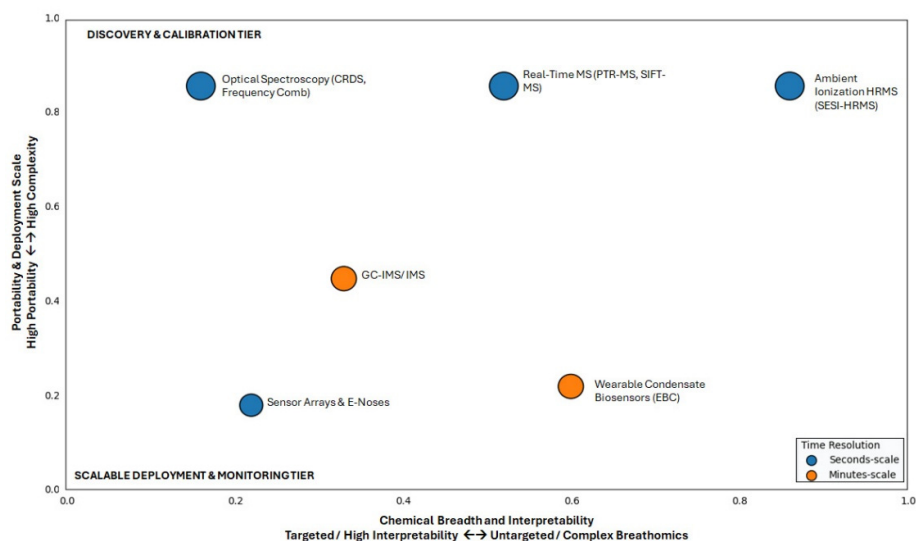
From a molecular perspective, breath biomarkers are best interpreted as chemical products of a limited set of biological processes rather than as disease labels. Most candidates fall into recurring chemical families, and each family points to a plausible origin in intermediary metabolism, redox chemistry, microbial transformation, or xenobiotic handling. This framing helps explain why single-molecule claims often fail outside curated cohorts, while multicomponent patterns and kinetics can remain informative under realistic heterogeneity [9,11]. Ketones, such as acetone, and related alcohols formed via reversible redox chemistry can signal shifts in hepatic energy balance and systemic NADH availability. Isoprene is an abundant endogenous hydrocarbon that is linked to isoprenoid biosynthesis and can vary with metabolic state and ventilation. Carbonyls, including aldehydes and ketones, often increase with oxidative and inflammatory stress because lipid peroxidation and secondary oxidation pathways generate volatile fragments, a mechanism that has motivated real-time monitoring of simple hydrocarbons in acute settings [76]. Nitrogen-containing species span nitric oxide and its downstream chemistry in the airway, as well as ammonia linked to urea handling and microbial urease activity. Sulfur-containing volatiles reflect methionine and cysteine turnover and microbial metabolism, and they can be strongly shaped by oral and airway microbiota. Many high-value airway markers are not volatile. Exhaled breath condensate provides access to ions and redox-active molecules that are difficult to measure in gas-phase panels, including nitrite, nitrate, ammonium, hydrogen peroxide, lipid peroxidation products, and selected proteins. The translational challenge is that these species are present in a dilute, protocol-sensitive matrix, so collection efficiency, normalization strategy, and stability controls become part of the molecular assay definition [17]. These molecular origins map naturally onto the technology stack summarized in Scheme 1. Proton transfer and selected ion flow tube mass spectrometry emphasize oxygenated VOCs and other proton-affinity-driven volatiles, optical spectroscopy anchors small inorganic gases through calibrated absorption physics, and ambient ionization, high-resolution workflows extend coverage toward polar and semi-volatile metabolites. Sensor arrays and wearable condensate systems can then be tuned toward priority chemical families, if humidity, drift, and quality control are engineered into the measurement chain.

## 3. Real Time Platforms for Breath Molecular Measurements

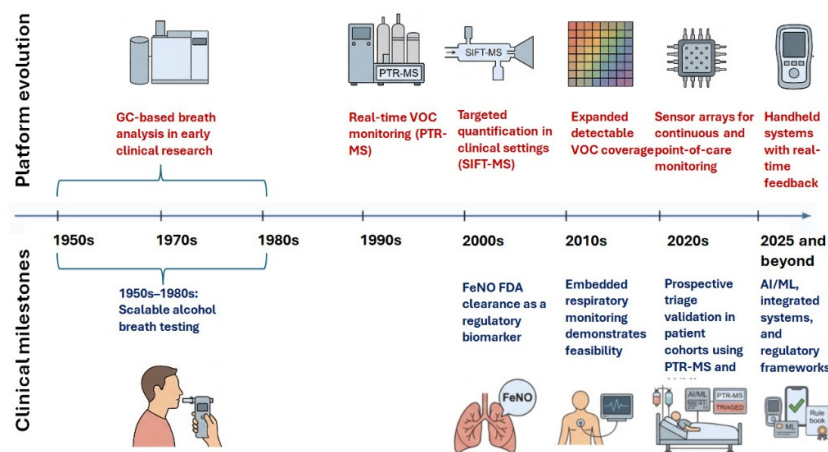
Real-time breath analysis is not built around a single technology. It consists of a group of analytical approaches, each with its own balance of chemical coverage, speed, cost, portability, and ease of interpretation. In practice, this has led to a layered technical structure. Reference-grade instruments are most useful at the discovery stage, where the aim is to identify candidate markers, assign chemistry where possible, and monitor calibration and drift. Smaller or less expensive systems become more relevant once the question shifts toward routine monitoring, where rapid turnaround, simple operation, and minimal disruption to workflow are often more important than complete molecular attribution. Benchmarking studies provide an important point of comparison across these platforms. Standardized exposure and washout protocols, such as peppermint challenge designs, are particularly useful because they can be repeated across instruments and across sites [27–29]. Figures 1–3 summarize the overall translational pathway, the current technological landscape, and the main stages in the development of breath-analysis platforms. Table 1 summarizes recurrent breath analytes and analyte classes together with their likely origin, main clinical context, suitable platform types, and the main factors that complicate interpretation.



**Figure 1.** The journey of a breath biomarker from source to clinical decision. Schematic traces endogenous production and exogenous exposure through blood-to-alveolar exchange and airway-surface chemistry, followed by sampling fractionation, inlet conditioning, analytical measurement, feature extraction, and clinical inference layers. It emphasizes where confounders enter, where standardization acts, and where dynamic monitoring converts variability into interpretable trajectories.



**Figure 2.** Technological ecosystem of real-time breath analysis. Platforms are positioned by chemical breadth and interpretability on the x-axis and portability and deployment scale on the y-axis. The landscape separates into a discovery and calibration tier and a scalable monitoring tier, highlighting the need for hybrid workflows that link reference platforms with portable devices. Abbreviations: CRDS, cavity ring-down spectroscopy; EBC, exhaled breath condensate; GC-IMS, gas chromatography–ion mobility spectrometry; HRMS, high-resolution mass spectrometry; MS, mass spectrometry; PTR, proton transfer reaction; SESI, secondary electrospray ionization; SIFT, selected ion flow tube.



**Figure 3.** Evolution of breath-analysis platforms and milestones toward clinical adoption. The upper timeline summarizes major developments in analytical platforms, from early gas chromatography to real-time mass spectrometry, targeted ion–molecule methods, broader VOC coverage, sensor arrays, and emerging handheld systems. The lower timeline highlights corresponding milestones in clinical uptake, including alcohol breath testing, FeNO clearance, embedded respiratory monitoring, prospective triage validation, and the current move toward AI-assisted, workflow-integrated, and regulation-aware breath diagnostics.

**Table 1.** Representative breath analytes and analyte classes detectable in real-time platforms: chemical class, dominant origin, clinical contexts, platforms, and key translational considerations.

Analyte or Analyte Class	Chemical Class	Dominant Origin or Pathway	Most Credible Clinical Contexts	Best-Fit Real-Time Platform(s)	Key Translational Caution or Interpretation
Acetone	Ketone	Systemic fatty-acid oxidation and ketone handling	Metabolic monitoring, exercise physiology, diabetes, ketoacidosis	PTR-ToF-MS, SIFT-MS, chemiresistive sensor arrays	Strong real-time signal and deployable-sensor relevance, but strongly influenced by fasting status, diet, and exertion
Isoprene	Hydrocarbon terpene	Systemic isoprenoid metabolism with strong perfusion and ventilation dependence	Exercise physiology, perfusion-linked monitoring, peri-procedural dynamics	PTR-ToF-MS, SIFT-MS	Best interpreted as a dynamic physiology signal rather than a disease-specific biomarker

Ammonia	Inorganic gas	Systemic urea and nitrogen handling, with oral contribution	Renal failure, dialysis adequacy, uremia, serial metabolic monitoring	IMS or GC-IMS, SIFT-MS, CRDS, targeted sensors	Among the most physiologically interpretable breath markers, but oral chemistry, humidity, and ambient contamination must be controlled
Nitric oxide (FeNO)	Inorganic gas	Airway epithelial iNOS and type 2 inflammatory signaling	Asthma phenotyping, steroid response monitoring, selected chronic airway disease	Targeted FeNO analyzers, chemiluminescence, electrochemical or optical platforms	Already clinically standardized, but reflects inflammatory endotype more than disease label
Aldehydes and related carbonyls (e.g., hexanal, nonanal)	Aldehydes and carbonyls	Airway and systemic lipid peroxidation and oxidative stress	Oxidative injury, inflammatory phenotyping, adjunct infection or oncology panels	PTR-ToF-MS with careful inlet control, SESI-HRMS, orthogonal GC-MS confirmation	High information content, but reactive species are vulnerable to inlet memory, humidity effects, and surface losses
Short-chain fatty acids (acetate, propionate, butyrate)	Carboxylic acids	Gut microbial fermentation with systemic appearance in breath	Gut-lung axis studies, diet-microbiome response, metabolic inflammation	SRI-ToF-MS or PTR-based MS with TD-GC-MS confirmation	Mechanistically attractive, but strongly diet-responsive and usually requires structural confirmation
Nitrite and hydrogen	Nitrosative and	Airway lining fluid	Longitudinal airway	EBC microfluidic and	Useful for repeated

peroxide in EBC	oxidative airway- surface markers	and condensate- phase inflammator y chemistry	inflammatio n and oxidative- stress monitoring, wearable EBC workflows	electrochemical sensing platforms	airway monitoring, but dilution normalization, salivary contamination, and collection consistency are critical
--------------------	--	---	---	--------------------------------------	--

### 3.1. Real-Time Mass Spectrometry: Broad Chemical Coverage with Fast Kinetics

PTR-MS and PTR-ToF-MS enable tracking of breath chemistry on the timescale of physiological changes [30,31]. Their analytical advantage is soft chemical ionization with limited fragmentation, which enables continuous monitoring of chemically distinct VOC families during rest, exercise, feeding, and acute illness [32]. The measured spectrum includes abundant low-molecular-weight metabolites and lower-abundance compounds linked to oxidative stress, inflammation, and host-microbial co-metabolism. Representative pathway-linked examples are summarized in Table 1. Acetone is often interpreted in relation to beta oxidation and ketone handling, whereas isoprene is associated with mevalonate-derived isoprenoid turnover. Aldehydes and related carbonyls can increase during lipid peroxidation, and sulfur volatiles and short-chain fatty acids are frequently shaped by diet and airway or gut microbial activity. In most settings, these signals are most useful as readouts of pathway activity rather than as disease-specific markers. PTR-based instruments, therefore, sample a mixture whose components differ in origin, transport behavior, and time course. In perturbation studies, kinetics often carry more information than a single concentration snapshot. Many VOCs change over minutes, so features such as onset, peak timing, and washout can reflect production, distribution, and clearance more directly than an isolated measurement [33]. Repeatability improved once sampling shifted from poorly defined exhaled air to controlled breath fractions. Buffered end-tidal sampling reduces dead-space variation, while higher-resolution time-of-flight designs reduce mass overlap and improve feature separation, allowing more consistent detection across sessions [12,34,35]. These engineering steps help ensure that measured variation reflects biology rather than differences in breathing maneuvers or collection methods. Prospective respiratory cohorts further show that PTR-ToF-MS breath prints can retain useful performance under heterogeneous clinical conditions, particularly when evaluation uses realistic differentials rather than healthy controls alone [8,36]. A persistent limitation of real-time MS is quantitative integrity for reactive and surface-active VOC classes.

Some of the compounds most informative about inflammation and oxidative stress are also the hardest to measure accurately. Aldehydes, certain sulfur-containing species, and other reactive VOCs often carry the most mechanistically relevant information, but they are also particularly sensitive to inlet memory, humidity, and surface losses. If the front end is not carefully designed and calibrated under humid breath-like conditions, the profile may shift toward more stable surrogate compounds that are easier to detect but less closely linked to the underlying biology [37]. This is one reason layered analytical strategies are often preferred when pathway traceability is important [38,39]. In such workflows, real-time mass spectrometry provides kinetics and throughput, while a smaller group of higher-value features is checked periodically by an orthogonal method. The confirmation of short-chain fatty acids by TD-GC-MS provides a good example of this approach [40]. SIFT-MS complements PTR-based methods by offering direct quantification based on established ion-molecule reaction kinetics [41,42]. A practical advantage is that it can be used with clinical gas lines and ventilator circuits, which makes breath monitoring easier to incorporate into routine patient care rather than treating it as a separate procedure [43]. At the discovery end, ambient-ionization high-

resolution platforms such as SESI-HRMS extend coverage toward more polar and semi-volatile metabolites [44]. The limitation is that the data are highly context-dependent. Ingestion effects and diurnal variation can easily dominate the signal unless they are directly measured and accounted for [25,26]. In asthma, some of the most convincing HRMS studies have therefore been those built around intervention, where breath chemistry was used to distinguish response phenotypes and metabotypes rather than to carry the greater burden of cross-sectional disease classification [45]. Real-time mass spectrometry remains the primary reference platform in breath diagnostics because it can identify promising breath markers, detect key confounding factors early, and track signal changes over time. That information can then be used to build smaller, simpler systems for routine clinical use, including compact separation platforms and sensor-based devices.

### 3.2. Ion Mobility Platforms: Portable Detection with Constrained Attribution

Ion mobility spectrometry, including GC-IMS, has found a practical niche in breath analysis. Compared with full mass spectrometry, these systems are smaller, run faster, and are easier to operate, which is why they remain attractive for point-of-care use and for repeated measurements in clinic-based studies. These systems are useful because they separate ionized breath compounds according to how quickly they move through gas. Rather than trying to identify many compounds simultaneously, they simplify a complex breath mixture into a smaller set of measurable signals. For focused clinical applications, such as ammonia monitoring in renal disease, that level of information is often sufficient. However, handling humidity is a technical challenge. Breath is a wet matrix, and water alters the ion population within the IMS cell by influencing clustering and the reactant-ion balance. Peak position and peak intensity can therefore drift even when the underlying breath composition has not changed. This is one example of why humidity cannot be treated as a minor operating variable. Ion mobility spectrometry (IMS) and gas chromatography coupled to IMS (GC-IMS) sit in the deployment-friendly zone because they can be compact, relatively fast, and less demanding to operate than full mass spectrometry. Their main limitation is that the mobility signature depends strongly on matrix conditions, which are difficult to maintain constant during routine breath sampling. Water vapor shifts the reactant ion balance and ion clustering, which can move peak position and intensity and reduce the reliability of library matching and multivariate feature spaces unless humidity is measured and actively managed. This matters in practice because exhaled breath is saturated with water, and deviations from the calibration envelope can translate into drift in peak position or quantitative recovery. Deployable IMS workflows therefore need an engineered humidity strategy, combined with routine quality control checks that flag when spectra have moved outside the calibration domain. These platforms are most persuasive when the analyte space is deliberately constrained, and the sampling envelope can be standardized, for example, in dialysis-associated ammonia monitoring or in structured screening sessions [46].

### 3.3. Optical Spectroscopy: Precise Quantification for Selected Analytes

Optical spectroscopy has generally been developed as a targeted and metrologic approach to breath analysis rather than as a discovery platform [47,48]. Photoacoustic implementations have been described in parallel for a range of selected analytes [49,50]. Its main advantage lies in the mid-infrared region, where many breath constituents exhibit strong, highly specific absorption bands. Under these conditions, concentration can be directly derived from light absorption, with less ambiguity than is typically encountered in pattern-based methods. Frequency-comb and other cavity-enhanced systems can detect multiple gases in real time with high sensitivity, making them particularly suitable for applications that require both specificity and calibration stability. This is especially relevant for continuous measurements of CO<sub>2</sub>, CO, NO, ammonia, and selected small VOCs with strong infrared signatures [7]. In this role, optical methods are used less to search broadly for new biomarkers and more to provide traceable measurements of compounds whose physiological interpretation is already reasonably well established. A further strength of optical methods is their ability to resolve isotopologues. This extends breath analysis from measuring concentration alone to

examining origin and turnover. Real-time measurements of  $^{13}\text{C}$ - and  $^{18}\text{O}$ -labeled  $\text{CO}_2$  with per-mil precision have already been demonstrated at minute-scale integration, with stable tracking across the alveolar plateau and good preservation of isotope fidelity [51]. This type of measurement is well-suited to tracer studies and metabolic phenotyping, and it may also serve as an external reference for validating or recalibrating broader breathomics workflows. The limitations of optical spectroscopy arise from the same feature that gives it strength. These systems only detect compounds that are addressed by the chosen wavelengths. They do not provide broad chemical coverage by default. In addition, water vapor and  $\text{CO}_2$  dominate substantial parts of the infrared spectrum, so routine breath measurements often depend on careful control of humidity, temperature, and optical-path stability. For these reasons, optical spectroscopy is best regarded as a reference method for targeted biomarkers and physiology-based tests, and as a useful calibration layer within broader breath diagnostic systems, rather than as a stand-alone solution for open-ended VOC discovery. Nevertheless, photoacoustic approaches are beginning to show disease-specific potential in obstructive airway disease [52].

### 3.4. Sensor Arrays and Contactless Transducers

Sensor-based approaches are appealing for breath analysis because they can be small, inexpensive, and low-power, features that make repeated sampling outside specialist laboratories more realistic. The difficulty lies in the sample itself. Exhaled breath is warm, saturated with water, and chemically complex, while most sensor materials respond to classes of compounds rather than to a single analyte. As a result, the output is usually a composite pattern influenced by chemistry, humidity, flow, and the sensor's recent history. Progress in this area has come less from the discovery of one ideal sensing material than from treating the device as a full measurement system, including the sensing layer, the package, the calibration strategy, and the method used to interpret the signal. This is illustrated by the development of array-based systems with scalable electronics. Complementary metal-oxide semiconductor (CMOS)-integrated graphene and metal-oxide arrays have demonstrated rapid responses, along with practical approaches to drift management and interference correction, both of which are essential if such systems are to move beyond proof-of-concept studies [53]. Terbium-doped  $\text{SnO}_2$  yolk-shell chemiresistors were designed to reduce the effect of humidity while preserving sensitivity to targets such as acetone, which remains of interest in metabolic monitoring [54]. Device format matters as well. A Pd- $\text{WO}_3$  acetone sensor incorporated into a disposable mask provides a good example of how chemical sensing can be adapted for repeated use in a simple, acceptable form for the wearer [55]. Other transduction approaches widen the range of possible designs when low-burden operation is the main goal. A magnetoelastic contactless sensor, for example, has been used to distinguish breath from ambient air while simultaneously tracking humidity and selected gases, suggesting a simpler route for wearable monitoring when exact molecular identification is not essential [56]. Flexible ammonia sensors based on tungsten-doped vanadium dioxide and carbon composites provide a related example, particularly for renal and environmental monitoring, where a wearable format may be advantageous [57]. It is also worth noting that breath-pattern sensing is not merely an accessory to chemical detection. In many cases, it provides the context needed to interpret chemistry. QCM-based humidity sensors, smart textiles, and optical-fiber systems can all detect breathing patterns and pauses over short time scales [58–60]. They may therefore help define when a signal was generated and under what conditions breathe. These developments suggest a practical role for sensor-based systems. They are unlikely to replace platforms that offer clear molecular attribution, but they become much more viable when humidity and drift are handled explicitly and when the chemical signal is interpreted alongside breathing metadata and calibration procedures that help preserve stable inference across users and environments.

### 3.5. Wearable EBC Sensing

Mask-based condensers combined with simple microfluidic handling are beginning to change how exhaled breath condensate is collected [61]. Instead of being limited to slow, supervised sampling in research settings, EBC can increasingly be obtained repeatedly during routine activity. This is important because EBC is more than condensed breath water. It contains airway-lining fluid microdroplets and dissolved compounds, thereby providing access to ions and other low-volatility markers that are only weakly represented in gas-phase VOC profiles. In practical terms, EBC extends breath analysis toward airway-surface chemistry, including markers of inflammation and oxidative stress that are often better suited to longitudinal monitoring than to one-off diagnosis. The main obstacles at this stage are less about whether such biomarkers exist and more about how they are collected and measured. Condensate volume is influenced by breathing patterns, humidity, and cooling efficiency, so dilution and normalization need to be addressed explicitly if changes in collection are not to be mistaken for changes in biology. Oral carryover and salivary contamination remain persistent concerns, particularly for analytes that are abundant in the mouth. Temperature control is also important, since some targets may degrade, volatilize, or shift equilibrium during collection and storage. Electrochemical approaches are attractive in this setting because they enable the direct measurement of selected redox and lipid-peroxidation markers in a humid matrix [62,63]. Their clinical usefulness, however, will depend on whether the full platform can provide stable within-person trajectories, supported by defensible normalization and built-in quality checks, rather than isolated measurements that are difficult to compare from one session to the next.

### 3.6. Platform Comparison and Translational Readiness

When these platforms are considered side by side, the main translational issue is not which performs best in abstract analytical terms, but which can produce a result that remains stable, interpretable, and useful in the setting in which it is intended to be used. In practice, the different platform families make different compromises. Some offer broader chemical coverage, others offer greater portability, faster turnaround, or simpler calibration. For this reason, platform comparison is less a matter of ranking technologies than of asking which approach is best suited to a particular clinical purpose. As the field develops, the emphasis is beginning to shift away from analytical performance considered in isolation and toward clinical usefulness. In most situations, the goal is not to catalog every detectable molecule in breath, but to obtain a readout that can support a defined clinical action. Large COVID-19 studies have been valuable in this respect because they tested workflow integration and performance under heterogeneous conditions. At the same time, they do not show that the field has reached the same level of maturity across other diseases or across routine seasonal settings. The next stage will therefore depend on broader prospective studies across a wider range of indications, together with benchmarking strategies that are not tied to a single disease wave or sampling context. Standardized exposure and washout protocols are useful in this regard, since they can be repeated across centers and used to examine the robustness of sampling, calibration, and inference. Table 2 compares the major real-time breath platform families based on analytical scope, translational constraints, and likely clinical role.

**Table 2.** Comparative analysis of real-time breath analytical platforms: performance metrics and clinical fit.

Platform family	Primary targets	Typical time resolution	Deployment setting	Main translational limitation	Best fit clinical use	Key examples
PTR-MS and ToF variants (PTR-MS,	Broad VOC patterns; oxygenated	0.1 to 1 s	Reference grade instrument	Cost and complexity; structural	Discovery and calibration	[8,12,32,36]

PTR-ToF-MS, CI-ToF-MS)	and some reactive VOCs with inlet control		s in clinical studies	assignment often needs confirmation; humidity and inlet effects can bias reactive VOCs	tier; high throughput screening where kinetics add value	
Sampling and breath fraction control interfaces (BET, CO <sub>2</sub> gating, noncontact, mask inlets)	Breath fraction definition (end tidal vs mixed); maneuver quality metrics	Breath by breath	Cross-cutting pre-analytical layer	Adds hardware and protocol discipline; noncontact designs increase dilution and ambient sensitivity	Essential enabler for cross-site comparability and model robustness	[12,61,93]
SIFT-MS	Targeted VOCs and small gases with direct quantification	Seconds	Bedside capable in structured environments	Limited chemical breadth; depends on predefined analytes and sampling line handling	Operating room and ICU monitoring; targeted panels requiring absolute values	[43]
SESI-HRMS and related ambient ionization HRMS	Untargeted polar and semi-volatile metabolites; high-dimensional fingerprints	Seconds	Reference grade discovery	Context sensitivity and identification burden; strong leakage risk without rigorous design	Mechanistic discovery; response phenotyping; individualized baseline studies	[24–26,45]
IMS and GC-IMS	Constrained VOC panels; small polar	Tens of seconds to minutes	Point-of-care oriented	Humidity and matrix sensitivity; drift and library	Targeted screening and monitoring in structured workflows	[46]

	gases such as ammonia			matching constraints		
Optical spectroscopy (frequency comb, CRDS, photoacoustic, mid IR)	Selected small gases and isotopologues, especially CO <sub>2</sub> ; tracer-based physiology	Seconds to minutes	Specialized reference layer	Limited analyte set tied to absorption lines; photonic complexity	Reference quantification; isotope physiology; calibration anchor for hybrid systems	[7,46,51]
Chemiresistive sensor arrays and electronic noses (MOX, graphene MOX, doped oxides)	Priority gases and pattern signatures (acetone, ammonia, mixed VOC surrogates)	Seconds	Wearable or low-cost edge	Drift, humidity interference, and cross-sensitivity; interpretability limitations	High frequency monitoring and screening adjuncts where usability dominates	[53–55,57]
Resonant and breath pattern transducers (QCM, magnetoelastic, fiber, and textile sensors)	Humidity and breathing pattern; adjunct gating and adherence	Sub-second to seconds	Wearables	Limited chemical specificity without added functional layers	Breath maneuver QC, adherence and physiologic context for chemical sensing	[56,58–60]
EBC microfluidics with electrochemical biosensing (mask platforms)	Nonvolatile ions and selected inflammatory or redox markers	Minutes	Wearable monitoring	Dilution normalization; oral contamination control; fouling and carryover	Longitudinal airway inflammation and oxidative stress monitoring	[61–63]
Breath aerosol and particle-phase	Particle-borne and microdroplet-associated	Seconds	Research and mechanistic studies	Strong dependence on particle generation and	Mechanistic transport studies and pharmacokin	[24]

analysis (online)	analytes, including nonvolatile species			sampling geometry	etic applications	
----------------------	--	--	--	----------------------	----------------------	--

#### 4. Clinical Translation: From Molecular Signature to Diagnosis

In breath diagnostics, the main challenge is usually not detecting a chemical signal but interpreting that signal in a real clinical setting. A breath profile may be statistically distinct and still have limited practical value if it is not linked to an actionable clinical insight. For this reason, the intended use of breath diagnostics in translational studies needs to be defined at the outset, whether the aim is screening, triage and rule-out, treatment adjustment, early warning, or surveillance after therapy. The intended use then defines the choice of comparator groups, and the study can be judged in relation to the consequences of false-positive and false-negative results within the test's workflow. Many breath signals reflect processes that are biologically important but not unique to a single disease, including airway inflammation, oxidative stress, microbiome-related change, and broader shifts in systemic metabolism. Such signals may still be clinically useful, but they do not necessarily provide specificity on their own. This is particularly relevant in higher-stakes applications such as oncology, where breath is likely to be most useful when interpreted alongside other clinical data. In practice, this means combining breath features with routine clinical variables, physiologic measurements, or imaging findings so that the result aligns more closely with how decisions are already made in the clinic. Table 3 summarizes representative validation studies in which sampling strategy, cohort heterogeneity, and evaluation design were aligned with defined clinical use.

**Table 3.** Key validation and translational studies across disease areas and platform types.

Disease area/use-case	Platform & modality	Cohort & setting	Key performance/findings	Translational note	Study
<b>ACUTE INFECTIOUS DISEASE TRIAGE</b>					
COVID-19 screening	PTR-TOF-MS breathomics + machine learning + metadata	173 participants (67 COVID-19 positive)	Sensitivity 98%, specificity 74%, PPV 72% NPV 98% AUC 0.961	Supports rapid front-end screening in prospective hospital triage; metadata integration improved performance	[8]
COVID-19 biomarker	PTR-TOF-MS + machine	1137 participants; independent	Accuracy overall; 81.2% overall; 97.3% in	Overlap with influenza-	[36]

identification and screening	learning framework	test cohort n=340	participants >55 years	like breathprints + ion-level limitations highlight need for realistic differentials	
Ventilator-associated lower respiratory tract infection	Untargeted GC-MS breath analysis	108 invasively ventilated ICU patients with suspected VA-LRTI	Sensitivity 98% at specificity 49%; NPV 96%	Strong rule-out potential in ventilated ICU workflows; supports breath as an adjunct to antimicrobial decision-making in suspected VA-LRTI	[65]
<b>RESPIRATORY DISEASE &amp; AIRWAY BIOCHEMISTRY</b>					
Continuous airway & metabolic monitoring	Mask-based EBC harvesting + electrochemical biosensors	Controls + COPD/asthma/post-COVID groups	EBC nitrite vs FeNO r=0.795; EBC ammonium vs serum urea r=0.846	Minute-scale workflow; validated under real-world activity	[61]
Pharmacometabolomics; bronchodilator responsiveness	SESI-HRMS breath analysis	34 pediatric asthma patients; 38 visits	333 features altered post-salbutamol; metabolotypes linked to poor responsiveness	Demonstrates response phenotyping embedded in outpatient workflow	[45]
<b>METABOLIC PHYSIOLOGY &amp; GUT-LINKED METABOLITES</b>					

Exercise stress testing (CPET)	SESI-HRMS	13 healthy adults (7 female)	33 metabolites change across intensity: pathway-level shifts (e.g., TCA/tryptophan)	Supports dynamic cardiometabolic phenotyping & rehab monitoring	[73]
SCFA monitoring; gut-linked volatiles	SRI-ToF-MS + TD-GC-MS confirmation	Method characterization + breath demonstration (humid/dry)	SCFA repeatability <15% (acetic/propionic/butyric); structural confirmation by TD-GC-MS	Establishes analytical credibility; underscores orthogonal confirmation + standardized sampling	[40]
<b>PERIPROCEDURAL &amp; CRITICAL CARE MONITORING</b>					
Perioperative monitoring (anesthesia/ventilation)	Online SIFT-MS integrated into the ventilation circuit	5 anesthetized patients	Propofol + acetone/isoprene/water vapor; rapid responses to operative events	Demonstrates feasibility integrated into existing clinical infrastructure	[43]
<b>RENAL &amp; UREMIC METABOLISM MONITORING</b>					
Dialysis adequacy; uremic metabolism	IMS + cavity ring-down spectroscopy	20 hemodialysis + 20 controls; serial sampling	Breath ammonia decreases during dialysis; correlates with BUN and Kt/V	Clear monitoring use case with interpretable physiology and actionability	[46]
<b>ONCOLOGY (EARLY-STAGE EXEMPLAR)</b>					
In vivo validation in thoracic	Cloud-connected	364 adults with clinical and/or	Validation ROC-AUC 0.83;	Multicenter prospective	[89]

oncology outpatients	eNose analysis of exhaled breath	radiological suspicion of lung cancer from 2 thoracic oncology outpatient clinics; external validation cohort n = 121	sensitivity 94%; specificity 63%; PPV 79%; NPV 89%; performance consistent across tumour characteristics, disease stage, centers, and clinical characteristics	external validation in a clinically realistic referral population, aligned with intended- use deployment	
<b>CROSS-CUTTING ENABLERS</b>					
Healthy baseline variability mapping	SESI-HRMS + bioinformat ics curation	31 participants	227 features uniquely identify 28/31; 37 signals associated with time of day	Justifies individualiz ed baselines + diurnal control in models	[26]

#### 4.1. Infectious Disease Triage and Syndromic Screening

Infectious disease triage is one setting in which delay in diagnosis matters most. Decisions on isolation, bed placement, treatment, and escalation of care are often taken before laboratory confirmation is available. Breath sampling is attractive in this context because it is quick, non-invasive, repeatable, and does not compete with the same laboratory capacity as swabs or blood tests. The biological rationale is that acute infection, along with the host response, can alter the levels and patterns of VOCs and small gases over short periods [64]. At the molecular level, infection reshapes breath chemistry through two coupled sources. The first is host immunometabolism. Pattern recognition signaling rapidly shifts substrate use and redox balance, and activated neutrophils and macrophages generate reactive oxygen and nitrogen species. These oxidants initiate lipid peroxidation in membranes and lipoproteins, yielding volatile carbonyls such as aldehydes that can rise alongside inflammation. The second source is microbial metabolism and microbe-driven remodeling of the airway surface, which can introduce nitrogen and sulfur-containing volatiles and alter the balance of small organic acids. These mechanisms help explain why breath signals can change within hours, but they also clarify the specificity problem in syndromic illness. Many of the same oxidative and inflammatory pathways are shared by viral, bacterial, asthma exacerbation, and chronic lung disease flare. This is one reason breath performs best as an early probability signal that accelerates decisions, rather than as a standalone molecular diagnosis [64]. Breath is therefore being explored not as a replacement for nucleic-acid testing, but as a way to provide an early, rapid signal in diagnosis and treatment. The studies that have shown the most promise have generally used breath as a front-end screening tool. In that role, the result can contribute to rapid rule-out, help identify who should proceed to confirmatory testing, and support prioritization when the pre-test probability is uncertain. Liangou et al. [36] reported that PTR-ToF-MS models maintained useful screening performance across a large multisite dataset despite substantial clinical heterogeneity, although the stronger performance observed in older subgroups is probably better interpreted in relation to the study population than as a general benchmark. Roquencourt et al. [8] also showed that rapid breath analysis could be incorporated into a prospective hospital triage workflow. Breath-based approaches

have also been explored in ventilator-associated lower respiratory tract infections, where an untargeted analysis demonstrated feasibility in a critical care setting [65]. At the same time, these studies point to an obvious limitation. In syndromic respiratory illness, high sensitivity is only part of the problem. Specificity is constantly challenged by common alternatives, such as influenza-like illness, bacterial pneumonia, and exacerbations of chronic lung disease, as well as by changes in prevalence from one season to the next. For this reason, breath triage is likely to be most useful when it is linked to a clear downstream action. A probabilistic result may be used to trigger confirmatory testing, to support earlier de-escalation of isolation in lower-risk patients with a negative signal, or to prompt more directed work-up when the signal is positive in a higher-risk setting. In the end, the value of breath testing in this area is better judged by what it changes in the workflow than by discrimination metrics alone.

#### 4.2. Respiratory Disease, Inflammation, and Treatment Response Phenotyping

Respiratory disease is an obvious area in which breath analysis might be expected to perform well, given the proximity of the signal to the organ of interest. At the same time, it is also an area marked by substantial clinical heterogeneity [66,67]. Symptoms such as cough and dyspnea are shared by asthma, COPD, post-infectious syndromes, heart failure, and anxiety, so detection of a single compound is rarely specific enough to establish a diagnosis. For this reason, breath is more likely to be useful when it is used to describe an inflammatory state or response to treatment rather than to classify disease as simply present or absent [68–70]. The strongest molecular anchor in airway breath testing is nitric oxide biology. In type 2 inflammation, cytokine signaling upregulates inducible nitric oxide synthase in the airway epithelium, increasing nitric oxide production and raising FeNO. Because nitric oxide chemistry is directly connected to inflammatory signaling and steroid responsiveness, FeNO has become clinically interpretable when sampling is standardized [17,18]. That same pathway is also evident in condensate chemistry, as nitric oxide is converted to nitrite and nitrate in airway lining fluid, and these ions can be recovered in exhaled breath condensate under controlled collection conditions [19–23]. In parallel, oxidative pathways contribute additional molecular classes. Peroxide-generating reactions and lipid peroxidation can yield redox-active markers in condensate and reactive carbonyls in the gas phase. This explains why aldehydes, peroxide-related species, and malondialdehyde-type signals often recur in airway inflammation studies, even though they rarely map to one diagnosis on their own [63]. In that setting, the relevant comparison is often not between patients and healthy controls, but between different physiological states within the same individual following a defined intervention. This is where real-time breathomics may offer a practical advantage. By capturing coordinated changes across multiple metabolites, it can reflect shifts in airway metabolism, oxidative stress, and energetics. In pediatric asthma, SESI-HRMS identified hundreds of features that changed after salbutamol administration and also distinguished metabolites associated with poor bronchodilator responsiveness [45]. This intervention anchored design is molecularly attractive because it reduces reliance on cross-sectional disease labels and instead tests whether breath tracks defined biochemical shifts. Bronchodilator response changes ventilation distribution and airway caliber within minutes, and that can immediately alter the exchange kinetics of highly volatile metabolites, while anti-inflammatory treatment is expected to modify epithelial redox state and inflammatory mediator production over longer windows. When the comparison is within the same patient before and after a known perturbation, breath features can be interpreted as pathway-responsive markers rather than as static disease fingerprints [45]. This is of interest because it relates directly to a clinical question that already shapes management, namely, whether a patient is likely to respond to bronchodilator treatment and how far anti-inflammatory therapy may need to be intensified. Wearable and mask-based EBC systems extend this approach beyond the clinic by making repeated sampling more feasible in daily life, where within-person change becomes the main source of information. A smart-mask platform, for example, enabled minute-scale EBC collection and measured airway and metabolic markers that correspond with established clinical measures in cohorts with asthma, COPD, and post-COVID

conditions [61]. The next question is whether such measurements can improve outcomes in practice [71]. That includes whether they can identify exacerbations before symptoms become obvious, reduce admissions through earlier intervention, or support more targeted prescribing by helping distinguish inflammatory relapse from noninflammatory dyspnea.

#### 4.3. Metabolic Physiology, Exercise, and Gut-Linked Metabolites

Metabolic applications of breath analysis are moving away from single-marker claims and toward a more physiological reading of the breath profile as it changes over time. This is important because several of the compounds measured in breath are closely tied to metabolic state. The biochemistry behind the most common metabolic breath volatiles is well defined. Acetone reflects ketone handling, since hepatic ketogenesis produces acetoacetate and  $\beta$ -hydroxybutyrate, and acetoacetate can yield acetone that partitions efficiently into alveolar air. Its breath level, therefore, integrates fatty acid oxidation rate, hepatic output, and peripheral utilization, and it is expected to shift with fasting, exercise intensity, and insulin availability rather than with a single disease label. Isoprene is often linked to the mevalonate pathway and cholesterol-related isoprenoid metabolism, but in breath, it is also strongly shaped by perfusion and ventilation dynamics, which is why exercise and recovery experiments show marked time structure. Real-time platforms add molecular value here because they can track pathway-linked kinetics rather than isolated concentrations, including detection of metabolites associated with central carbon metabolism during graded exertion [72,73]. Acetone is commonly linked to fat oxidation and ketone metabolism, while isoprene has been associated with aspects of cholesterol metabolism and muscle-related physiology. Other volatile signals change rapidly during exertion as perfusion and ventilation shift. In this setting, timing matters. A single concentration measured at a single time point may be less informative than tracking how the breath signal changes over time. This broader view is supported by studies showing that tricarboxylic-acid-related metabolites and other intermediates can be detected in exhaled breath when sampling is fast enough to follow exercise and recovery [72]. Exercise studies have also shown that multiple metabolites change with increasing intensity, consistent with a coordinated metabolic response rather than the behavior of a single biomarker [73]. One possible route toward wider use is the combination of portable sensors with models trained against physiological comparators. Breath analyzers coupled to deep-learning approaches have already been used to estimate exercise-related fat burning from oral and alveolar breath features, suggesting that deployable systems may provide useful metabolic readouts when the validation strategy is sufficiently rigorous [74]. A further area of interest is the contribution of gut-derived metabolites. Gut-linked volatiles can also be described in terms of explicit pathways. Short-chain fatty acids arise from microbial fermentation of dietary substrates through acetyl-CoA and propionyl-CoA-centered routes, and a fraction of these products enters the circulation and can be detected in breath when measurement sensitivity and sampling control are sufficient [40]. The same principle applies to microbially modified aromatic compounds and sulfur-containing metabolites, which can act as readouts of host microbial co-metabolism rather than host metabolism alone [14–16]. This molecular framing helps with interpretation, because it makes clear why diet, fiber intake, and antibiotic exposure can shift breath profiles in reproducible ways that are biologically real but not necessarily disease specific [40]. Microbial metabolism generates VOCs that can enter the circulation and later appear in breath. Short-chain fatty acids are an obvious example, and their measurement in breath, supported by orthogonal confirmation, provides a useful analytical starting point for this line of work [40]. The main difficulty is interpretation. These compounds are often strongly influenced by diet and recent exposure. As a result, a signal may be analytically reproducible but clinically uninformative. Progress in this area is therefore likely to depend on study designs that explicitly record and model diet and exposure, rather than treating them as secondary sources of noise.

#### 4.4. Peri-Procedural and Critical Care Monitoring

High acuity settings offer a direct route to adoption because the sampling interface can be stabilized by existing infrastructure such as ventilator circuits and anesthesia workflows. When flow and timing are controlled, breath readouts can be interpreted as continuous physiology rather than as isolated snapshots. SIFT-MS integrated into ventilation lines has demonstrated real time tracking of multiple exhaled compounds in perioperative care, showing that breath measurement can be embedded into complex environments without adding behavioral burden [43]. Oxidative injury is a second mechanistically anchored target. During cardiac surgery, ischemia and reperfusion can drive lipid peroxidation and generate volatile end products such as ethylene, which has been monitored in real time, provided that circuit effects and ambient contamination are explicitly managed [76]. Other rapid approaches underline the same monitoring-first logic by prioritizing timeliness and workflow integration over exhaustive molecular attribution [75].

#### 4.5. Renal and Uremic Metabolism Monitoring

Breath ammonia is one of the more physiologically interpretable markers in breath analysis because it is closely related to urea metabolism and nitrogen balance [77,78]. In patients with renal failure, rising blood urea levels and changes in acid-base status increase the availability of ammonia and ammonium, which in turn can be reflected in exhaled breath. This has made breath ammonia an attractive candidate for serial monitoring during dialysis. In hemodialysis studies using ion mobility and optical methods, breath ammonia was higher at baseline in dialysis patients, fell during treatment, and correlated with blood urea nitrogen and adequacy measures such as Kt/V [46]. These observations are important because they link the breath signal to a physiological process that is already familiar in renal care. They also show that the marker can be measured repeatedly without burdening the patient and that the results can be related to a clear clinical question. For these reasons, renal monitoring provides a useful test case for the development of sampling interfaces, calibration procedures, and longitudinal analytical approaches that may later be applied in less constrained disease settings, where specificity is more difficult to achieve.

#### 4.6. Oncology: Specificity, Clinical Realism, and Multimodal Integration

Cancer has long been regarded as one of the most promising areas for breath diagnostics, yet it is also among the most difficult to establish specificity in. The problem is not simply whether altered VOC patterns can be detected, but what those patterns actually reflect. Oxidative stress, inflammation, treatment effects, smoking exposure, and microbiome-related changes can all contribute to the breath profile of a patient being assessed for cancer. None of these processes is unique to malignancy, and all are common in the at-risk populations in which cancer is most likely to develop. For that reason, confounding does not arise as a separate problem after the fact. It is part of the clinical pathway itself. A given breath signal may therefore reflect tumor biology, the effects of therapy, or background comorbidity. Metabolic reprogramming is a plausible contributor to some signatures. Altered central carbon metabolism and mitochondrial function can shift ketone handling and redox balance, which in turn can influence acetone and carbonyl-related readouts. Changes in lipid turnover and in isoprenoid synthesis may also contribute to aldehyde profiles and isoprene dynamics. However, these pathway-level changes overlap with those of infection and chronic inflammatory disease, making specificity a design problem rather than an analytical one. This is why claims of disease specificity depend so heavily on careful comparator design and rigorous validation. The same concern has appeared repeatedly in the literature, as summarized in a meta-analysis of VOC-based breath tests, which identified ongoing methodological weaknesses, including variation in sampling methods and a lack of multicenter validation [79]. More recent reviews have described encouraging pooled performance but also point to substantial heterogeneity, the repeated observation of nonspecific VOCs, and the limited value of classical case-control designs in clinically realistic settings [80]. These studies suggest that the key issue in oncology is not whether a signal can

be found, but whether that signal remains informative when tested against the differential diagnoses encountered in routine practice.

Markar et al. [81], Amal et al. [82], and Hong et al. [83] each moved oncology breath research away from simple case–control separation and toward more clearly defined diagnostic questions. Steenhuis et al. [84] applied the same line of thinking in colorectal cancer follow-up. In lung cancer, Steenhuis et al. [85] reviewed the electronic-nose literature and reported promising pooled sensitivity and specificity. Zhao et al. [86], in turn, examined malignant pleural mesothelioma against a comparator group of asbestos-exposed but asymptomatic individuals. Other studies have compared suspected lung cancer with COPD and related chronic inflammatory lung diseases [87–89]. These examples make the central point clear. The clinically relevant comparisons are not cancer versus health, but malignant versus benign nodules in smokers, cancer versus chronic inflammatory lung disease, and recurrence versus post-treatment inflammation. Each of these situations poses a different problem and therefore requires a different balance between sensitivity, specificity, and tolerance of false-positive results. Breath is also more likely to be accepted in oncology when it is used as part of a broader evidence framework rather than as a stand-alone signal. This is already reflected in work combining breath metabolites with imaging and routine clinical variables, where multi-omics fusion improved discrimination in the assessment of breast lesions [90]. Interest is also increasing in more complex AI models designed to learn hierarchical breath representations [91]. Their promise, however, should be interpreted cautiously. Their clinical value will depend on locked development, strict external holdout sets, and comparator groups that represent real diagnostic uncertainty. Without these safeguards, such models may reflect clinical site environmental features, sampling routines, or prevalence differences rather than disease biology. For that reason, the most credible translational route in oncology is likely to involve AI-supported data fusion, multicenter external validation, transparent subgroup reporting, and study designs anchored to meaningful clinical outcomes. *In vitro* headspace studies remain useful, but mainly as a source of mechanistic hypotheses. VOC mixtures released from tissue may separate under controlled conditions, yet this cannot replace *in vivo* validation in clinical populations [92]. More than any other application area, oncology demonstrates why breath analysis should be incorporated into a broader diagnostic framework rather than presented as a definite answer in isolation.

## 5. Assay Standardization, Metrology, and Model Robustness

For a breath-based test to be clinically useful, several conditions must be met from the outset. Sampling must be controlled, confounding factors must be measured explicitly, the analytical chain must be traceable, and the resulting models must remain reliable when used outside the study in which they were developed. The field is no longer limited by a lack of measurable chemistry. A number of platforms can now generate rich, time-resolved readouts and, under controlled conditions, distinguish between clinical groups. The more difficult question is whether the same result remains stable when the method is used across different patients, rooms, seasons, devices, and clinical workflows. That is why reproducibility has become central to translation. In other words, breath analysis is shaped less by analytical sensitivity alone than by factors that occur before and around the measurement, including pre-analytical handling, control of confounding factors, and the way the model is built and maintained. The attraction of breath analysis lies in how closely it reflects physiology and environment. That is also what makes it vulnerable. The measured signal depends on how the subject breathes, which fraction of the breath is sampled, the recent history of the inlet, and what the subject has inhaled or ingested shortly beforehand.

### 5.1. Standardized Sampling and Breath-Fraction Control

Breath fraction is a key parameter and should not be treated as a minor technical point [13]. Dead-space air, end-tidal air, airway aerosol, and condensate microdroplets do not represent the same material, either chemically or biologically, and the end result changes accordingly. Even so, terms such as late expiratory, end-exhaled, and alveolar are still used inconsistently in the literature,

which makes comparison between studies more difficult than it should be. For this reason, fraction control needs to be planned deliberately and described in enough detail for the protocol to be reproduced elsewhere. The sampling method should explain exactly how the patient breathes during collection, including whether the sample is taken during normal breathing or a deep breath, whether a breath hold is used, and whether exhalation is performed against resistance. It should also make clear how the relevant phase of exhalation is selected, whether by CO<sub>2</sub> gating, flow or pressure triggers, or buffered end-tidal capture. The interface used for collection also matters. Mouthpieces, masks, and non-contact arrangements each distinctly affect dilution and contamination. Acceptance criteria are therefore essential. These usually include the quality of the CO<sub>2</sub> plateau, checks for leaks, and a minimum exhaled volume. When these conditions are not met, the sample is better discarded than corrected afterward, since post hoc correction can shift the analysis away from biology and toward artifact. Breath sampling can remain easy for patients and practical in routine care, but it can only be standardized if the collection process is properly instrumented [27–29]. Non-contact and mask-based systems reduce the amount of coaching required, yet they are also more vulnerable to dilution and to interference from room air. Under those circumstances, paired room-air sampling and breath-cycle metadata become part of the specimen rather than optional background information [93]. The same applies when particle-associated biomarkers are being studied. In that setting, the patient's breathing pattern and the sampling setup should both be recorded, because the number of particles in exhaled breath can be influenced by airway behavior and by how the sample is collected [24]. Put simply, if two centers cannot reproduce the same breathing pattern under the same instructions and acceptance criteria, they are not performing the same assay, even if they use the same detector and the same model.

### 5.2. Confounding Variables as Measured Assay Parameters

Several of the factors discussed in Section 2.4 are better regarded as part of the assay than as background noise to be filtered out. Ambient exposure, recent diet, oral chemistry, smoking, and time of day can all shape the measured breath profile [5,6,25,26]. For that reason, it is usually better to record these variables explicitly and account for them in the analysis than to deal with them only by excluding participants later. In practical terms, this starts with the sampling itself. Breathing mechanics should be documented, including flow, CO<sub>2</sub> profile, confirmation of end-tidal sampling, and the quality of the breathing action during sampling. The surrounding environment also needs attention. Where possible, room air should be sampled alongside breath, and relevant details such as ventilation conditions or obvious exposure events should be noted. The same applies to the participant's immediate physiological state. Food intake, smoking or vaping, medication timing, recent activity, stress, and pain may all influence the resulting profile and should be considered during both study design and interpretation. The choice of comparator group is part of the same problem. Control groups should reflect the clinical setting in which the test is meant to be used, rather than relying mainly on healthy volunteers. This becomes particularly evident in larger screening studies, where breath-based classification performs well only when differences in sampling conditions, participant background, and site-specific practice are built into the assay instead of being treated as an afterthought [36]. Confounding, then, is not just a statistical nuisance. It is part of the circumstances under which breath is produced, collected, and interpreted, and it needs to be measured on those terms.

### 5.3. Metrology, Biomarker Identification, and Quality Control

Analytical capability alone is not enough to establish a clinically useful breath test. A breath-based system also requires a measurement framework that defines what is being measured, how calibration is performed, how drift is monitored, and how results are kept comparable across instruments and over time [38,39,42]. In practice, three recurring metrology issues determine whether a reported breath signature can be relied upon beyond a single study. The first concerns reactive and surface-active VOCs. These compounds are particularly susceptible to distortion during sampling

and analysis. Inlet materials, humidity, ion chemistry, and memory effects can all alter their apparent abundance and, in doing so, change the feature set presented to the model. This matters because aldehydes and sulfur-containing compounds are often linked to oxidative stress and inflammation, yet these same pathways are not specific to a single disease state, particularly in oncology. Studies that have improved inlet design and calibrated measurements under humid breath-like conditions have shown that quantitative performance can be strengthened, but they also make a broader point. Reactive VOCs should be treated as a metrology problem before they are treated as a modelling problem [37]. The second issue concerns untargeted breath signatures. These broad signal patterns are useful in discovery studies, but they are less reliable as clinical endpoints unless the key contributing features are stable, well-defined, and reproducible. Many untargeted studies report mass-to-charge features, putative formulas, or multivariate pattern vectors. This is acceptable at an early stage, when the aim is to identify candidate panels. It becomes less satisfactory when those same features are intended to support clinical decisions. At that stage, stable feature definitions are required, along with internal references where feasible and periodic orthogonal confirmation of features that are expected to carry diagnostic or monitoring value. The short-chain fatty acid work provides a useful example of this type of confirmation strategy [40]. The third issue is quality control. In breath diagnostics, quality control cannot be added at the end of the workflow. It has to be built into both the device and the analytical process from the outset. This includes blanks, ambient references, drift sentinels, and performance-control mixtures, together with acceptance limits that can stop reporting when measurements fall outside the calibrated range. Without these safeguards, models may fail quietly, particularly in distributed settings where operators, environments, and routines differ from site to site. For this reason, the translational pathway is likely to remain tiered. Untargeted fingerprinting may continue to play an important role in discovery, but clinical implementation will depend on features that are calibrated, traceable, and interoperable whenever the underlying chemistry allows it.

#### 5.4. Bioinformatics, Overfitting, and Model Governance

Breathomics is particularly prone to overfitting, meaning a model may appear accurate on the development dataset but perform poorly when tested in new patients or in different clinical settings. Breathomics datasets often contain many measured features, while the number of patients is relatively small. Because the signal also changes with time and context, this creates a high risk of unstable models and overfitting. Under these conditions, AI modeling does not necessarily define disease. It may instead end up defining where and how the samples were collected. Clinic-specific room air, inlet history, cartridge batches, operator technique, and small differences in breathing maneuvers can all become embedded in the data, especially when cases and controls are not collected under the same workflow or within the same period. The risk of data leakage is practical and immediate, because models may otherwise learn features of the collection setting rather than the underlying biology. Even among healthy participants, breath profiles can retain stable personal and contextual structure. In one SESI-HRMS dataset, 227 features were sufficient to identify 28 of 31 participants, while 37 features varied with time of day [26]. If a spectrum contains enough information to indicate who was sampled and when, it may also contain enough information to reveal where and how the sample was obtained, unless the study design actively prevents this. This is particularly important in oncology, where effect sizes may be modest and confounding is common. In that setting, controlling leakage should be treated as a design requirement no less important than calibration of the assay itself. Instrumental drift and batch effects add a further layer of difficulty. They can change apparent discrimination without reflecting any true biological difference. Reported AUROC values have been shown to vary by approximately 0.1–0.2, depending on the preprocessing strategy and the choice of drift or batch correction applied to the same dataset [94]. In breathomics, the meaning of batch extends beyond a single analytical run. It may also reflect room-air composition, seasonal exposure patterns, inlet conditioning, maintenance schedules, staffing changes, or shifts in protocol. If these boundaries are ignored during validation, a model may separate the collection

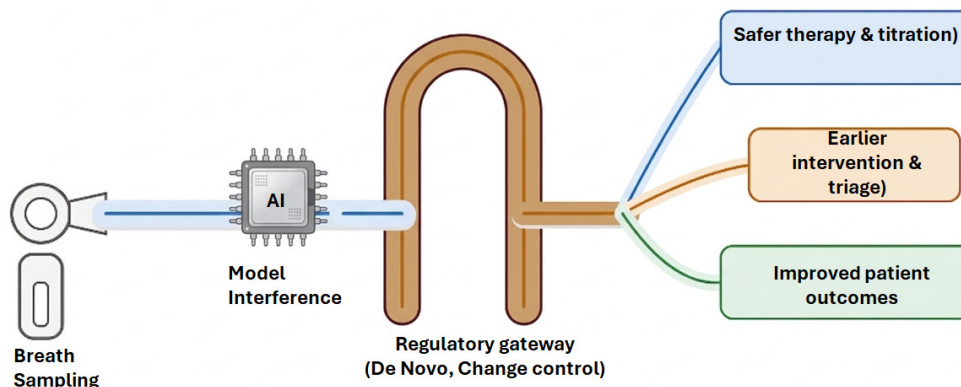
context rather than the physiology. For this reason, clinical breath bioinformatics should rely less on random data splits and more on validation strategies that test generalizability directly. Cross-site and cross-period holdouts should become routine. Data splits should respect batch structure, for example, through site-level or time-blocked evaluation. Feature selection and analysis plans should be defined in advance whenever possible, and models should be locked before prospective testing in cohorts that reflect the intended workflow. Drift and batch surveillance should also be built into the analytical pipeline itself, with quality-control rules and acceptance thresholds that indicate when an instrument or a deployment has moved outside its calibrated operating range.

The use of AI is becoming increasingly important in real-time breath diagnostics. Most breath-analysis platforms generate multivariate spectra, and these are influenced not only by disease-related chemistry but also by humidity, breath fraction, flow, and background room air, all factors that should be measured and controlled. In PTR-ToF-MS and related workflows, machine-learning methods have already been used to convert large ion-feature datasets into rapid screening outputs in clinical settings [8,36]. This shows that algorithmic analysis can operate on timescales relevant to triage, provided that sampling and labeling are carefully controlled. Breath, however, is not a static specimen. It is a time-structured signal, meaning that kinetic information can improve interpretation by helping distinguish endogenous production from inhaled contaminants and by revealing responses to intervention that may be obscured when spectra are averaged. In practice, model performance is often shaped before classification is even attempted. Quality-control and preprocessing steps are therefore critical. These include selecting the appropriate breath phase, confirming end-tidal sampling using CO<sub>2</sub> structure, and parallel sampling of ambient air. Such steps reduce the risk that a classifier will rely on dead-space composition or exposure carryover rather than on physiological information. Questions of governance and privacy are also directly relevant. Multicenter learning will be necessary if breath models are to generalize across settings, but breath profiles may still be indirectly identifiable through stable personal patterns or unusual exposures. Federated learning and privacy-first analytical approaches offer one way to build and test models across sites without pooling raw patient-level data [95,96]. Their value, however, depends on transparent reporting of performance across sites, demographic groups, comorbidities, and collection environments. For this reason, breath models should not be viewed as fixed outputs of a single study. They should be treated as clinical assets that require ongoing maintenance. Seasonal exposure drift, workflow changes, and instrument aging are all expected. If updates are made informally or not documented at all, performance loss is likely to be gradual, difficult to detect, and clinically important.

### 5.5. Regulatory Pathways for Device–Algorithm Systems

Many clinically useful breath-based systems will combine a physical device with software that can be updated over time. This creates a regulatory challenge that extends beyond analytical performance alone. There is already precedent in breath diagnostics from FeNO systems, where exhalation control, operating conditions, and quality control are built into the device itself. For newer breathomics platforms, particularly those without a clear predicate device, the regulatory route is more likely to follow a risk-based framework [97]. The issue becomes more complicated when the product includes an algorithm that contributes directly to the result, because in that setting the software is not simply a reporting tool but part of the diagnostic function. Two regulatory points follow from this. First, the intended use of the system determines the level of evidence required. A general wellness claim falls within a lower-risk, non-diagnostic category, whereas a diagnostic or therapy-guiding claim places the system within medical-device regulation and requires evidence that matches that purpose [98]. Second, any future model updates need to be anticipated and managed in advance. Predetermined change-control plans are intended to make such updates part of a formal regulatory process rather than an informal technical adjustment [99]. This is particularly relevant in breath diagnostics, where dataset shift is likely to occur as devices age, workflows change, and exposure patterns vary over time. For that reason, regulatory planning needs to begin early and

develop in parallel with the measurement strategy and inference pipeline. Quality systems, drift control, auditability, and change management are not secondary administrative matters, but part of the product itself. Figure 4 outlines this device-and-algorithm pathway and shows how it links to validation.



**Figure 4.** Regulatory pathway for breath diagnostic device-and-algorithm systems. This process encompasses intended use, risk classification, evidentiary pathway, algorithm change control, and validation linkage across the product lifecycle.

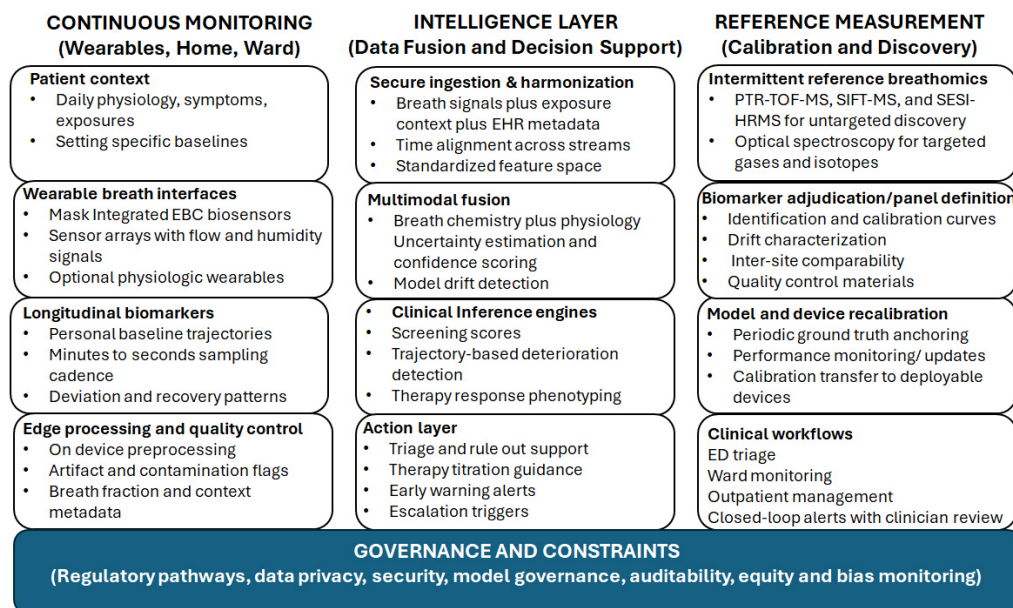
#### 5.6. Outcome-Based Validation and Clinical Implementation

Accuracy, sensitivity, and AUC describe how well a system separates groups, but by themselves, they do not show that the result is clinically useful. For breath testing to have value in practice, it must improve clinical management. This may involve shortening the time to isolation decisions in syndromic respiratory illness, prompting earlier escalation when deterioration is genuine, supporting de-escalation when confirmatory testing is negative, reducing exacerbations through earlier treatment, or improving stewardship by helping distinguish inflammatory relapse from noninflammatory symptoms. Validation should therefore extend beyond diagnostic performance and examine what happens downstream when the test is used under routine clinical conditions. Relevant outcomes include decreased time to decision, days of isolation avoided, antibiotic exposure, admissions, ICU transfers, exacerbation frequency, length of stay, and the consequences of false-positive and false-negative results. Wherever possible, such studies should be prospective, embedded in the clinical pathway in which the test is intended to operate, and follow prespecified analysis plans. Claims of implementation also require evidence that the system remains reliable outside a single center or a tightly controlled study. Breath-based platforms need to show that measurements are reproducible across sites, seasons, operators, and devices, with clear reporting of sampling adherence, ambient conditions, and quality-control failures. This is particularly important because the clinical value of breath testing may be greatest in settings where conventional laboratory turnaround is difficult to achieve. In this respect, breath diagnostics should be viewed not only as an analytical task, but also as an implementation task. Sampling, metrology, inference, governance, and clinical treatment all need to be developed together and evaluated as parts of the same assay. The validation pathway for breath biomarkers also differs from that of more established diagnostic modalities. Imaging biomarkers are usually linked to lesion localization, standardized acquisition methods, and familiar interpretive frameworks. Liquid-biopsy markers, by contrast, often rely on defined molecular targets that can be compared directly with tissue pathology or genotyping. Breath biomarkers rarely have either of these advantages but are more susceptible to pre-analytical variability, environmental carryover, and dependence on multivariate patterns. Their strength lies in frequent, non-invasive sampling, which enables the tracking of physiological changes over short

timescales. For that reason, validation in this field should draw on the discipline used for established biomarker pathways, while giving particular attention to issues that are specific to breath, such as breath-fraction control, measurement of confounders, robustness in longitudinal use, and clinical utility defined in relation to the intended use.

## 6. Future Directions: Integration into Clinical Workflows

Once validation has been aligned with intended use, the next challenge is implementation in routine care. Future progress will depend not only on the ability to acquire repeated breath measurements, but also on whether those measurements can be translated into decisions reliable enough to guide effective clinical action. Figure 5 summarizes this logic. As discussed by Rattray et al. [100], the value of repeated sampling lies in its ability to follow temporal change rather than depend on isolated measurements. For this to be useful in practice, breath systems will need individualized baselines, structured incorporation of breathing patterns and exposure metadata, and clear rules for handling measurements obtained under unstable or poorly calibrated conditions. The resulting outputs will also need to be presented in a form that clinicians can act upon, for example, as a probability of deterioration or recovery linked to recommendations for repeat sampling, confirmatory testing, escalation, or continued monitoring. When uncertainty is high, this should be stated explicitly rather than masked by an apparently confident label. A further issue is dataset shift. In breath diagnostics, seasonal exposures, device aging, inlet history, and differences in workflow can all alter the measured feature space even when the underlying biology has not changed. This means that quality control, calibration transfer, and drift surveillance will need to be incorporated into routine lifecycle management rather than treated as occasional technical checks. This heightens the need for more advanced and complete electronic health records, and for artificial intelligence electronic health systems capable of highlighting and integrating disparate data sources over time to support clinical action, rather than simply applying machine learning to datasets in isolation. Finally, multicenter learning will be necessary if breath systems are to generalize across settings, but this also raises privacy concerns, particularly for wearable platforms that generate dense longitudinal profiles.



**Figure 5.** An integrated future architecture for breath-based diagnostics and monitoring. This conceptual systems diagram links wearable longitudinal monitoring, intermittent reference measurement, multimodal inference, and clinical decision outputs within regulatory and privacy constraints.

## 7. Conclusions

Breath diagnostics are most likely to enter clinical practice first in settings where sampling can be standardized, comparator groups reflect real clinical differentials, and the resulting readout can be linked to a clear clinical response. At present, these conditions are best met in respiratory infection triage, peri-procedural monitoring, dialysis-linked metabolic assessment, and the longitudinal management of inflammatory airway disease, where timing matters and sampling interfaces are relatively stable. The main barrier to wider clinical use is no longer the ability to detect breath signals, but the ability to obtain measurements that remain reproducible across different sites and workflows. From a molecular perspective, the strongest breath claims will be those that link signals to defensible biochemical provenance and transport. Examples include acetone as a readout of ketone handling, isoprene as a physiology-sensitive product of isoprenoid metabolism, ammonia as a proxy for urea and nitrogen handling in renal impairment, nitric oxide as an airway signal tied to inducible nitric oxide synthase activity in type 2 inflammation, and reactive carbonyls as downstream products of lipid peroxidation during oxidative stress. Because many of these pathways converge across diseases, breath signatures should be framed as biochemical states rather than as single-molecule disease labels, and key discriminatory features should be supported by traceable metrology, appropriate calibration in humid matrices, and orthogonal confirmation when molecular identity is central to interpretation. Progress will therefore depend on standardized breath-fraction control, clinically realistic study design, pathway-anchored feature definitions, and appropriately governed devices and algorithm systems that support better decision-making rather than simply separate groups. Under these conditions, breath analysis may shift from a promising research tool to a reliable clinical assay that reports molecular physiology at the time scale of care delivery.

**Acknowledgments:** The authors gratefully acknowledge the interdisciplinary academic environment and scientific discussions within the University of Oklahoma, OU Health Children's Clinic, and OU Health Stephenson Cancer Center, which helped shape the scope and clinical framing of this review.

**Author Contributions:** Conceptualization, V.T.; writing—original draft preparation, V.T.; literature curation, S.Y., P.W.S., and S.V.; writing—review and editing, N.T., N.N., N.C., S.Y., J.P.M., D.H., D.L.T., and T.V.; supervision, T.V. All authors have read and agreed to the published version of the manuscript.

**Conflicts of Interest:** The authors declare no conflict of interest.

**Funding:** The authors thank OU Children's Hospital, Oklahoma City, for seed funding, and the OU Vice President of Research and Partnerships, Norman, for supporting funding through the Bridge Funding Investment Program.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Not applicable.

## Abbreviations

AI	artificial intelligence
AUC	area under the curve
AUROC	area under the receiver operating characteristic curve
BET	buffered end-tidal
BUN	blood urea nitrogen
CI-ToF-MS	chemical ionization time-of-flight mass spectrometry
CMOS	complementary metal-oxide-semiconductor
CO	carbon monoxide
CO <sub>2</sub>	carbon dioxide
COPD	chronic obstructive pulmonary disease

COVID-19 coronavirus disease 2019  
CPET cardiopulmonary exercise testing  
CRDS cavity ring-down spectroscopy  
EBC exhaled breath condensate  
FeNO fractional exhaled nitric oxide  
FDA U.S. Food and Drug Administration  
GC gas chromatography  
GC-IMS gas chromatography–ion mobility spectrometry  
HRMS high-resolution mass spectrometry  
ICU intensive care unit  
IMS ion mobility spectrometry  
Kt/V dialysis adequacy index (dialyzer clearance × dialysis time / volume of distribution of urea)  
MOX metal oxide  
MS mass spectrometry  
NO nitric oxide  
NPV negative predictive value  
PCB printed circuit board  
PMA premarket approval  
PPV positive predictive value  
PTR proton transfer reaction  
PTR-MS proton transfer reaction mass spectrometry  
PTR-ToF-MS proton transfer reaction time-of-flight mass spectrometry  
QC quality control  
QCM quartz crystal microbalance  
SCFA short-chain fatty acid(s)  
SESI secondary electrospray ionization  
SESI-HRMS secondary electrospray ionization high-resolution mass spectrometry  
SIFT selected ion flow tube  
SIFT-MS selected ion flow tube mass spectrometry  
SRI-ToF-MS selective reagent ionization time-of-flight mass spectrometry  
TCA tricarboxylic acid  
TD-GC-MS thermal desorption gas chromatography–mass spectrometry  
ToF time-of-flight  
VOC volatile organic compound  
VOCs volatile organic compounds

## References

1. WHA76.5: Strengthening Diagnostics Capacity. Available online: [https://apps.who.int/gb/ebwha/pdf\\_files/WHA76/A76\\_R5-en.pdf](https://apps.who.int/gb/ebwha/pdf_files/WHA76/A76_R5-en.pdf) (accessed on 24 March 2026).
2. Amann, A.; Miekisch, W.; Schubert, J.; Buszewski, B.; Ligor, T.; Jezierski, T.; Pleil, J.; Risby, T. Analysis of exhaled breath for disease detection. *Annu. Rev. Anal. Chem.* 2014, 7, 455–482. <https://doi.org/10.1146/annurev-anchem-071213-020043>
3. Pereira, J.A.M.; Porto-Figueira, P.; Cavaco, C.; Taunk, K.; Rapole, S.; Dhakne, R.; Nagarajaram, H.; Câmara, J.S. Breath analysis as a potential and non-invasive frontier in disease diagnosis: An overview. *Metabolites* 2015, 5, 3–55. <https://doi.org/10.3390/metabo5010003>
4. Sharma, A.; Kumar, R.; Varadwaj, P. Smelling the disease: Diagnostic potential of breath analysis. *Mol. Diagn. Ther.* 2023, 27, 321–347. <https://doi.org/10.1007/s40291-023-00640-7>
5. Issitt, T.; Wiggins, L.; Veysey, M.; Sweeney, S.T.; Brackenbury, W.J.; Redeker, K. Volatile compounds in human breath: Critical review and meta-analysis. *J. Breath Res.* 2022, 16, 024001. <https://doi.org/10.1088/1752-7163/ac5230>

6. Brinkman, P.; Wilde, M.; Ahmed, W.; Wang, R.; van der Schee, M.; Abuhelal, S.; Schaber, C.; Cunoosamy, D.; Clarke, G.W.; Maitland-van der Zee, A.-H.; Dahlén, S.-E.; Siddiqui, S.; Fowler, S.J. Fulfilling the promise of breathomics: Considerations for the discovery and validation of exhaled volatile biomarkers. *Am. J. Respir. Crit. Care Med.* 2024, 210, 1079–1090. <https://doi.org/10.1164/rccm.202305-0868TR>
7. Liang, Q.; Chan, Y.-C.; Changala, P.B.; Nesbitt, D.J.; Ye, J.; Toscano, J. Ultrasensitive multispecies spectroscopic breath analysis for real-time health monitoring and diagnostics. *Proc. Natl. Acad. Sci. USA* 2021, 118, e2105063118. <https://doi.org/10.1073/pnas.2105063118>
8. Roquencourt, C.; Salvator, H.; Bardin, E.; Lamy, E.; Farfour, E.; Naline, E.; Devillier, P.; Grassin-Delyle, S. Enhanced real-time mass spectrometry breath analysis for the diagnosis of COVID-19. *ERJ Open Res.* 2023, 9, 00206-2023. <https://doi.org/10.1183/23120541.00206-2023>
9. Miekisch, W.; Sukul, P.; Schubert, J.K. Diagnostic potential of breath analysis: Focus on the dynamics of volatile organic compounds. *TrAC Trends Anal. Chem.* 2024, 180, 117977. <https://doi.org/10.1016/j.trac.2024.117977>
10. Ethics and Governance of Artificial Intelligence for Health: Guidance on Large Multi-Modal Models. Available online: <https://www.who.int/publications/i/item/9789240084759> (accessed on 24 March 2026).
11. Drabińska, N.; Flynn, C.; Ratcliffe, N.; Belluomo, I.; Myridakis, A.; Gould, O.; Fois, M.; Smart, A.; Devine, T.; de Lacy Costello, B. A literature survey of all volatiles from healthy human breath and bodily fluids: The human volatilome. *J. Breath Res.* 2021, 15, 034001. <https://doi.org/10.1088/1752-7163/abf1d0>
12. Herbig, J.; Titzmann, T.; Beauchamp, J.; Kohl, I.; Hansel, A. Buffered end-tidal (BET) sampling—a novel method for real-time breath gas analysis. *J. Breath Res.* 2008, 2, 037008. <https://doi.org/10.1088/1752-7155/2/3/037008>
13. Horváth, I.; Barnes, P.J.; Loukides, S.; Sterk, P.J.; Högman, M.; Olin, A.-C.; Amann, A.; Antus, B.; Baraldi, E.; Bikov, A.; Boots, A.W.; Bos, L.D.; Brinkman, P.; Bucca, C.; Carpagnano, G.E.; Corradi, M.; Cristescu, S.; de Jongste, J.C.; Dinh-Xuan, A.-T.; et al. A European Respiratory Society technical standard: Exhaled biomarkers in lung disease. *Eur. Respir. J.* 2017, 49, 1600965. <https://doi.org/10.1183/13993003.00965-2016>
14. Smolinska, A.; Tedjo, D.I.; Blanchet, L.; Bodelier, A.; Pierik, M.J.; Masclee, A.A.M.; Dallinga, J.; Savelkoul, P.H.M.; Jonkers, D.M.A.E.; Penders, J.; van Schooten, F.-J. Volatile metabolites in breath strongly correlate with gut microbiome in CD patients. *Anal. Chim. Acta* 2018, 1025, 1–11. <https://doi.org/10.1016/j.aca.2018.03.046>
15. Neyrinck, A.M.; Rodriguez, J.; Zhang, Z.; Nazare, J.-A.; Bindels, L.B.; Cani, P.D.; Maquet, V.; Laville, M.; Bischoff, S.C.; Walter, J.; Delzenne, N.M. Breath volatile metabolome reveals the impact of dietary fibres on the gut microbiota: Proof of concept in healthy volunteers. *eBioMedicine* 2022, 80, 104051. <https://doi.org/10.1016/j.ebiom.2022.104051>
16. Hernandez-Leyva, A.J.; Berna, A.Z.; Bui, M.H.; Liu, Y.; Rosen, A.L.; Lint, M.A.; Whiteside, S.A.; Jaeger, N.; McDonough, R.T.; Joardar, N.; Santiago-Borges, J.; Tomera, C.P.; Luo, W.; Odom John, A.R.; Kau, A.L. The gut microbiota shapes the human and murine breath volatilome. *Cell Metab.* 2026, online ahead of print. <https://doi.org/10.1016/j.cmet.2025.12.013>
17. Dweik, R.A.; Boggs, P.B.; Erzurum, S.C.; Irvin, C.G.; Leigh, M.W.; Lundberg, J.O.; Olin, A.-C.; Plummer, A.L.; Taylor, D.R. An official ATS clinical practice guideline: Interpretation of exhaled nitric oxide levels (FENO) for clinical applications. *Am. J. Respir. Crit. Care Med.* 2011, 184, 602–615. <https://doi.org/10.1164/rccm.9120-11ST>
18. Khatri, S.B.; Iaccarino, J.M.; Barochia, A.; Soghier, I.; Akuthota, P.; Brady, A.; Covar, R.A.; Debley, J.S.; Diamant, Z.; Fitzpatrick, A.M.; Kaminsky, D.A.; Kenyon, N.J.; Khurana, S.; Lipworth, B.J.; McCarthy, K.; Peters, M.; Que, L.G.; Ross, K.R.; Schneider-Futschik, E.K.; et al. Use of fractional exhaled nitric oxide to guide the treatment of asthma: An official American Thoracic Society clinical practice guideline. *Am. J. Respir. Crit. Care Med.* 2021, 204, e97–e109. <https://doi.org/10.1164/rccm.202109-2093ST>
19. Grob, N.M.; Aytekin, M.; Dweik, R.A. Biomarkers in exhaled breath condensate: A review of collection, processing and analysis. *J. Breath Res.* 2008, 2, 037004. <https://doi.org/10.1088/1752-7155/2/3/037004>
20. Hunt, J. Exhaled breath condensate: An overview. *Immunol. Allergy Clin. North Am.* 2007, 27, 587–596. <https://doi.org/10.1016/j.iac.2007.09.001>

21. Montuschi, P. Analysis of exhaled breath condensate in respiratory medicine: Methodological aspects and potential clinical applications. *Ther. Adv. Respir. Dis.* 2007, 1, 5–23. <https://doi.org/10.1177/1753465807082373>
22. Maniscalco, M.; Candia, C.; Fuschillo, S.; Ambrosino, P.; Paris, D.; Motta, A. Exhaled breath condensate (EBC) in respiratory diseases: Recent advances and future perspectives in the age of omic sciences. *J. Breath Res.* 2024, 18, 045001. <https://doi.org/10.1088/1752-7163/ad7a9a>
23. Horváth, I.; Hunt, J.; Barnes, P.J.; Alving, K.; Antczak, A.; Baraldi, E.; Becher, G.; van Beurden, W.J.C.; Corradi, M.; Dekhuijzen, R.; Dweik, R.A.; Dwyer, T.; Effros, R.; Erzurum, S.; Gaston, B.; Gessner, C.; Greening, A.; Ho, L.-P.; Hohlfeld, J.; Jöbssis, Q.; et al. Exhaled breath condensate: Methodological recommendations and unresolved questions. *Eur. Respir. J.* 2005, 26, 523–548. <https://doi.org/10.1183/09031936.05.00029705>
24. Chen, X.; Zhang, K.; Yin, Z.; Fang, M.; Pu, W.; Liu, Z.; Li, L.; Sinues, P.; Dallmann, R.; Zhou, Z.; Li, X. Online real-time monitoring of exhaled breath particles reveals unnoticed transport of nonvolatile drugs from blood to breath. *Anal. Chem.* 2021, 93, 5005–5008. <https://doi.org/10.1021/acs.analchem.1c00509>
25. Gisler, A.; Lan, J.; Singh, K.D.; Usemann, J.; Frey, U.; Zenobi, R.; Sinues, P. Real-time breath analysis of exhaled compounds upon peppermint oil ingestion by secondary electrospray ionization-high resolution mass spectrometry: Technical aspects. *J. Breath Res.* 2020, 14, 046001. <https://doi.org/10.1088/1752-7163/ab9f8b>
26. Sasiene, Z.J.; LeBrun, E.S.; Schaller, E.; Mach, P.M.; Taylor, R.; Candelaria, L.; Glaros, T.G.; Baca, J.; McBride, E.M. Real-time breath analysis towards a healthy human breath profile. *J. Breath Res.* 2024, 18, 026003. <https://doi.org/10.1088/1752-7163/ad1cf1>
27. Henderson, B.; Ruzskiewicz, D.M.; Wilkinson, M.; Beauchamp, J.D.; Cristescu, S.M.; Fowler, S.J.; Salman, D.; Di Francesco, F.; Koppen, G.; Langejürgen, J.; Holz, O.; Hadjithekli, A.; Moreno, S.; Pedrotti, M.; Sinues, P.; Slingers, G.; Wilde, M.; Lomonaco, T.; Zanella, D.; Zenobi, R.; Focant, J.-F.; Grassin-Delyle, S.; Franchina, F.A.; Malásková, M.; Stefanuto, P.-H.; Pugliese, G.; Mayhew, C.A.; Thomas, C.L.P. A benchmarking protocol for breath analysis: The peppermint experiment. *J. Breath Res.* 2020, 14, 046008. <https://doi.org/10.1088/1752-7163/aba130>
28. Henderson, B.; Slingers, G.; Pedrotti, M.; Pugliese, G.; Malásková, M.; Bryant, L.; Lomonaco, T.; Ghimenti, S.; Moreno, S.; Cordell, R.; Harren, F.J.M.; Schubert, J.; Mayhew, C.A.; Wilde, M.; Di Francesco, F.; Koppen, G.; Beauchamp, J.D.; Cristescu, S.M. The peppermint breath test benchmark for PTR-MS and SIFT-MS. *J. Breath Res.* 2021, 15, 046005. <https://doi.org/10.1088/1752-7163/ac1fcf>
29. Pham, Y.L.; Yu, R.; Beauchamp, J. Cross-validation of the peppermint benchmarking experiment across three analytical platforms. *J. Breath Res.* 2023, 17, 046003. <https://doi.org/10.1088/1752-7163/ace8b2>
30. Zhan, X.; Duan, J.; Duan, Y. Recent developments of proton-transfer reaction mass spectrometry (PTR-MS) and its applications in medical research. *Mass Spectrom. Rev.* 2013, 32, 143–165. <https://doi.org/10.1002/mas.21357>
31. Casas-Ferreira, A.M.; del Nogal-Sánchez, M.; Pérez-Pavón, J.L.; Moreno-Cordero, B. Non-separative mass spectrometry methods for non-invasive medical diagnostics based on volatile organic compounds: A review. *Anal. Chim. Acta* 2019, 1045, 10–22. <https://doi.org/10.1016/j.aca.2018.07.005>
32. Lindinger, W.; Hansel, A.; Jordan, A. On-line monitoring of volatile organic compounds at pptv levels by means of proton-transfer-reaction mass spectrometry (PTR-MS): Medical applications, food control and environmental research. *Int. J. Mass Spectrom. Ion Process.* 1998, 173, 191–241. [https://doi.org/10.1016/S0168-1176\(97\)00281-4](https://doi.org/10.1016/S0168-1176(97)00281-4)
33. Winkler, K.; Herbig, J.; Kohl, I.; Beauchamp, J.; Trefz, P.; Miekisch, W.; Schubert, J.K. Real-time metabolic monitoring with proton transfer reaction mass spectrometry. *J. Breath Res.* 2013, 7, 036006. <https://doi.org/10.1088/1752-7155/7/3/036006>
34. Jordan, A.; Haidacher, S.; Hanel, G.; Hartungen, E.; Märk, L.; Seehauser, H.; Schottkowsky, R.; Sulzer, P.; Märk, T.D. A high resolution and high sensitivity proton-transfer-reaction time-of-flight mass spectrometer (PTR-TOF-MS). *Int. J. Mass Spectrom.* 2009, 286, 122–128. <https://doi.org/10.1016/j.ijms.2009.07.005>

35. Graus, M.; Müller, M.; Hansel, A. High resolution PTR-TOF: Quantification and formula confirmation of VOC in real time. *J. Am. Soc. Mass Spectrom.* 2010, 21, 1037–1044. <https://doi.org/10.1016/j.jasms.2010.02.006>
36. Liangou, A.; Tasoglou, A.; Huber, H.J.; Wistrom, C.; Brody, K.; Menon, P.G.; Bebekoski, T.; Menschel, K.; Davidson-Fiedler, M.; DeMarco, K.; Salphale, H.; Wistrom, J.; Wistrom, S.; Lee, R.J. A method for the identification of COVID-19 biomarkers in human breath using proton transfer reaction time-of-flight mass spectrometry. *EClinicalMedicine* 2021, 42, 101207. <https://doi.org/10.1016/j.eclinm.2021.101207>
37. Pugliese, G.; Trefz, P.; Brock, B.; Schubert, J.K.; Miekisch, W. Extending PTR-based breath analysis to real-time monitoring of reactive volatile organic compounds. *Analyst* 2019, 144, 7359–7367. <https://doi.org/10.1039/C9AN01478K>
38. Smith, D.; Španěl, P. On the importance of accurate quantification of individual volatile metabolites in exhaled breath. *J. Breath Res.* 2017, 11, 047106. <https://doi.org/10.1088/1752-7163/aa7ab5>
39. Wüthrich, C.; Käser, T.; Zenobi, R.; Giannoukos, S. Internal standard addition system for online breath analysis. *Anal. Chem.* 2024, 96, 10871–10876. <https://doi.org/10.1021/acs.analchem.4c01924>
40. Meurs, J.; Sakkoula, E.; Cristescu, S.M. Real-time non-invasive monitoring of short-chain fatty acids in exhaled breath. *Front. Chem.* 2022, 10, 853541. <https://doi.org/10.3389/fchem.2022.853541>
41. Španěl, P.; Smith, D. Progress in SIFT-MS: Breath analysis and other applications. *Mass Spectrom. Rev.* 2011, 30, 236–267. <https://doi.org/10.1002/mas.20303>
42. Belluomo, I.; Boshier, P.R.; Myridakis, A.; Vadhwana, B.; Markar, S.R.; Španěl, P.; Hanna, G.B. Selected ion flow tube mass spectrometry for targeted analysis of volatile organic compounds in human breath. *Nat. Protoc.* 2021, 16, 3419–3438. <https://doi.org/10.1038/s41596-021-00542-0>
43. Boshier, P.R.; Cushnir, J.R.; Mistry, V.; Knaggs, A.; Španěl, P.; Smith, D.; Hanna, G.B. On-line, real time monitoring of exhaled trace gases by SIFT-MS in the perioperative setting: A feasibility study. *Analyst* 2011, 136, 3233–3237. <https://doi.org/10.1039/C1AN15356K>
44. Blanco, F.G.; Vidal-de-Miguel, G. Breath analysis by secondary electro-spray ionization-mass spectrometry to interrogate biologically significant metabolites non-invasively. *Crit. Rev. Anal. Chem.* 2023, 53, 825–837. <https://doi.org/10.1080/10408347.2021.1981226>
45. Zeng, J.; Usemann, J.; Singh, K.D.; Jochmann, A.; Trachsel, D.; Frey, U.; Sinues, P. Pharmacometabolomics via real-time breath analysis captures metabolotypes of asthmatic children associated with salbutamol responsiveness. *iScience* 2024, 27, 111446. <https://doi.org/10.1016/j.isci.2024.111446>
46. Neri, G.; Lacquaniti, A.; Rizzo, G.; Donato, N.; Latino, M.; Buemi, M. Real-time monitoring of breath ammonia during haemodialysis: Use of ion mobility spectrometry (IMS) and cavity ring-down spectroscopy (CRDS) techniques. *Nephrol. Dial. Transplant.* 2012, 27, 2945–2952. <https://doi.org/10.1093/ndt/gfr738>
47. Henderson, B.; Khodabakhsh, A.; Metsälä, M.; Ventrillard, I.; Schmidt, F.M.; Romanini, D.; Ritchie, G.A.D.; te Lintel Hekkert, S.; Briot, R.; Risby, T.; Marczin, N.; Harren, F.J.M.; Cristescu, S.M. Laser spectroscopy for breath analysis: Towards clinical implementation. *Appl. Phys. B* 2018, 124, 161. <https://doi.org/10.1007/s00340-018-7030-x>
48. Metsälä, M. Optical techniques for breath analysis: From single to multi-species detection. *J. Breath Res.* 2018, 12, 027104. <https://doi.org/10.1088/1752-7163/aa8a31>
49. Navas, M.J.; Jiménez, A.M.; Asuero, A.G. Human biomarkers in breath by photoacoustic spectroscopy. *Clin. Chim. Acta* 2012, 413, 1171–1178. <https://doi.org/10.1016/j.cca.2012.04.008>
50. Dumitras, D.C.; Petrus, M.; Bratu, A.-M.; Popa, C. Applications of near infrared photoacoustic spectroscopy for analysis of human respiration: A review. *Molecules* 2020, 25, 1728. <https://doi.org/10.3390/molecules25071728>
51. Zhou, T.; Wu, T.; Wu, Q.; Chen, W.; Wu, M.; Ye, C.; He, X. Real-time monitoring of <sup>13</sup>C- and <sup>18</sup>O-isotopes of human breath CO<sub>2</sub> using a mid-infrared hollow waveguide gas sensor. *Anal. Chem.* 2020, 92, 12943–12949. <https://doi.org/10.1021/acs.analchem.0c01586>
52. Nidheesh, V.R.; Mohapatra, A.K.; Kartha, V.B.; Chidangil, S. Multiwavelength photoacoustic breath analysis sensor for the diagnosis of lung diseases: COPD and asthma. *ACS Sens.* 2023, 8, 4111–4120. <https://doi.org/10.1021/acssensors.3c01316>

53. Wu, T.-C.; De Luca, A.; Zhong, Q.; Zhu, X.; Ogbeide, O.; Um, D.-S.; Hu, G.; Albrow-Owen, T.; Udrea, F.; Hasan, T. Inkjet-printed CMOS-integrated graphene-metal oxide sensors for breath analysis. *npj 2D Mater. Appl.* 2019, 3, 42. <https://doi.org/10.1038/s41699-019-0125-3>
54. Kwak, C.-H.; Kim, T.-H.; Jeong, S.-Y.; Yoon, J.-W.; Kim, J.-S.; Lee, J.-H. Humidity-independent oxide semiconductor chemiresistors using terbium-doped SnO<sub>2</sub> yolk-shell spheres for real-time breath analysis. *ACS Appl. Mater. Interfaces* 2018, 10, 18886–18894. <https://doi.org/10.1021/acsami.8b04245>
55. Verma, A.; Yadav, B.C. Development and integration of a hierarchical Pd/WO<sub>3</sub> acetone-sensing device for real-time exhaled breath monitoring with disposable face mask. *J. Hazard. Mater.* 2024, 463, 132872. <https://doi.org/10.1016/j.jhazmat.2023.132872>
56. Peña, A.; Aguilera, J.D.; Matatagui, D.; de la Presa, P.; Horrillo, C.; Hernando, A.; Marín, P. Real-time monitoring of breath biomarkers with a magnetoelastic contactless gas sensor: A proof of concept. *Biosensors* 2022, 12, 871. <https://doi.org/10.3390/bios12100871>
57. Afzal, U.; Wang, K.; Liang, J.; Fatima, A.; Zhang, L.; Wu, T.; Haider, H.S.; Mustafa, S.; Ma, K. High-performance flexible gas sensors based on W-VO<sub>2</sub>/1D-carbon composites for real-time ammonia detection in breath and environmental monitoring. *ACS Sens.* 2025, 10, 9033–9043. <https://doi.org/10.1021/acssensors.5c03382>
58. Chang, Q.; Wu, D.; Huang, Y.; Liang, C.; Liu, L.; Liu, H.; He, Y.; Huang, Q.; Qiu, J.; Tang, X. A lead-free K<sub>2</sub>CuBr<sub>3</sub> microwires-based humidity sensor realized via QCM for real-time breath monitoring. *Sens. Actuators B Chem.* 2022, 367, 132112. <https://doi.org/10.1016/j.snb.2022.132112>
59. Roudjane, M.; Bellemare-Rousseau, S.; Drouin, E.; Belanger-Huot, B.; Dugas, M.-A.; Miled, A.; Messaddeq, Y. Smart T-shirt based on wireless communication spiral fiber sensor array for real-time breath monitoring: Validation of the technology. *IEEE Sens. J.* 2020, 20, 10841–10850. <https://doi.org/10.1109/JSEN.2020.2993286>
60. Srikrishnarka, P.; Patrakka, J.; Sun, Z.; Nonappa. Optical fiber sensor with dynamically responsive cladding for real-time breath pattern monitoring. *ACS Omega* 2025, 10, 44528–44540. <https://doi.org/10.1021/acsomega.5c06598>
61. Heng, W.; Yin, S.; Min, J.; Wang, C.; Han, H.; Shirzaei Sani, E.; Li, J.; Song, Y.; Rossiter, H.B.; Gao, W. A smart mask for exhaled breath condensate harvesting and analysis. *Science* 2024, 385, 954–961. <https://doi.org/10.1126/science.adn6471>
62. Maier, D.; Laubender, E.; Basavanna, A.; Schumann, S.; Güder, F.; Urban, G.A.; Dincer, C. Toward continuous monitoring of breath biochemistry: A paper-based wearable sensor for real-time hydrogen peroxide measurement in simulated breath. *ACS Sens.* 2019, 4, 2945–2951. <https://doi.org/10.1021/acssensors.9b01403>
63. Jafari, M.; Solhi, E.; Tagi, S.; Hasanzadeh, M.; Jouyban-Gharamaleki, V.; Jouyban, A.; Shadjou, N. Non-invasive quantification of malondialdehyde biomarker in human exhaled breath condensate using self-assembled organic-inorganic nanohybrid: A new platform for early diagnosis of lung disease. *J. Pharm. Biomed. Anal.* 2019, 164, 249–257. <https://doi.org/10.1016/j.jpba.2018.10.048>
64. Ahmed, W.M.; Lawal, O.; Nijssen, T.M.E.; Goodacre, R.; Fowler, S.J. Exhaled volatile organic compounds of infection: A systematic review. *ACS Infect. Dis.* 2017, 3, 695–710. <https://doi.org/10.1021/acsinfectdis.7b00088>
65. van Oort, P.M.P.; Nijssen, T.M.; White, I.R.; Knobel, H.H.; Felton, T.; Rattray, N.; Lawal, O.; Bulut, M.; Ahmed, W.; Artigas, A.; Povoia, P.R.; Martin-Loeches, I.; Weda, H.; Goodacre, R.; Schultz, M.J.; Dark, P.M.; Fowler, S.J.; Bos, L.D.; the BreathDx Consortium. Untargeted molecular analysis of exhaled breath as a diagnostic test for ventilator-associated lower respiratory tract infections (BreathDx). *Thorax* 2022, 77, 79–81. <https://doi.org/10.1136/thoraxjnl-2021-217362>
66. Ibrahim, W.; Carr, L.; Cordell, R.; Wilde, M.J.; Salman, D.; Monks, P.S.; Thomas, P.; Brightling, C.E.; Siddiqui, S.; Greening, N.J. Breathomics for the clinician: The use of volatile organic compounds in respiratory diseases. *Thorax* 2021, 76, 514–521. <https://doi.org/10.1136/thoraxjnl-2020-215667>
67. Bos, L.D.; Sterk, P.J.; Fowler, S.J. Breathomics in the setting of asthma and chronic obstructive pulmonary disease. *J. Allergy Clin. Immunol.* 2016, 138, 970–976. <https://doi.org/10.1016/j.jaci.2016.08.004>
68. Fens, N.; Zwinderman, A.H.; van der Schee, M.P.; de Nijs, S.B.; Dijkers, E.; Roldaan, A.C.; Cheung, D.; Bel, E.H.; Sterk, P.J. Exhaled breath profiling enables discrimination of chronic obstructive pulmonary disease

- and asthma. *Am. J. Respir. Crit. Care Med.* 2009, 180, 1076–1082. <https://doi.org/10.1164/rccm.200906-0939OC>
69. Fens, N.; Roldaan, A.C.; van der Schee, M.P.; Boksem, R.J.; Zwinderman, A.H.; Bel, E.H.; Sterk, P.J. External validation of exhaled breath profiling using an electronic nose in the discrimination of asthma with fixed airways obstruction and chronic obstructive pulmonary disease. *Clin. Exp. Allergy* 2011, 41, 1371–1378. <https://doi.org/10.1111/j.1365-2222.2011.03800.x>
  70. de Vries, R.; Dagelet, Y.W.F.; Spoor, P.; Snoey, E.; Jak, P.M.C.; Brinkman, P.; Dijkers, E.; Bootsma, S.K.; Elskamp, F.; de Jongh, F.H.C.; Haarman, E.G.; In 't Veen, J.C.C.M.; Maitland-van der Zee, A.-H.; Sterk, P.J. Clinical and inflammatory phenotyping by breathomics in chronic airway diseases irrespective of the diagnostic label. *Eur. Respir. J.* 2018, 51, 1701817. <https://doi.org/10.1183/13993003.01817-2017>
  71. Robroeks, C.M.; van Berkel, J.J.; Jöbssis, Q.; van Schooten, F.-J.; Dallinga, J.W.; Wouters, E.F.; Dompeling, E. Exhaled volatile organic compounds predict exacerbations of childhood asthma in a 1-year prospective study. *Eur. Respir. J.* 2013, 42, 98–106. <https://doi.org/10.1183/09031936.00010712>
  72. Tejero Rioseas, A.; Singh, K.D.; Nowak, N.; Gaugg, M.T.; Bruderer, T.; Zenobi, R.; Sinues, P.M.-L. Real-time monitoring of tricarboxylic acid metabolites in exhaled breath. *Anal. Chem.* 2018, 90, 6453–6460. <https://doi.org/10.1021/acs.analchem.7b04600>
  73. Osswald, M.; Kohlbrenner, D.; Nowak, N.; Spörri, J.; Sinues, P.; Nieman, D.; Sievi, N.A.; Scherr, J.; Kohler, M. Real-time monitoring of metabolism during exercise by exhaled breath. *Metabolites* 2021, 11, 856. <https://doi.org/10.3390/metabo11120856>
  74. Lee, B.; Lee, J.; Lee, H.-K.; Park, H.J.; Kwack, M.-J.; Kim, D.Y.; Park, I.; Lim, S.; Lee, D.-S. Breath analyzer for real-time exercise fat burning prediction: Oral and alveolar breath insights with CNN. *ACS Sens.* 2025, 10, 2510–2519. <https://doi.org/10.1021/acssensors.4c02502>
  75. Zhang, R.; Huang, W.; Li, G.; Hu, Y. Noninvasive strategy based on real-time in vivo cataluminescence monitoring for clinical breath analysis. *Anal. Chem.* 2017, 89, 3353–3361. <https://doi.org/10.1021/acs.analchem.6b03898>
  76. Cristescu, S.M.; Kiss, R.; te Lintel Hekkert, S.; Dalby, M.; Harren, F.J.M.; Risby, T.H.; Marczin, N. Real-time monitoring of endogenous lipid peroxidation by exhaled ethylene in patients undergoing cardiac surgery. *Am. J. Physiol. Lung Cell. Mol. Physiol.* 2014, 307, L509–L515. <https://doi.org/10.1152/ajplung.00168.2014>
  77. Chan, M.-J.; Li, Y.-J.; Wu, C.-C.; Lee, Y.-C.; Zan, H.-W.; Meng, H.-F.; Hsieh, M.-H.; Lai, C.-S.; Tian, Y.-C. Breath ammonia is a useful biomarker predicting kidney function in chronic kidney disease patients. *Biomedicines* 2020, 8, 468. <https://doi.org/10.3390/biomedicines8110468>
  78. Wagner, M.; Saad, S.; Killard, A.J. Validation of a sensor system for the measurement of breath ammonia using selected-ion flow-tube mass spectrometry. *J. Breath Res.* 2025, 19, 016006. <https://doi.org/10.1088/1752-7163/ad8e7d>
  79. Hanna, G.B.; Boshier, P.R.; Markar, S.R.; Romano, A. Accuracy and methodologic challenges of volatile organic compound-based exhaled breath tests for cancer diagnosis: A systematic review and meta-analysis. *JAMA Oncol.* 2019, 5, e182815. <https://doi.org/10.1001/jamaoncol.2018.2815>
  80. Jin, M.-J.; Li, E.-M.; Xu, L.-Y. Diagnostic accuracy of breath tests based on volatile organic compounds for cancer: A systematic review and meta-analysis. *Clin. Biochem.* 2025, 136, 110898. <https://doi.org/10.1016/j.clinbiochem.2025.110898>
  81. Markar, S.R.; Wiggins, T.; Antonowicz, S.; Chin, S.-T.; Romano, A.; Nikolic, K.; Evans, B.; Cunningham, D.; Mughal, M.; Lagergren, J.; Hanna, G.B. Assessment of a noninvasive exhaled breath test for the diagnosis of oesophagogastric cancer. *JAMA Oncol.* 2018, 4, 970–976. <https://doi.org/10.1001/jamaoncol.2018.0991>
  82. Amal, H.; Leja, M.; Funke, K.; Skapars, R.; Sivins, A.; Ancans, G.; Liepniece-Karele, I.; Kikuste, I.; Lasina, I.; Haick, H. Detection of precancerous gastric lesions and gastric cancer through exhaled breath. *Gut* 2016, 65, 400–407. <https://doi.org/10.1136/gutjnl-2014-308536>
  83. Hong, Y.; Che, X.; Su, H.; Mai, Z.; Huang, Z.; Huang, W.; Chen, W.; Liu, S.; Gao, W.; Zhou, Z.; Tan, G.; Li, X. Exhaled breath analysis using on-line preconcentration mass spectrometry for gastric cancer diagnosis. *J. Mass Spectrom.* 2021, 56, e4588. <https://doi.org/10.1002/jms.4588>
  84. Steenhuis, E.G.M.; Schoenaker, I.J.H.; de Groot, J.W.B.; Fiebrich, H.B.; de Graaf, J.C.; Brohet, R.M.; van Dijk, J.D.; van Westreenen, H.L.; Siersema, P.D.; de Vos tot Nederveen Cappel, W.H. Feasibility of volatile

- organic compound in breath analysis in the follow-up of colorectal cancer: A pilot study. *Eur. J. Surg. Oncol.* 2020, 46, 2068–2073. <https://doi.org/10.1016/j.ejso.2020.07.028>
85. Steenhuis, E.G.M.; Asmara, O.D.; Kort, S.; Papenhuijzen, M.H.G.; Veeger, N.J.G.M.; van den Heuvel, M.M.; van Geffen, W.H. The electronic nose in lung cancer diagnostics: A systematic review and meta-analysis. *ERJ Open Res.* 2025, 11, 00723-2024. <https://doi.org/10.1183/23120541.00723-2024>
86. Zhao, T.; Hu, X.; Peng, M.; Wei, W.; Fu, N.; Chen, C.; Chen, Z. Volatile organic compounds in exhaled human breath for the diagnosis of malignant pleural mesothelioma: A meta-analysis. *Front. Oncol.* 2025, 15, 1537767. <https://doi.org/10.3389/fonc.2025.1537767>
87. Kort, S.; Brusse-Keizer, M.; Schouwink, H.; Citgez, E.; de Jongh, F.H.; van Putten, J.W.G.; van den Borne, B.; Kastelij, E.A.; Stolz, D.; Schuurbijs, M.; van den Heuvel, M.M.; van Geffen, W.H.; van der Palen, J. Diagnosing non-small cell lung cancer by exhaled breath profiling using an electronic nose: A multicenter validation study. *Chest* 2023, 163, 697–706. <https://doi.org/10.1016/j.chest.2022.09.042>
88. de Vries, R.; Farzan, N.; Fabius, T.; de Jongh, F.H.C.; Jak, P.M.C.; Haarman, E.G.; Snoey, E.; In 't Veen, J.C.C.M.; Dagelet, Y.W.F.; Maitland-van der Zee, A.-H.; Lucas, A.; van den Heuvel, M.M.; Wolf-Lansdorf, M.; Muller, M.; Baas, P.; Sterk, P.J. Prospective detection of early lung cancer in patients with COPD in regular care by electronic nose analysis of exhaled breath. *Chest* 2023, 164, 1315–1324. <https://doi.org/10.1016/j.chest.2023.04.050>
89. Buma, A.I.G.; Muntinghe-Wagenaar, M.B.; van der Noort, V.; de Vries, R.; Schuurbijs, M.M.F.; Sterk, P.J.; Schipper, S.P.M.; Meurs, J.; Cristescu, S.M.; Hiltermann, T.J.N.; van den Heuvel, M.M. Lung cancer detection by electronic nose analysis of exhaled breath: A multicentre prospective external validation study. *Ann. Oncol.* 2025, 36, 786–795. <https://doi.org/10.1016/j.annonc.2025.03.013>
90. Yang, Y.; Long, H.; Feng, Y.; Tian, S.; Chen, H.; Zhou, P. A multi-omics method for breast cancer diagnosis based on metabolites in exhaled breath, ultrasound imaging, and basic clinical information. *Heliyon* 2024, 10, e32115. <https://doi.org/10.1016/j.heliyon.2024.e32115>
91. Lee, B.; Lee, J.; Noh, H.; Bahn, H.-K.; Jeon, J.-H.; Park, I.; Jheon, S.; Lee, D.-S. Advanced breath analysis through hierarchical deep convolutional neural network for multi-cancer screening. *npj Digit. Med.* 2026, 9, 138. <https://doi.org/10.1038/s41746-025-02319-1>
92. Venkatraman, P.; Sureka, C.S. An in-vitro study for early detection and to distinguish breast and lung malignancies using the PCB technology based nanodosimeter. *Sci. Rep.* 2019, 9, 380. <https://doi.org/10.1038/s41598-018-36805-2>
93. Mochalski, P.; Shuster, G.; Leja, M.; Unterkofler, K.; Jaeschke, C.; Skapars, R.; Gasenko, E.; Polaka, I.; Vasiljevs, E.; Shani, G.; Mitrovics, J.; Mayhew, C.A.; Haick, H. Non-contact breath sampling for sensor-based breath analysis. *J. Breath Res.* 2019, 13, 036001. <https://doi.org/10.1088/1752-7163/ab0b8d>
94. Märtens, A.; Holle, J.; Mollenhauer, B.; Wegner, A.; Kirwan, J.; Hiller, K. Instrumental drift in untargeted metabolomics: Optimizing data quality with intrastudy QC samples. *Metabolites* 2023, 13, 665. <https://doi.org/10.3390/metabo13050665>
95. Rieke, N.; Hancox, J.; Li, W.; Milletari, F.; Roth, H.R.; Albarqouni, S.; Bakas, S.; Galtier, M.N.; Landman, B.A.; Maier-Hein, K.; Ourselin, S.; Sheller, M.; Summers, R.M.; Trask, A.; Xu, D.; Baust, M.; Cardoso, M.J. The future of digital health with federated learning. *npj Digit. Med.* 2020, 3, 119. <https://doi.org/10.1038/s41746-020-00323-1>
96. Sadilek, A.; Liu, L.; Nguyen, D.; Kamruzzaman, M.; Serghiou, S.; Rader, B.; Ingerman, A.; Mellem, S.; Kairouz, P.; Nsoesie, E.O.; MacFarlane, J.; Vullikanti, A.; Marathe, M.; Eastham, P.; Brownstein, J.S.; Aguera y Arcas, B.; Howell, M.D.; Hernandez, J. Privacy-first health research with federated learning. *npj Digit. Med.* 2021, 4, 132. <https://doi.org/10.1038/s41746-021-00489-2>
97. De Novo Classification Request. Available online: <https://www.fda.gov/medical-devices/premarket-submissions-selecting-and-preparing-correct-submission/de-novo-classification-request> (accessed on 24 March 2026).
98. General Wellness: Policy for Low Risk Devices: Guidance for Industry and Food and Drug Administration Staff. Available online: <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/general-wellness-policy-low-risk-devices> (accessed on 24 March 2026).

99. Marketing Submission Recommendations for a Predetermined Change Control Plan for Artificial Intelligence-Enabled Device Software Functions: Guidance for Industry and Food and Drug Administration Staff. Available online: <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/marketing-submission-recommendations-predetermined-change-control-plan-artificial-intelligence> (accessed on 24 March 2026).
100. Rattray, N.J.W.; Hamrang, Z.; Trivedi, D.K.; Goodacre, R.; Fowler, S.J. Taking your breath away: Metabolomics breathes life in to personalized medicine. *Trends Biotechnol.* 2014, 32, 538–548. <https://doi.org/10.1016/j.tibtech.2014.08.003>

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.